

The Origin of Symbols in the Brain

Christian Balkenius, Peter Gärdenfors and Lars Hall

*Lund University Cognitive Science
Kungshuset, Lundagård
S-222 22 Lund, Sweden*

*Christian.Balkenius@lucs.lu.se,
Peter.Gardenfors@lucs.lu.se,
Lars.Hall@lucs.lu.se*

0. Introduction

Deacon's (1997) book is an interesting attempt to explain the critical aspects of the evolution of language as the learning of symbolic relationships. Deacon blurs the traditional distinction between syntax and semantics by arguing that the meaning of symbols is primarily determined via the combinatorial relations between symbols, and only secondarily via an indexical relation between a symbol and a referent (Deacon 1997, Ch. 3). However, this account of how acquisition of symbols involves multiple hierarchies of associative learning has proved rather difficult to understand (Hurford 1997), and even more difficult to incorporate into an explicit representational model. In this article, we want to use Deacon's theory as a platform for a more elaborated and precise model of symbol learning.

Our model will be presented in rough phylogenetic order, and will contain only those cognitive elements that are minimally required for the learning of symbols. These mechanisms have evolved for other purposes than symbolic processing, but taken together they form a substrate for the emergence of a symbolic ability. The purpose of this article is therefore to show that given the coevolutionary dynamics of Baldwinian evolution (i.e. that behavioral adaptation tends to precede biological/cognitive change, Deacon 1997, p345, see also Laland, Odling-Smee & Feldman, to appear), and some rather conservative assumptions about the cognitive capacities preceding the evolution of language, the end state of modern human language ability can be reached.

These capacities can be divided into three groups: (1) general learning abilities that are present in many animal species; (2) communicative capacities, some of

which are present in apes and some of which coevolve with language; and (3) learning of symbols in the form of symmetrical signs, i.e. relations between vocal sounds (or gestures) and representations, and also relations between such signs.

1. General Learning Abilities

1.1. *Classical Conditioning* is the ability to form associations from sensory representations to events that are rewarding or punishing and it critically depends on the amygdala (Rolls, 1995). This is an ability that is present even in very primitive organisms.

1.2. *Internal Inhibition* reverses the influence of conditioning when expected rewards or punishments are omitted. This is called extinction and is controlled by the orbital prefrontal cortex (Rolls, 1995).

1.3. *Expectation Learning*. More advanced organisms can associate more freely between sensory representations to predict the occurrence of an event based on another event also in cases when they are not connected with any innate needs (Mackintosh, 1983). However, these associations may still be limited and are not necessarily symmetrical.

1.4. *Context Learning*. An important step forward is the evolution of the hippocampus which makes contextual representations possible. Unlike simple associative learning that deals with the sensory representations in isolation, context learning is concerned with relations between representations. Initially, these representations are mainly spatial (Eichenbaum et al. 1999). The connections between objects and locations are mostly arbitrary which makes locations efficient cues for recall of objects (Balkenius

& Morén 2000). In neural network parlance, they make the representations of specific objects more orthogonal.

1.5. *Contextual Discrimination*. One important role played by contextual representations appears experimentally in omission training (Schmajuk & Holland 1998). While the initial learning is mainly context independent, extinction through omission of reward depends almost entirely on the context (Bouton 1991). This mechanism learns to suppress previously learnt associations when certain relations among representations are present. The prefrontal cortex has been implicated in this type of inhibitory learning (Fuster 1997, Rolls 1995) which has evolved from the more primitive extinction mechanism described above.

1.6. *Expectation Matching*. The above mechanism is driven by the omission of reward but many animals are also able to learn by expectation matching independent of any reward or punishment. When expectations are not met, a contextual inhibition develops for the incorrect expectations. This matching is one of the main functions of the septo-hippocampal system (Gray 1995).

2. Communication

The following capacities are connected with the evolution of communication. As Oliphant (to appear) makes clear, the difficulties inherent in moving from an innate to a learned system of communication lies not primarily in how to learn to associate a sound (or sign) with an object, but how to observe the correct sound-object meaning pairs.

2.1. *Joint attention*. By exploiting gaze contact, head-direction, body-posture, etc., individuals can establish joint attention to an object or event (Langton, Watt & Bruce 2000). Joint attention makes communication by ostension possible since one can then show that one wants to show something (Gomez 1998). Comparative evidence suggests that chimpanzees in the wild engage in such ostensive behavior (Leavens & Hopkins 1998).

2.2. *Vocalisation/miming*. If ostensive communication is combined with vocalisation or a miming gesture, signs for different kinds of objects can be established within a community. This sets up selection pressure for increased cortical control of the hands and facial muscles, and enables further enrichment of the system (Deacon 1997, ch. 8). The extent to which imitative and/or theory-of-mind-like abilities are required for, or a

consequence of, the development and maintenance of such a conventional system, remains unclear (Dennett 1996, Burling 1999, Schaal 1999).

2.3. *Spontaneous Generation*. The next step on the road to full symbolization is spontaneous production, or babbling. Most importantly, this was not a matter of one clever hominid suddenly opening its mouth and letting the world know what it thought. As Dennett (1997) explains, the habit of adding a ‘soundtrack’ to one’s activities must be driven into place *before* the words themselves are understood. “For a word to serve as a useful, manipulable label in the refinement of the resources of a brain, it must be a ready enhancer of sought-for associations that are already to some extent laid down in the system” (Dennett 1997, p 347).

3. Symbolicity

3.1 *Symmetrical Signs*. In a phylogenetic perspective, babbling most probably coevolved with the advent of crude indexical language use, but as a simple example of automatic, instinctual resource enhancement it had far-reaching cognitive consequences. In our model it lays the foundation for the development of *symmetrical* associations between a representation, its corresponding sound pattern and the patterns of muscle activity that produces it. It seems unlikely that nonhuman primates can acquire such symmetrical associations (Lowe & Dugdale 2000). We want to propose that this paves the way for an emerging Saussurian sign (de Saussure 1972). As a consequence, this opens up the all important class of *implicit negative evidence* that modern language learners rely upon.

3.2. *Words as Orthogonal Code*. Since words are arbitrary, they are in a similar position as a place representation to enhance storage of a sensory representation (Clark & Thornton 1997). These orthogonal codes make it possible for signs to enter into indexical associative relations (Deacon 1997) as well as contextual representations which form the basis for symbolic relations.

3.3. *Symbolic Discrimination*. The final mechanism required for symbolic learning is a generalization of the earlier context system. Instead of working on spatial locations, it now operates on symmetrical signs, where sound patterns take on the role previously played by spatial locations.

It is well known that the prefrontal cortex is involved in the representation of sequential structures and it is well fitted for the inhibition of "ungrammatical" sequences or combinations. We suggest that by being explicitly unrewarded or having expectations unconfirmed, combinations of tokens can take on combinatorial properties. This implies that the learning of symbols proceeds through the acquisition of context-sensitive rules that code for incorrect combinations. The context, in this case, mainly consists of other symbols. As a consequence, explicit negative feedback is not necessary for learning symbolic relations, in contrast to what has been claimed by Chomsky and his followers (cf. Regier 1996, Rhode and Plaut 1999, Schoenemann 1999, Morris, Cottrell & Elman, in press).

Our model also fits well with recent empirical studies of language learning in children. The constructive grammar accounts by Carpenter, Nagell, Tomasello (1998), Gillette, et al. (1999) and Tomasello (2000) all show that children tend to build their communicative competence in an incremental fashion by rote learning of pragmatically relevant sentences, and then eventually *discover* the common structure behind them. This process is described by Tomasello (2000) as children "not just combining words or isolated linguistic categories, they are combining pre-compiled linguistic constructions of various shapes, size, and level of abstractness." (p 243). As Tomasello admits, the *mechanisms* underlying this capacity (what he calls 'structure mapping' and 'analogy making') are still underspecified.

Our model attempts to address this problem and specify the mechanisms behind symbolic learning. Several parts of the model have already been tested in computer simulations (Balkenius & Morén 1999, 2000) and in the future we intend to perform simulations of the complete model.

References

Balkenius, C. & Morén, J. (1999) Dynamics of a classical conditioning model. *Autonomous Robots*, 1(7).

Balkenius, C. & Morén, J. (2000). A computational model of context processing, submitted to SAB 2000.

Bouton, M. E. (1991). Context and retrieval in extinction and in other examples of interference in

simple associative learning. In L. Dachowski & C. F. Flaherty (Eds.) *Current Topics in Animal Learning: Brain, Emotion and Cognition*, Erlbaum: Hillsdale, N. J.

Burling, R. (1999). The Cognitive Prerequisites for Language. *Psychology* 10(032).

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*. 255:63:4.

Clark, A., & Thornton, C. (1997). Trading spaces: computation, representation and the limits of uninformed learning. *Behavioral and Brain Sciences*. 20. 57-90.

De Saussure, F. (1972). *Course in General Linguistics*. London: Duckworth.

Deacon, T. (1997). *The Symbolic Species*. London: Penguin Books.

Dennett, D. C. (1996). *Kinds of Minds: Toward an Understanding of Consciousness*. Basic Books.

Dennett, D. C. (1997). How to do other things with words. *Philosophy: the journal of the Royal institute of philosophy*, Iss. 42.

Eichenbaum, H., Dudchenko, P., Wood, E., Shapiro, M., & Tanila, H. (1999) The Hippocampus, Memory, and Place Cells: Is It Spatial Memory or a Memory Space? *Neuron*, Vol. 23, pp. 209–226.

Fuster, J. M. (1997). *The Prefrontal Cortex: Anatomy, Physiology, and Neuropsychology of the Frontal Lobe*. Philadelphia: Lippincott-Raven.

Gillette, J., et al. (1999). Human simulations of vocabulary learning. *Cognition*. 73. 135-176.

Gomez, J-C., (1998). Some thoughts about the evolution of LADS: With Special reference to TOM & SAM. In Carruthers, P., & Boucher, J. (Eds.). *Language & Thought: Interdisciplinary themes*. Cambridge: Cambridge University Press.

Gray, J. A. (1995). A model of the limbic system and basal ganglia: Application to anxiety and schizophrenia. In Gazzaniga, M. S. (Ed.) *The cognitive neurosciences* (pp. 1165-1176). Cambridge, MA: MIT Press.

Hurford, J. (1997). Review of Deacon: *The Symbolic Species*, Allen Lane, The Penguin Press, London. *The Times Literary Supplement*, October 23rd, 1998, p.34.

Laland, K. N., Odling-Smee, J., & Feldman, M. W. (to appear) Niche Construction, Biological

Evolution and Cultural Change. *Behavioral and Brain Sciences*.

- Leavens, D. A., & Hopkins, W. D. (1998). Intentional communication by chimpanzees: A cross-sectional study of the use of referential gestures. *Developmental Psychology*. 34:5. 813-822.
- Lowe, F., Dugdale, N. (2000). Testing for symmetry in the conditional discrimination of language-trained chimpanzees. *Journal of the experimental analysis of behavior*. 73. 5-22
- Mackintosh, N. J. (1983). *Conditioning and associative learning*. Oxford: Oxford University Press.
- Morris W.C., Cottrell, G.W., & Elman, J.L. (in press). A connectionist simulation of the empirical acquisition of grammatical relations. In Stefan Wermter & Run Sun (Eds.) *Hybrid Neural Symbolic Integration*. Springer Verlag
- Oliphant, M. (to appear). The learning barrier: Moving from innate to learned systems of communication. *Adaptive Behavior*.
- Regier, T. (1996). *The Human Semantic Potential: Spatial Language and Constrained Connectionism*. Cambridge, MA: The MIT Press.
- Rhode, D. L. T. & Plaut, D. C. (1999). Language acquisition in the absence of explicit negative evidence: how important is starting small? *Cognition* 72. 67-109.
- Rolls, E. T. (1995). A theory of emotion and consciousness, and its application to understanding the neural basis of emotion. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 1091-1106). Cambridge, MA: MIT Press.
- Schaal, S. (1999). Is imitation learning the route to humanoid robots. *Trends in Cognitive Sciences*. 3:6. 233-242.
- Schmajuk, N. A. & Holland, P. C. (Eds.). (1998). *Occasion Setting: Associative Learning and Cognition in Animals*. Washington, DC: American Psychological Association.
- Schoenemann, P. T (1999). Syntax as an emergent characteristic of the evolution of semantic complexity. *Minds and Machines*. 9:3. 309-346.
- Langton, S, R. H., Watt, R. J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences*. 4:2. 50-59
- Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition*. 74. 209-253.