Pere Alberch's Developmental Morphospaces and the Evolution of Cognition

Sergio Balari

Departament de Filologia Catalana and Centre
de Lingüística Teòrica, Facultat de Lletres
Universitat Autònoma de Barcelona
Bellaterra (Barcelona), Spain
Sergi.Balari@uab.cat

Guillermo Lorenzo

Departamento de Filología Española, Facultad de Filología
Campus El Milán, Universidad de Oviedo, Oviedo, Spain
glorenzo@uniovi.es

Abstract

In this article we argue for an extension of Pere Alberch's notion of developmental morphospace into the realm of cognition and introduce the notion of cognitive phenotype as a new tool for the evolutionary and developmental study of cognitive abilities.

During the 1980s and until the time of his death in the late 1990s, the Catalan biologist Pere Alberch developed a theoretical model of morphological evolution. Alberch's model can in many ways be considered as a direct precursor of contemporary evolutionary developmental biology (EvoDevo) thinking (for details and an historical overview see García-Azkonobieta 2005; Etxeberria and Nuño de la Rosa in press). As we expound in detail in the next section, one of the characteristic features of Alberch's model is the idea that "morphological variation at the macroscopic level is *not* continuously distributed. Rather it is distributed among a finite set of discrete states" (Alberch 1980: 654); moreover, it is also nonchaotic, in the sense that "morphologies are *not* generated in a continuous and random fashion" (p. 654). This idea goes back at least to the work of 19th-century transcendental morphologists like É. Geoffroy Saint-Hilaire (1818–1822), I. Geoffroy Saint-Hilaire (1832–1837), and Owen (1848) as well as to that of some post-Darwinian naturalists like Bateson (1894); and it was directly implemented by Alberch in his notion of *parametric space*. For the purposes of these introductory remarks, suffice it to say that a parametric space is a theoretical representation of the finite collection of possible phenotypes as a function of a set of developmental constraints or *morphogenetic parameters* that is also finite. In Alberch's view, such spaces are to be seen not as static but as dynamic entities, where possible transitions from one phenotype to another may be defined as perturbations over one or more of these morphogenetic parameters, thus making some transitions more probable than others and, concomitantly, some phenotypes more probable than others.

Although Alberch's own experimental work concentrated on the study of the development of some very specific morphological patterns (like the number of digits per limb in different organisms), his model was intended to explain any case of morphological variation, including that of the nervous system. Indeed, Alberch's theories are particularly well suited to any attempt at explaining brain development and evolution, since, as most contemporary work in this field suggests, these processes are clearly nonrandom (Striedter 2005, 2006; see also Ebbesson 1980, 1984 and Edelman 1987 for some earlier proposals) and constrained to a finite number of possibilities (Hofman 2001). Now, inasmuch as the "mind" is just what "the brain does" (Searle 1985), this line of research should prove also very useful in the enterprise of broadening some EvoDevo concepts to aid in explaining the evolution of cognition, following an internalist trend that is gaining visibility (e.g., Griffiths and Stotz 2000; Amundson 2006; Griffiths 2007) as an alternative to the externalist positioning of evolutionary psychology (Barkow et al. 1992; Pinker 1997; Plotkin 1997; Buss 2005, 2007).

A sizeable literature is already devoted to explaining the evolution of patterns of neuroanatomical organization based on the idea that readjustments in the development of nervous systems are the main evolutionary source of diversity in this domain (see Parker et al. 2000, Falk and Gibson 2001, and Minugh-Purvis and McNamara 2001 on primate brain evolution). While this is an important factor linking standard EvoDevo applications to the study of cognition, we argue that it cannot be seen as the ultimate move in explaining the origins of complex forms of cognition. This is so because perturbations leading to minor quantitatively measurable changes in brain organization can nevertheless correlate with true qualitative gaps in cognitive skills. For this reason we defend the position that the adoption of an abstract and function-oriented viewpoint is an important complementary strategy to that of a purely anatomical perspective when dealing with the evolutionary study of cognition (on the issue of the levels of brain description, see Chomsky 1980; Marr 1982), a viewpoint that can nicely be accommodated within EvoDevo endeavors (see Gibson 1990, 2004, Parker and McKinney 1999, and Langer 2000 for some precedents of the idea in various cognitive areas). Actually, Alberch himself explored these avenues in his notebooks, where we find some reflections on the pattern-generating capacities of the human brain, invoking the computer metaphor (Laura Nuño de la Rosa, personal communication, 11 March 2008).

This article presents an extension of Alberch's notion of parametric space into the realm of cognition, suggesting that a parametric space of cognitive or computational phenotypes exists, paralleling the space of possible brain morphologies, such that each computational phenotype correlates with one or more brain phenotypes of which it is a computational characterization. The space of possible computational phenotypes is defined on the basis of the Chomsky hierarchy of computational levels of complexity (Chomsky 1956, 1959), for which a precise mathematical characterization is available in terms of computational regimes and resources. Thus, we have at our disposal a theoretical tool to provide a functional description of the necessary resources to perform some cognitive task, capable of complementing and possibly directing research at the level of brain morphology. At the end of this article we offer a concrete example of this application to the case of the human faculty of language.

Morphological Evolution and Related Concepts

Alberch (in particular 1980, 1989, and 1991) originated the concept of *morphological evolution* aimed at explaining the phylogeny of organic designs on the very same basis as modern EvoDevo thinking: changes in the parameters underlying the development of organisms are the main source of the evolutionary processes capable of introducing novelties into nature. He also contended that developmental systems possess the properties of complex dynamic systems (Thelen and Smith 1994; Kelso 1995), in which intricate interactions between genetic

and nongenetic factors relate nonlinearly with morphological outputs. Such nonlinearity basically means that once certain critical thresholds are reached, small perturbations of any of the morphogenetic parameters of the system are capable of bringing about wide-range consequences in development and, ultimately, in the evolution of entire lineages of organisms. In this connection, De Renzi et al. (1999: 625–626) have pointed out that Alberch anticipated the application of contemporary complex systems theory to organic development through his conception of developmental systems as dynamic nonlinear systems (see, in particular, Oster and Alberch 1982; see also Kauffman 1993, which includes a note of personal acknowledgment and several references to Alberch's work). As a matter of fact, at the time of his death Alberch was working on the project of formalizing his ideas in terms of chaos and complexity theory in a book titled *An Introduction to Chaos Theory and Complexity With Special Emphasis on Biological Sciences*, which he left unfinished (L. Nuño de la Rosa, personal communication, 11 March 2008). Alberch also shared with most EvoDevo theoreticians the conviction that the most common effects of those perturbations have to do with the timing and/or rates of growth of the developing structures, which as a result can diverge radically from closely related ones (Gould 1977; Smith 2001, 2002). *Heterochrony* was thus the main mechanism referred to by Alberch to explain the origins of most evolutionary novelties (Alberch et al. 1979; Alberch and Alberch 1981; see also Alberch 1985 and Alberch and Blanco 1996 for some critical comments on standard uses of the concept).

A more distinctive aspect of Alberch's model is the idea that systems of interactions underlying developmental sequences are rather stable and that changes in these sequences are mostly due to modifications in the values of one or another of the morphogenetic parameters of the system (Alberch 1989, 1991). In the study of complex dynamic systems, the concept of *control parameter* refers to the systemic component whose perturbations correlate with the emergence of new morphologies (a new pattern in the surface of a chemical solution, a new embryological state, a new form of behavior, and so on; Thelen and Smith 1994: 63–64). A control parameter is not, however, a central agent in the causation of phenotypic variation in that the effects of its perturbations do not immediately reflect on morphological outputs but on the other morphogenetic parameters (Thelen and Smith 1994: 112; Kelso 1995: 7). Within this model, development (even the development of novel forms) is always a function of the system as a whole. The idea of control parameter basically introduces the possibility of pinpointing a single parameter of the system as the starting point of the chain reaction leading to new morphologies.

For our own purposes, however, the most relevant aspect of Alberch's proposals is the contention that developmental systems foreshadow the scope of their attainable phenotypes,
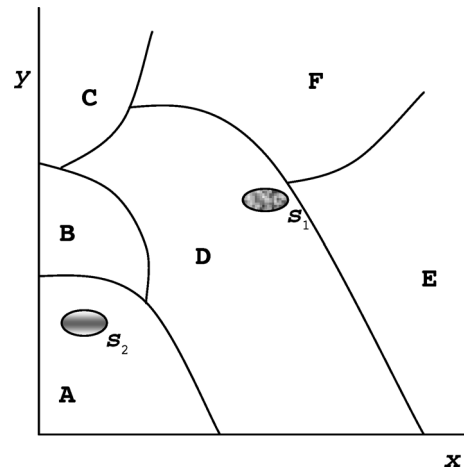
Figure 1.
An idealized parametric space defined by two abstract parameters x and y.

as well as the trajectories leading from one phenotypic state to another. It is the concept of *parametric space* that in Alberch's framework is in charge of theoretically representing the finite and discrete set of the possible outcomes of any developmental system (Alberch 1989, 1991). The main properties of parametric spaces are summed up in the following paragraphs (see Figure 1 as a point of reference).

A parametric space is a finite set of discrete phenotypes. The discontinuous character of phenotypic variation is captured in Figure 1 by the spaces labeled with capital letters. Each phenotype has a characteristic probability of coming into being, represented in Figure 1 by the extension that it occupies (D is thus the most probable phenotype, while B is the most improbable one). Moreover, each phenotype is also characterized by the relative probability of transforming itself into one or another of the neighboring phenotypes. In Figure 1 this aspect is represented by the extension of the line separating different phenotypes (A has a strong probability of turning into D, a low probability of turning into B, and no probability at all of turning into any of the remaining phenotypes).

Species are represented in Figure 1 by means of the oval items (s_1 and s_2). From a population point of view, the model incorporates the following contentions: Firstly, every species falls upon one or another phenotype (s_1 belongs to phenotype D, whereas s_2 fits in phenotype A). Secondly, the morphological stability of a species is a function of both (1) the probability of its phenotype (s_1 is, in principle, a more stable population than s_2) and (2) its proximity to a point of bifurcation to other phenotypes (s_1 is thus a rather unstable population within its phenotype, given its vicinity to the bifurcation leading to E and F). Finally, the proximity to a point of bifurcation as well as the relative propensity of its own phenotype to transform into one or another phenotype puts a certain population at the edge of undergoing a radical morphological reorganization (s_1, for instance, has a high propensity of acquiring the properties of

phenotype E). *Bifurcation*, a point within a parametric space in which a minimal perturbation is capable of bringing about qualitatively new morphologies, is thus another key concept of Alberch's framework (Oster and Alberch 1982).

From an evolutionary perspective, the idea that the geometry of parametric spaces works as a very strong constraining force, capable of counteracting that of natural selection, is an important corollary of this model. As explicitly stated by Alberch, this does not discredit the role of natural selection in the evolution of organic designs; instead, it redefines it as a filtering mechanism rather than as a creative force (Alberch 1980: 664, 1989: 46–48, 1991: 16; Oster and Alberch 1982: 455; see also Goodwin 1994: 143; Wagensberg 2004: 125).

As we argue in the next section, Alberch's model of morphological evolution can be easily extended to explain the evolution of complex cognitive functions.

Computational Phenotypes

As we noted in the introductory section, a detailed characterization of all the defining properties of a specific brain phenotype is not always possible, but we have at our disposal a powerful theoretical tool that makes it possible to identify its main features at the computational level and to elaborate concrete proposals about what structures and neural organizations could be associated with these properties. Our concrete proposal is to assume the existence of a parametric space of cognitive phenotypes containing a minimum of four phenotypes, in direct correspondence with the four levels of computational complexity in the Chomsky hierarchy, namely, type 3 (regular), type 2 (context-free), type 1 (context-sensitive), and type 0 (unrestricted). For ease of exposition, we limit our space to four phenotypes, but it should be taken into account that after Chomsky's original formulation, the complexity space has been enriched to accommodate further levels (Chomsky 1963; Aho 1968; Hopcroft and Ullman 1979: ch. 14; Joshi 1985; Joshi et al. 1991; Vijay-Shanker and Weir 1994). Since nothing in this exposition hinges on the exact number of phenotypes, we stick to a space of four, which, following the numbering in the hierarchy, we will label CP_3, CP_2, CP_1, and CP_0, respectively. Just as in Figure 1 above, these would be arranged in a parametric space, representing the probability of each phenotype and the probability of jumping from one to the others. The only exception here will be CP_0 because we assume that the probability of transition from any other phenotype to CP_0 is actually 0. In this sense, our parametric space is a *theoretical morphospace* in the sense of Rasskin-Gutman (2005: 214–215), including possible (both actual and potential) and impossible phenotypes, as opposed to an *empirical morphospace*, excluding the latter. Its exclusion is justified by natural limitations on the parametric factors involved, as it seems reasonable to do with the case of CP_0, given the proper-

ties associated with unrestricted systems (see Chomsky 1959 for discussion of this point).

What is important here is that whatever the specific morphological properties of some brain phenotype, but provided that we are able to associate it with one computational phenotype, we have a precise computational characterization for it. In fact, we know that CP_3 has a computational power equivalent to that of a finite-state automaton, CP_2 is equivalent to a pushdown automaton, CP_1 to a linear-bounded automaton, and, finally, CP_0 is equivalent to a Turing machine.

Space limitations prevent us from giving a precise formal description of each of these systems, but they can be safely understood as kinds of abstract computers working with specific processing regimes and with different computational resources. This sketchy characterization is nevertheless sufficient to derive an interesting conclusion with potentially relevant repercussions in the study of the evolution of cognitive abilities. In fact, as a number of recent mathematical results demonstrate (Weir 1992, 1994; Joshi and Schabes 1997), the only difference between the different levels of complexity lies in a single computational resource, viz., memory, meaning that the progression up the scale of complexity is simply a function of the changes introduced in the memory system, with no other modification of any fundamental property of the computational system being necessary. We are, therefore, in a position not just of being able to characterize precisely our computational phenotypes but also of being able to determine the kinds of alterations of the developmental system that are necessary to facilitate the transition from one phenotype to the other and, concomitantly, eventually to characterize the alterations at the morphological level necessary to implement one or another computational regime. The obvious corollary of this is that the evolution of higher cognitive skills may be understood as a historical process in which newer and more sophisticated cognitive skills have emerged not as the result of radical rearrangements in the computational regimes of nervous systems but just as the result of brains having at their disposal larger and more sophisticated systems of memory.

To conclude this section, note, however, that the parametric space of cognitive computational phenotypes possesses the very same properties as the morphospaces that Alberch proposed. That is, it shows a noncontinuous distribution of phenotypic variation, with discrete and easily identifiable states. Thus, the transition from one state to another is, in fact, a "jump" that is only made possible once a specific critical point is attained (possibly as a result of the accumulation of small gradual changes). In this connection, it is nevertheless important to take into account that the fact that there is discontinuity in *variation* does not contradict the idea of gradual change at the level of *processes* (as already pointed out by Bateson 1894: 13–17). This is, in fact, a fundamental feature of the concept of "critical-point emergence" to which we shall appeal here

(for details, see Reid 2007: ch. 8). According to our proposal, then, the evolution of what we call "the computational mind" would have consisted in a historical process in which complex interactions among genetic and epigenetic factors during the individual developmental process of the nervous system would have given rise to qualitatively differentiated phenotypes via a system of "critical-point" emergent processes. Such phenotypes would have been able to act as the material support for richer and varied computational regimes, as we will see in the next section.

Some Examples: Language, Birdsong, and Knots

Language is often seen "as one of the chief distinctions between man and the lower animals," as Darwin (1871) put it. Language is complex—but not *too* complex. In fact, as several years of investigations of the structural complexities of human language have demonstrated, the computational resources necessary to process the most structurally complex human utterance possible appear to be only slightly beyond the capacity of a context-free system (our CP_2) and within the complexity space of what Joshi (1985) termed *mild context-sensitivity* (i.e., within the lower area of our CP_1) but in any case, never beyond what are known as *indexed systems* (see Aho 1968, Pullum 1986, and Radzinski 1991 for critical discussion).

That said, and given our previous considerations on the role of memory in connection to greater computational complexity, we can identify the major evolutionary event that gave rise to modern human linguistic abilities—or, to be precise, the computational system underlying the faculty of language in the narrow sense of Hauser et al. (2002)—as one that enhanced the working memory of a preexisting computational system in the direction of context-sensitivity (CP_1). In fact, following Lieberman (2006), there is enough evidence for such an anatomical precursor in the form of the cortico-subcortico-cortical circuits associated with the regulation of different aspects of mobility, cognition, and emotion whose main subcortical component is located in the basal ganglia (see also Cummings 1993). In this model the role of the cortical areas is that of working memory space, with the basal ganglia acting as a cognitive pattern generator to make up what Lieberman characterizes as an iterative sequencing machine. From an evolutionary point of view the basal ganglia, as opposed to the cortex, have been described as highly conservative structures among amniotes (Reiner et al. 1984), such that no greater interspecific differences are expected at the level of the subcortical structure acting as the material support of the system's pattern-generation procedure, leaving all room for innovation to those cortical areas that, as pointed out above, provide working space memory to the system.

This is all very sketchy (but see Balari and Lorenzo 2009 for a concrete and detailed proposal of the emergence of lan-

guage along these lines); however, we believe it is illustrative enough of the applicability of our notion of computational phenotype in the construction of scientific hypotheses about the evolution and development of higher cognitive functions in humans and other species. Indeed, we believe that one of the virtues of our proposal is that it can easily be applied to the cognitive abilities of other species.

An important claim of our proposal is that the computational core engine of the vertebrate brain is located in the striatum, with the basal ganglia functioning as an iterative sequencing machine. The participation of the basal ganglia in the performance of linguistic tasks appears to be unquestionable (Tettamanti et al. 2005), and more interestingly, they also appear to play a crucial role in singing behavior in birds, with striking neurochemical similarities (Ding and Perkel 2002; Gale and Perkel 2005; Sasaki et al. 2006; Cornil et al. 2008; Huang and Hessler 2008). In addition, at the genetic level, the human variety of FOXP2 is expressed, among many other loci, in the basal ganglia (see Benítez Burraco 2008, 2009 for a detailed overview), as it is the case of its avian homolog (Rochefort et al. 2007). All this evidence suggests that birdsong and language share a common neural substrate and, consequently, that the observed differences between the two cognitive abilities must be due to other properties of the computational systems associated with them. One of these differences, according to the proposal developed here, must lie at the level of the resources available to perform these tasks. In fact, we already noted that, computationally, language would correspond to our CP_1, whereas birdsong typically shows the properties of CP_3 (Todt and Hultsch 1998; Okanoya 2002), and in any case, it never appears to go beyond those of CP_2 (Gentner et al. 2006). Clearly, however, this cannot be the only difference because language is also a symbolic system that precisely because it has specific computational properties, also possesses the property of discrete infinity. This is important because, as pointed out by Lorenzo (2006), it is perfectly plausible to assume that in the animal kingdom certain behaviors are observed that suggest the presence of a symbolic system (or the rudiments thereof) or the use of very complex recursive patterns (which does not imply the simultaneous presence of complex recursion and symbolism). To extend and perfect working memory space is, then, a necessary condition for the emergence of language but not a sufficient one. This dissolves a potential objection to our notion of computational phenotype, viz., that it may not necessarily be the case that this phenotype (and its corresponding morphological phenotype) correlates with the presence of language. In fact, this observation is correct, and our CP_1 might well correspond to nonlinguistic "minds" that would nevertheless be capable of producing complex recursive patterns within other areas of cognition such as motor sequences or melodic sequences. Remember that, according to Lieberman's (2006) model, the

basal ganglia comprise a sequence of cognitive patterns, but cognitive patterns may be of many different sorts, and the sequencer, just because of connectivity and working space limitations, might have access to only a single type of pattern or to a reduced collection of them (e.g., motor patterns and melodic patterns) but not others (symbolic patterns, assuming these are even available). In this connection, it is particularly interesting to pay attention to, for instance, the construction behavior of weaver birds, which are capable of making some types of knots (Hansell 2000, 2005). Making a knot requires the application of an operation over a part of the constructed figure and keeping it in active memory until the moment at which the operation completing the figure is executed, which requires complex computations perhaps within the power of CP_1 (Camps and Uriagereka 2006).

We need to emphasize that the computational system to which we refer must not be identified with the faculty of language in the broadest sense of the term (Hauser et al. 2002). Moreover, since our model does not presuppose that this system is language-specific (or human-specific, for that matter), it may perfectly be understood as one of the "infrastructural" elements (Oller 2005) rather than as a "component" of language per se. In this sense, this system would correspond to a level of cognitive analysis comparable to other "facilitating" elements of the linguistic capacity, such as shared attention (Carpenter et al. 1998; Tomasello and Farrar 1986) or vocal dexterity (Ploog 2002; Oller 2005). We wish to point out in this connection that our model offers the possibility of explaining these infrastructural elements of the language faculty, invoking the very same factors from which we derive the complexity of its computational regime. On the one hand, it is known that controlled attention, of which shared attention can be assumed to be just a particular kind, may be enhanced as a function of the working-memory resources associated with it (see Desimone 1996; Awh and Jonides 2001; Curtis and D'Esposito 2003; see also the overview in Klingberg 2009). On the other hand, vocal dexterity is explained by Striedter (2005: 324–326) as an effect of the "connectional invasion" of the medulla and the spinal cord by cortical neurons as a by-product of cortical growth.

We believe that these considerations have important methodological implications, especially with respect to the application of the comparative method when seeking precursors of language or of some of its constituent properties like complex recursion. Indeed, one of the most direct consequences is that formal grammar and automata theory may prove to be an extremely useful tool at the time of assessing the abilities and capabilities shown by other animal species, a point also argued for by many other researchers (see, e.g., Hauser et al. 2002; Fitch and Hauser 2004; O'Donnell et al. 2005; Camps and Uriagereka 2006; Gentner et al. 2006; Uriagereka 2008). However, as we have just seen, there is no guarantee that animal species different from ours show the very same abilities within the very same areas of cognition, which is perfectly compatible with our suggestion that a computational phenotype may correlate with many different morphological phenotypes corresponding to different cognitive abilities. Thus, our proposal is less prone to be the target of the kind of criticisms (Perruchet and Rey 2005; Pullum and Rogers 2006; Rogers and Pullum in press) that were targeted at work along similar lines. Finally, just as our proposal is compatible with the possibility of nonlinguistic expert minds associated with our CP_1—which would, for example, explain the absence of any clear evidence of symbolic behavior in the Neandertal archaeological record (Balari et al. 2008; Benítez Burraco et al. 2008)—it is also compatible with the idea that the computational regime subserving the human language faculty also subserves other, nonlinguistic aspects of human mentality, and, consequently that the narrow faculty of language of Hauser et al. (2002) is, in fact, not specific to language.

Conclusion

Our work may be inscribed within a line of research characterized by the appeal to both internalist principles and computational complexity. We believe that, *pace* Lewontin's (1998) pessimistic stance with respect to the evolutionary study of cognition, attempts to construct evolutionary explanations of complex cognitive abilities will eventually open new pathways to find empirical support for Darwin's (1871) contention that "nevertheless the difference in mind between man and the higher animals, great as it is, certainly is one of degree and not of kind."

Acknowledgments

References

Aho AV (1968) Indexed grammars: An extension of context-free grammars. Journal of the ACM 15: 647–671.

Alberch P (1980) Ontogenesis and morphological diversification. American Zoologist 20: 653–667.

Alberch P (1985) Problems with the interpretation of developmental sequences. Systematic Zoology 34: 46–58.

Alberch P (1989) The logic of monsters: Evidence for internal constraint in development and evolution. Geobios 12 (mémoire spécial): 21–57.

Alberch P (1991) Del gen al fenotipo: sistemas dinámicos y evolución morfológica. Revista Española de Paleontología (número extraordinario "El

estudio de la forma orgánica y sus consecuencias en Paleontología Sistemática, Paleontología y Paleontología Evolutiva"): 13–19.

Alberch P, Alberch J (1981) Heterochronic mechanisms of morphological diversification and evolutionary change in the neotropical salamander, *Bolitoglossa occidentalis* (Amphibia: Plethodontidae). Journal of Morphology 167: 249–264.

Alberch P, Blanco MJ (1996) Evolutionary patterns in ontogenetic transformation. International Journal of Developmental Biology 40: 845–858.

Alberch P, Gould SJ, Oster GF, Wake DB (1979) Size and shape in ontogeny and phylogeny. Paleobiology 5: 296–317.

Amundson RA (2006) EvoDevo as cognitive psychology. Biological Theory 1: 10–11.

Awh E, Jonides J (2001) Overlapping mechanisms of attention and spatial working memory. Trends in Cognitive Science 5: 119–126.

Balari S, Benítez Burraco A, Camps M, Longa VM, Lorenzo G, Uriagereka J (2008) ¿Homo loquens neanderthalensis? En torno a las capacidades simbólicas y lingüísticas del Neandertal. Munibe Antropologia—Arkeologia 59: 3–24.

Balari S, Lorenzo G (2009) Computational phenotypes: Where the theory of computation meets Evo-Devo. Biolinguistics 3: 2–60.

Barkow JH, Cosmides L, Tooby J, eds (1992) The Adapted Mind: Evolutionary Psychology and the Generation of Culture. Oxford: Oxford University Press.

Bateson W (1894) Materials for the Study of Variation Treated with Special Regard to Discontinuity in the Origin of Species. London: Macmillan.

Benítez Burraco A (2008) FOXP2 y la biología molecular del lenguaje: Nuevas evidencias. II. Aspectos moleculares e implicaciones para la ontogenia y la filogenia del lenguaje. Revista de Neurología 46: 351–359.

Benítez Burraco A (2009) Genes y Lenguaje: Aspectos ontogenéticos, filogenéticos y cognitivos. Barcelona: Reverté.

Benítez Burraco A, Longa VM, Lorenzo G, Uriagereka J (2008) Also sprach Neanderthalis. . . Or did she? Biolinguistics 2: 225–232.

Buss DM, ed (2005) The Handbook of Evolutionary Psychology. Hoboken, NJ: Wiley.

Buss DM (2007) Evolutionary Biology: The New Science of the Mind. Boston: Allyn and Bacon.

Camps M, Uriagereka J (2006) The Gordian knot of linguistic fossils. In: The Biolinguistic Turn: Issues on Language and Biology (Rosselló J, Martín J, eds), 34–65. Barcelona: PPU.

Carpenter M, Nagell K, Tomasello M, Butterworth G, Moore C (1998) Social cognition, joint attention, and communicative competence from 9 to 15 months of age. Monographs of the Society for Research in Child Development 63: 1–174.

Chomsky N (1956) Three models for the description of language. IRE Transactions on Information Theory 2: 113–124.

Chomsky N (1959) On certain formal properties of grammars. Information and Control 2: 137–167.

Chomsky N (1963) Formal properties of grammars. In: Handbook of Mathematical Psychology, Vol. 2 (Luce RD, Bush RR, Galanter E, eds), 323–418. New York: Wiley.

Chomsky N (1980) Rules and Representations. New York: Columbia University Press.

Cornil CA, Castelino CB, Ball GF (2008) Dopamine binds to α_2-adrenergic receptors in the song control of zebra finches (*Taeniopygia guttata*). Journal of Chemical Neuroanatomy 35: 202–215.

Cummings JL (1993) Frontal–subcortical circuits and human behavior. Archives of Neurology 50: 873–880.

Curtis CE, D'Esposito M (2003) Persistent activity in the prefrontal cortex during working memory. Trends in Cognitive Science 7: 415–423.

Darwin C (1871) The Descent of Man, and Selection in Relation to Sex. London: John Murray.

De Renzi M, Moya A, Peretó J (1999) Obituary. Evolution, development and complexity in Pere Alberch (1954–1998). Journal of Evolutionary Biology 12: 624–626.

Desimone R (1996) Neural mechanisms for visual memory and their role in attention. Proceedings of the National Academy of Science USA 93: 13494–13499.

Ding L, Perkel DJ (2002) Dopamine modulates excitability of spiny neurons in the avial basal ganglia. Journal of Neuroscience 22: 5210–5218.

Ebbesson SOE (1980) The parcellation theory and its relation to interspecific variability in brain organization, evolutionary and ontogenetic development and neuronal plasticity. Cell and Tissue Research 213: 179–212.

Ebbesson SOE (1984) Evolution and ontogeny of neural circuits. Behavioral and Brain Sciences 7: 321–366.

Edelman GM (1987) Neural Darwinism: The Theory of Neuronal Group Selection. New York: Basic Books.

Etxeberria A, Nuño de la Rosa L (in press) A world of opportunity within constraint: Pere Alberch's early Evo-Devo. In Pere Alberch: The Creative Trajectory of an Evo-Devo Biologist (Rasskin-Gutman D, De Renzi M, eds). València: Universitat de València.

Falk D, Gibson KR, eds (2001) Evolutionary Anatomy of the Primate Cerebral Cortex. Cambridge: Cambridge University Press.

Fitch WT, Hauser MD (2004) Computational constraints on syntactic processing in a nonhuman primate. Science 303: 377–380.

Gale SD, Perkel DJ (2005) Properties of dopamine release and uptake in the songbird basal ganglia. Journal of Neurophysiology 93: 1871–1879.

García-Azkonobieta T (2005) Evolución, desarrollo y (auto)organización. Un estudio sobre los principios filosóficos de la evo-devo. Doctoral Dissertation, Universidad del País Vasco, Donostia, Spain.

Gentner TQ, Fenn KM, Margoliash D, Nusbaum H (2006) Recursive syntactic pattern learning by songbirds. Nature 440: 1204–1207.

Geoffroy Saint-Hilaire É (1818–1822) Philosophie anatomique, 2 Vols. Paris: J.-B. Baillière.

Geoffroy Saint-Hilaire I (1832–1837) Histoire générale et particulière des anomalies de l'organisation chez l'homme et les animaux, 4 Vols. Paris: J.-B. Baillière.

Gibson KR (1990) New perspectives on instincts and intelligence: Brain size and the emergence of hierarchical mental construction skills. In: "Language" and Intelligence in Monkeys and Apes (Parker ST, Gibson KR, eds), 97–128. New York: Cambridge University Press.

Gibson KR (2004) Human brain evolution: Developmental perspectives. In: Biology and Knowledge Revisited: From Neurogenesis to Psychogenesis (Parker ST, Langer J, Milbrath C, eds), 123–143. Mahwah, NJ: Erlbaum.

Goodwin B (1994) How the Leopard Changed Its Spots: The Evolution of Complexity. London: Phoenix.

Gould SJ (1977) Ontogeny and Phylogeny. Cambridge, MA: Belknap Press.

Griffiths PE (2007) Evo-Devo meets the mind: Towards a developmental evolutionary psychology. In: Integrating Evolution and Development (Brandon R, Sansom R, eds), 195–226. Cambridge, MA: MIT Press.

Griffiths PE, Stotz K (2000) How the mind grows: A developmental perspective on the biology of cognition. Synthese 122: 29–51.

Hansell MH (2000) Bird Nests and Construction Behaviour. Cambridge: Cambridge University Press.

Hansell MH (2005) Animal Architecture. Oxford: Oxford University Press.

Hauser MD, Chomsky N, Fitch WT (2002) The faculty of language: What is it, who has it, and how did it evolve? Science 298: 1569–1579.

Hofman MA (2001) Brain evolution in hominids: Are we at the end of the road? In: Evolutionary Anatomy of the Primate Cerebral Cortex (Falk D, Gibson KR, eds), 113–127. Cambridge: Cambridge University Press.

Hopcroft JE, Ullman JD (1979) Introduction to Automata Theory, Languages, and Computation. Reading, MA: Addison-Wesley.

Huang YC, Hessler NA (2008) Social modulation during songbird courtship potentiates midbrain dopaminergic neurons. PloS ONE 3(10): e3281.

Joshi AK (1985) Tree adjoining grammars: How much context-sensitivity is required to provide reasonable structural descriptions? In: Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives (Dowty DR, Karttunen L, Zwicky AM, eds), 206–250. Cambridge: Cambridge University Press.

Joshi AK, Schabes Y (1997) Tree adjoining grammars. In: Handbook of Formal Languages, Vol. 3 (Rozenberg G, Salomaa A, eds), 69–126. Berlin: Springer.

Joshi AK, Vijay-Shanker K, Weir D (1991) The convergence of mildly context-sensitive grammar formalisms. In: Foundational Issues in Natural Language Processing (Sells P, Shieber S, Wasow T, eds), 31–81. Cambridge, MA: MIT Press.

Kauffman S (1993) The Origins of Order: Self-Organization and Selection in Evolution. New York: Oxford University Press.

Kelso JAS (1995) Dynamic Patterns: The Self-Organization of Brain and Behavior. Cambridge, MA: MIT Press.

Klingberg T (2009) The Overflowing Brain: Information Overload and the Limits of Working Memory. Oxford: Oxford University Press.

Langer J (2000) The heterochronic evolution of primate cognitive development. In: Biology, Brains, and Behavior: The Evolution of Human Development (Parker ST, Langer J, McKinney ML, eds), 215–235. Santa Fe, NM: School of American Research Press.

Lewontin RC (1998) The evolution of cognition: Questions we will never answer. In: An Invitation to Cognitive Science. Vol. 4: Methods, Models, and Conceptual Issues, 2nd ed. (Scarborough D, Sternberg S, eds), 107–132. 2nd ed. Cambridge, MA: MIT Press.

Lieberman P (2006) Toward an Evolutionary Biology of Language. Cambridge, MA: Harvard University Press.

Lorenzo G (2006) El vacío sexual, la tautología natural y la promesa minimalista. Madrid: Antonio Machado.

Marr D (1982) Vision. San Francisco: Freeman.

Minugh-Purvis N, McNamara KJ, eds (2001) Human Evolution Through Developmental Change. Baltimore: Johns Hopkins University Press.

O'Donnell TJ, Hauser MD, Fitch WT (2005) Using mathematical models of language experimentally. Trends in Cognitive Sciences 9: 284–289.

Okanoya K (2007) Sexual display as a syntactic vehicle: The evolution of syntax in birdsong and human language through sexual selection. In: The Transition to Language (Wray A, ed), 46–62. Oxford: Oxford University Press.

Oller DK (2005) The natural logic of communicative possibilities: Modularity and presupposition. In: Modularity: Understanding the Development and Evolution of Natural Complex Systems (Callebaut W, Rasskin-Gutman D, eds), 409–434. Cambridge, MA: MIT Press.

Oster GF, Alberch P (1982) Evolution and bifurcation of developmental programs. Evolution 36: 444–459.

Owen R (1848) On the Archetype and Homologies of the Vertebrate Skeleton. London: John Van Voorst.

Parker ST, Langer J, McKinney ML, eds (2000) Biology, Brains, and Behavior: The Evolution of Human Development. Santa Fe, NM: School of American Research Press.

Parker ST, McKinney ML (1999) Origins of Intelligence: The Evolution of Cognitive Development in Monkeys, Apes, and Humans. Baltimore: Johns Hopkins University Press.

Perruchet P, Rey A (2005) Does the mastery of center-embedded linguistic structures distinguish humans from non-human primates? Psychonomic Bulletin and Review 12: 307–313.

Pinker S (1997) How the Mind Works. New York: Norton.

Ploog D (2002) Is the neural basis of vocalization different in non-human primates and *Homo sapiens*? In: The Speciation of *Homo Sapiens* (Crow TJ, ed), 121–135. London: British Academy.

Plotkin H (1997) Evolution in Mind: An Introduction to Evolutionary Psychology. London: Alan Lane.

Pullum GK (1986) Footloose and context-free. Natural Language and Linguistic Theory 4: 409–414.

Pullum GK, Rogers J (2006) Animal pattern-learning experiments: Some mathematical background. http://www.lel.ed.ac.uk/~gpullum/MonkeyMath.pdf.

Radzinski D (1991) Chinese number-names, tree adjoining languages, and mild context-sensitivity. Computational Linguistics 17: 277–299.

Rasskin-Gutman D (2005) Modularity: Jumping forms within morphospace. In: Modularity: Understanding the Development and Evolution of Natural Complex Systems (Callebaut W, Rasskin-Gutman D, eds), 207–219. Cambridge, MA: MIT Press.

Reid RGB (2007) Biological Emergences: Evolution by Natural Experiment. Cambridge, MA: MIT Press.

Reiner A, Brauth SE, Karten HJ (1984) Evolution of the amniote basal ganglia. Trends in Neurosciences 7: 320–325.

Rochefort C, He X, Scotto-Lomassese S, Scharff C (2007) Recruitment of FOXP2-expressing neurons to Area X varies during song development. Developmental Neurobiology 67: 809–817.

Rogers J, Pullum GK (in press) Aural pattern recognition experiments and the subregular hierarchy. UCLA Working Papers in Linguistics—Proceedings of the Mathematics of Language 10.

Sasaki A, Sotnikova TD, Gainetdinov RR, Jarvis ED (2006) Social context-dependent singing-regulated dopamine. Journal of Neuroscience 26: 9010–9014.

Searle JR (1985) Minds, Brains and Science. Cambridge, MA: Harvard University Press.

Smith KK (2001) Heterochrony revisited: The evolution of developmental sequences. Biological Journal of the Linnean Society 73: 169–186.

Smith KK (2002) Sequence heterochrony and the evolution of development. Journal of Morphology 252: 82–97.

Striedter GF (2005) Principles of Brain Evolution. Sunderland, MA: Sinauer.

Striedter GF (2006) Précis of *Principles of Brain Evolution*. Behavioral and Brain Sciences 29: 1–36.

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

Tettamanti M, Moro A, Messa C, Moresco RM, Rizzo G, Carpinelli A, Matarrese M, Fazio F, Perani D (2005) Basal ganglia and language: Phonology modulates dopaminergic release. Brain Imaging 16: 397–401.

Todt D, Hultsch H (1998) How songbirds deal with large amounts of serial information: Retrieval rules suggest a hierarchical song memory. Biological Cybernetics 79: 487–500.

Tomasello M, Farrar MJ (1986) Joint attention and early language. Child Development 57: 1454–1463.

Uriagereka J (2008) Desperately evolving syntax. In: The Evolution of Language: Proceedings of the 7th International Conference (EVOLANG7) (Smith ADM, Smith K, Ferrer i Cancho R, eds), 331–337. Singapore: World Scientific.

Vijay-Shanker K, Weir D (1994) The equivalence of four extensions of context-free grammars. Mathematical Systems Theory 27: 511–546.

Wagensberg J (2004) La rebelión de las formas o cómo perseverar cuando la incertidumbre aprieta. Barcelona: Tusquets.

Weir D (1992) A geometric hierarchy beyond context-free languages. Theoretical Computer Science 104: 235–261.

Weir D (1994) Linear iterated pushdowns. Computational Intelligence 10: 431–439.