Some thoughts on biolinguistics

Alguns pensamentos sobre a biolinguística

Cedric Boeckx

Abstract: Naïve depictions of the biology of language are unable to treat the real complexity observed by biologists at all levels of analysis, and consequently they do not bring us closer to an accurate depiction of the nature of human language and the human mind. The aim of this essay is to show that if a real biolinguistics is intended to be achieved we ought to be compelled to go beyond these depictions.

Keywords: Biolinguistics. Human language. Human mind. Human genome. Biology.

Resumo: Retratos ingênuos da biologia da linguagem são incapazes de tratar a real complexidade observada por biólogos em todos os níveis de análise, e consequentemente não nos conduzam mais próximos de um relato preciso sobre a natureza da linguagem humana e da mente humana. O objetivo deste artigo é mostrar que se se propõe a realizar uma biolinguística efetiva, então deveríamos ser compelidos a ir além destes relatos.


* Portions of this contribution are reproduced from Boeckx, C., & Benitez-Burraco, A. (2014). [Biolinguistics 2.0. In Fujita, K., Fukui, N., Yusa, N., & Ike-Uchi, M. (Eds.), The design, development and evolution of human language: Biolinguistics Explorations, 7-30. Tokyo: Kaitakusha], with kind permission from my co-author, the editors of the volume, and the publisher. Thanks especially to Antonio Benítez-Burraco for his collaboration over the years, and to the students of my research group (Biolinguistics Initiative Barcelona). Preparation of this work was supported by funds from the Spanish Ministry of Economy and Competitiveness (grants FFI2013-43823-P and FFI2014-61888-EXP), as well as funds from a Marie Curie International Reintegration Grant from the European Union (PIRG-GA-2009-256413), research funds from the Fundació Bosch i Gimpera, and from the Generalitat de Catalunya (2014-SGR-200).

** Prof. Dr., Catalan Institute for Advanced Studies and Research (ICREA) & Department of Linguistics, Universitat de Barcelona, Gran Via de les Corts Catalanes, 585 Barcelona, 08007, Spain. <cedric.boeckx@gmail.com>.
1 Introduction

Reflections on biolinguistics go back to the very origin of the field, spearheaded by Noam Chomsky and Eric Lenneberg. In many respects, early practitioners were not doing something very different from what biolinguists are doing today. But in some respects—and this is the point of the reflections that follow—they were. Recent advances that bear on the practice of biolinguistics require a shift of perspective, which will be the focus of the rest of this paper.

My aim here is to briefly review what has been done up to now in the field of the biology of language, what some people are doing now, and what I think we should do in the future to gain some insight into the biological underpinnings of the human faculty for language. A central claim I’ll be making is that on the whole the research program of biolinguistics have been more controversial than they should have been because of certain (naïve) assumptions that dominated the field for several decades, on both sides of bio-linguistics: naïve views were entertained both from the linguistic side, regarding biology, and from the biological side, regarding linguistics. I will, however, focus on the linguistics side in what follows.

In general terms I will claim that, for all its success, work on bio-linguistics, especially work by linguists of a Chomskyan persuasion, where the biological orientation has been made most explicit, has resulted in a naïve biologization of language. There is ample evidence that this conceptualization of the biological foundations of language is unable to properly deal with the attested complexity observed at the genetic, neurobiological, developmental, or even evolutionary level. Because this approach is still majority in the field, hence it threatens the validity of the enterprise. It certainly generates a lot of needless controversy, and harmful dichotomies.

2 Biologizing Language

The biologization of language is surely an outstanding achievement of modern linguistics. In the fifties the Chomskyan revolution brought about a radical shift of focus in language studies, which laid the foundations of this change of paradigm. Since then, other pieces of evidence have supported this important turn in linguistics.

Firstly, the way in which language is acquired by the child, which suggests that language “learning” mechanisms are biased or constrained in specific ways. This led to the view that the systems of representation and computation underlying language are independent from other
cognitive capacities, but also that language acquisition is only possible because of some inborn grammatical knowledge.

Secondly, the fact that certain language deficits appear whenever certain brain areas are damaged or functionally impaired after a stroke, a trauma, or a tumoral process. These acquired language disorders led to the view that specific brain areas process specific language components. Aphasia studies are typical examples of this.

Thirdly, the fact that certain language deficits recurrently appear during growth whenever certain genes are mutated. These developmental language disorders led to the view that specific genes affect the development of specific brain areas that process specific components of language. Dyslexia is a typical example. This is an inherited developmental disorder which entails a deficit in phonological processing and thus reading and spelling difficulties (Ramus et al. 2003). Different brain areas are underactive or overactive in dyslexics during language processing (Shaywitz et al. 1998). Ultimately, different candidate genes for the disorder, located in different chromosomes, have been identified (see Gibson and Gruen 2008 for a review).

Fourthly, the existence of language-related components in other extant species. Allegedly, some key features of human language can be identified in the cognitive lives of other animals (including in their systems of communication). For instance, some primates are said to use and acquire rudimentary symbolic systems (Cheney and Seyfarth 1990). Similarly, bird songs are usually described in terms of strings of notes that contain motifs, strophes, and the like (Berwick et al. 2011).

Lastly, the existence of language-related components in some extinct species. Hence, different putative “fossils of language” have been identified in some hominin species: human-like speech and hearing organs in Homo heidelbergensis or Homo neanderthalensis (Martínez et al. 2004), Broca’s area in Homo erectus (Holloway, 1983), different symbolic artefacts in Neanderthals (e.g. d’Errico et al. 2003), and even the hominin variants of language related genes have been cloned (paradigmatically, FOXP2; see below.)

This last achievement nicely exemplifies how current methodologies in biology have contributed to gain a better understanding of the biological underpinnings of language and also to find real (deep) homologues of the faculty in other species.

On the whole, we are now in a position to accurately know which brain areas are active during language processing. Similarly, we have identified many of the genes that contribute to the development and the initial wiring of these neuronal devices during growth. Moreover, molecular biology tools allow us to accurately know the structure and the
function of these genes: where they are located in the genome, how they are organized, where they are expressed, what products (e.g., proteins) they encode, what functional partners they have, and ultimately, which biological function they contribute to at the brain level. Clearly, our understanding of the faculty of language is going to greatly benefit from this new source of biological data (Boeckx 2013).

Interestingly, we can explore their structural and functional properties in other organisms. In some model species we are now in a position to mutate them in order to know what phenotypic disturbances appear whenever a defective version of these genes is present in the organism. Of course, this is also of outstanding interest for the analysis of language disorders in our own species. At present, the role of all these important-for-language brain areas and genes in many different extant species has been characterized. Actually, modern paleogenomics and paleoneurology could allow us in a future to do the same in some extinct species.

**FOXP2** is perhaps the best example of what happens when this research plan is brought to completion (see Fisher and Scharff 2009, Enard 2011, or Graham and Fisher 2013 for comprehensive reviews of the topics mentioned below). The gene was identified in a family suffering from a language disorder that was inherited in a Mendelian fashion (Lai et al. 2011). Since then we have been able to sequence the whole gene including its regulatory regions, and to identify a dozen of pathogenic variants.

Moreover, we exactly know where it is expressed in the brain, which other genes **FOXP2** interacts with, which neuronal network it contributes to assemble and which computational properties this circuitry is endowed with. We know as well which brain areas are underactive or overactive during language processing in people bearing a defective version of the gene.

We have identified **FOXP2** homologues in other species and we have modified its expression pattern in some model organisms. The observed deficits quite closely resemble the speech problems exhibited by people with a defective version of the gene. On the whole, we have found that the neuronal circuitry FOXP2 contributes to is quite ancient in evolutionary terms, with clear homologues in birds and mammals. Ultimately, the Neanderthal version of the gene has even been sequenced, which has allowed us to ask new questions regarding what the gene does.

This successful research program was once expected to cast light onto some of the central problems of linguistics, including language development in the child and language knowledge in the adult, and of course, the biological nature of the human faculty for language, and its universal basis. However, **FOXP2** nicely exemplifies as well the difficulties we have to face in our task and importantly, the shortcomings of most...
current biological approaches to language. In some important sense, FOXP2 is a nightmare for everybody, not just for linguists (cp. Piattelli-Palmarini and Uriagereka 2011). Let us turn to this very briefly.

People bearing a mutated copy of the gene not only show language deficits, but other cognitive and even motor problems (Vargha-Khadem et al. 1995, Watkins et al. 2002, Padovani et al. 2010). In fact, comparative studies suggest that this gene could be more closely related to the externalization of language than to the central computational system (Kurt et al. 2012). Moreover, the brain areas in which the gene is expressed (paradigmatically, the basal ganglia) are also active during the processing of non-linguistic stimuli and/or are impaired in people affected by non-linguistic disorders (e.g. Huntington’s disease or parkinsonism) (Graybiel 1995, Gusella and MacDonald 2006). Additionally, the gene plays some important roles outside the brain (Shu et al. 2001). Ultimately, paleogenomic data, which had confirmed that Neanderthals exhibited the human variant of the gene (at least the same coding sequence) (Krause et al. 2007) are not straightforwardly compatible with other data, archaeological or paleoanthropological, which suggests that syntax (in its modern sense) is an innovation of anatomically-modern humans (bringing with it a full-fledged grammatical competence) (see Tattersall 1998, Mellars 2005, Mithen 2006 among many others).

As I suggested above, this nightmarish scenario can be seen as a plain consequence of the naïve biologization of language done till now. My main point here is that these naïve depictions of the biology of language are unable to treat the real complexity observed by biologists at all levels of analysis, and consequently they do not bring us closer to an accurate depiction of the nature of human language and the human mind. Consequently, we are compelled to go beyond them if a real biolinguistics is intended to be achieved.

3 Naïve Biolinguistics

But first, let me provide more facets of this “naïve” biolinguistics, since they are so standard, and so ‘second nature’ that they may look harmless to many.

Firstly, language features are claimed to be directly rooted in the genome. A linguistic genotype is explicitly postulated (see, e.g., Anderson and Lightfoot 1999). This linguistic genotype is further equated to a Universal Grammar. Ultimately, nativism is conflated with geneticism. This linguistic genotype is further assumed to be uniform across the species (pathologies aside) (see Lightfoot 1999, Wexler 2003, or Falcaro et al. 2008 as examples).
However, what we actually observe is quite different. Usually, genes contribute to different physiological functions in different places and times during development. At the same time, complex traits are always polygenic by nature. Moreover, language genes are polymorphic, with some pathogenic variants giving rise to language disorders, but with others affecting language growth in the normal population. Additionally, people bearing a mutated copy of a gene usually show different degrees of affectedness or can simultaneously develop different disorders (or none!). Ultimately, the mutation of two functionally related genes can give rise to different language and/or cognitive deficits or disorders in different populations (see the discussion in State 2011 on CNTNAP2 (one of FOXP2 targets) and in Benítez-Burraco 2012 on FOXP2 itself).

It is clear that genes do not work as many linguists claim. It is also clear that we need to improve our characterization of genes and particularly, their role in the context of development. A direct link between the genotype and the phenotype is not only simplistic, but biologically untenable, given the way in which genes contribute to development and how developmental processes actually take place. Genes are not blueprints. Developmental processes also depend on non-genetic factors (Oyama et al. (eds.) 2001, Griffiths and Gray 2004).

Secondly, at the brain level, “language areas” are assumed to exist. Allegedly, these areas only process specific features/components/operations of language. However, we observe that the same brain regions can be structurally/functionally impaired in subjects exhibiting different deficits and/or different (including non-linguistic) disorders (e.g. the ventral portion of the occipito-temporal region is underactive in dyslexics (Shaywitz et al. 1998), but functional anomalies in this area have also been linked to a non-linguistic disorder known as prosopagnosia (Sorger et al. 2007)). Perhaps, they are multifunctional by nature. Or perhaps they perform broader, more basic or primitive computations. Moreover, it is quite difficult to draw a precise map of the neural substrate of language, since the limits of the involved brain areas are rather changeable from one subject to another and during growth (and of course, in different clinical conditions) (Prat and Just 2011; see Fedorenko and Kanwisher 2009 for discussion). Not to mention the fact that neurolinguistic research has not lead to principled neurobiological insights. Definitely, maps are not explanations (Poeppel 2012).

It is clear that the linguistic brain does not work as is still too often assumed. On the contrary, it seems that brain areas perform basic kinds of computations that are recruited for different, high level cognitive functions (see Poeppel and Embick 2005 for discussion). It is also clear that we need a better characterization of this linguistic brain.
Thirdly, all human beings (pathological instances aside) are supposed to be endowed by the same, homogeneous faculty of language. This faculty is further construed as one of the modules encompassing the human mind/brain. Ultimately, this module is thought to be present ab initio (e.g. Chomsky 1986, Linebarger 1995, Smith and Tsimpli 1995).

However, different pieces of evidence suggest that the human faculty for language is not actually (so) uniform within the species. We have already reviewed some of them, genetic, and neurobiological. In fact, some of this evidence is not new. For example, different linguistic modalities can coexist in the same subject, as bilingual people in oral and sign languages prove (Emmorey and McCullough 2009). Moreover, psycholinguistic measures are varied across the normal (and of course, the impaired) population (Fenson et al. 2000). And in truth, one important piece of evidence is the very existence of language disorders, which plausibly represent different breakdowns of the faculty that are qualitatively diverse by nature. Moreover, developmental trajectories followed by language acquisition, while encompassing similar milestones, are yet diverse (particularly at the cognitive/neurobiological levels) (Bates et al. 1988, Dehaene et al. 1997). As expected, language ontogeny in pathological populations is even more diverse (see Thomas et al. 2009 for discussion).

Importantly, similar cognitive profiles can rely on different brain architectures (Karmiloff-Smith, 2010). It seems then that there can be many ways of implementing a (more or less) functional faculty of language at the term of growth (see Hancock and Bever 2013 for discussion). Additionally, major changes in the brain architecture and function usually take place across development. “Modules are not born; they are made” (Bates et al. 1988: 284; see Karmiloff-Smith 2010 for discussion), although their basic wiring is achieved before birth, plausibly, genetically-guided (Balaban 2006).

At the same time, modularity pervades biological systems (Wagner 1996, Kitano 2004). However, different types of modules actually exist (Winther 2001, Breuker et al. 2006). Cognitive capacities such as language are very probably cross-modular by nature, that is, they result from the interplay of diverse brain areas performing specific, low-level activities (Griffiths 2007). On the whole, it is clear that the language is not implemented at the brain level as is still too often thought.

Notice, however, that this widespread variation is just one side of the coin. Let us have a quick look at the other side. For example, at the neurobiological level we also observe that anatomical variability is quite constrained. Hence, myelinization patterns, receptor maps, cytoarchitectonic probability maps, and other structural features can be
confidently established (Zilles and Amunts 2009). Similarly, functional variability seems to be constrained as well, to the extent that regions of interest (ROIs) can be always identified (Fedorenko et al. 2011). In sum, although variation is omnipresent, the brain still exhibits a robust structure when processing language (Grodzinsky 2010).

At the molecular level, it is true that the “linguistic genotype” is actually polymorphic. But at the same time, the initial wiring of the linguistic brain is similarly (and regularly) achieved in all subjects (Benítez-Burraco 2009).

When it comes to language growth in the child, we observe that developmental itineraries are also constrained (although not fully predetermined), as was in fact already noted in Lenneberg (1967). Probably, it is the ontogeny of language disorders which more clearly reveals the real nature of the problem. What we recurrently observe in pathological populations is that (see Sirois et al. 2008 and Karmiloff-Smith 2009 for a detailed discussion):

1. Diffuse effects on the brain and on cognitive capacities/abilities are the norm;
2. Deficits in low-level, more generalized processes usually manifest as disturbances of upper, more specialized processes, which ultimately give rise to shortcomings in even higher-level, more specific capacities;
3. Importantly, impaired, delayed or deviant systems are still adaptive. Remember that (substantially preserved) linguistic abilities can be achieved in spite of deeper cognitive impairments.
4. At the same time, breakdowns and compensations, whenever they occur, do not proceed randomly. In other words, phenotypic outcomes which are erratically diverse are not observed.

It is then possible that there exist not so many ways of implementing a functional faculty of language at the term of growth.

At the same time, it is also clear that we need a better characterization of language development (both in the impaired and unimpaired population), one that truly takes the dynamics of development into account (instead of idealizing it away), and ultimately, of how language emerges at the term of growth as “one component of the human mind”, to use Chomsky’s terminology.

Lastly, it has usually been assumed that language has evolved from animal communication by descent with modification (e.g. Pinker and Bloom 1990). However, we actually observe more symbolic complexity as well as more complex (sound) strings in phylogenetically-distant species (Cheney and Seyfarth 1990, Berwick et al. 2011, Cäsar et al. 2013). Additionally, we must address the form-function problem. This problem
entails that we cannot automatically infer modern functions (that is, a modern faculty of language) from human-like, language-related biological structures, like the descended larynx, the mirror neurons or Broca’s areas, or even the derived sequence of “language genes” (see Balari et al. 2013 for discussion). Even evidence of symbolism can be problematic. Symbolic cultures are opaque by nature, whereas linguistic meaning is open, productive by nature (Eco 1976, Silverstein 1976). Ultimately, they can be compatible with other, non-human mental architectures (e.g. Wynn and Coolidge 2004 on the Neanderthal mind). And conversely, modern functions can exist even if some human-like, language-related biological structure is absent (e.g., sign languages are full-fledged natural languages, but they do not make use of the speech-hearing organs). Absence of evidence is not evidence of absence.

Contrary to functions or behaviours, biological structures do exhibit a measurable degree of evolutionary continuity that allows making justified inferences from them regarding language evolution. Accordingly, we should rather rely on the evolution of the neuronal architecture (and even of language-related biological structures) if we want to reach biologically-grounded conclusions about the evolution of language. Ultimately, it seems that evolution (and specifically, the evolution of language) does not take place as has all too often been assumed.

In fact, current evolutionary theory tells us that it is not behaviors that evolve, but biological structures supporting them (Wagner and Altenberg 1996, Carroll 2005, Love 2007) (this is another face of the “form-function” problem), and also that other mechanisms can account for evolutionary novelties such as language, to the extent that phenotypic novelties seem to be largely reorganizational rather than a product of innovative genes (see West-Eberhard 2003, 2005 for a detailed discussion).

It is clear that we need as well a better characterization of how language has evolved in the species.

4 What is to be done?

Fortunately Biolinguistics is progressively moving from this “folk” (i.e., naïve) conception of the biology of language, to more biologically-grounded hypotheses about the nature of language, language development, and language evolution in the species. Instead of singling out works that strike me as going in the right direction, I will highlight a few general traits of the recent literature. Specifically, we must:

i) succeed in integrating into current (bio)linguistic theory novel theoretical paradigms within biology, particularly, in the areas of genetics, developmental biology, neurobiology, and evolution;
ii) grant biology the ultimate role of constraining our “linguistic” hypotheses about language: if to be correct, a linguistic theory must be biologically plausible. For instance, this necessarily requires meeting David Poeppel and David Embick’s (2005) granularity mismatch problem, and above all a willingness to reconstruct linguistic theory from the bottom up, relying as much as possible on generic operations.

Eventually these steps would allow us to satisfactorily address all the problems and shortcomings we have previously faced. In particular, the following assumptions must be adopted:

– genes are not blueprints
– the innate cannot be conflated with the genetic
– developmental processes also depend of non-genetic factors
– there is always an indirect link between the genotype and the phenotype
– developmental itineraries are constrained, but not fully predetermined (in other words, development is both plastic and canalized)
– only biological structures (performing specific activities) are the final output of developmental processes
– functions (that is, forms of behaviour) usually result from the interplay of different biological structures; at the same time, one biological structure can contribute to more than one function
– biological structures (but not the functions they contribute to) are the real evolutionary loci
– biological systems are both robust (i.e. resistant to change) and evolvable (i.e. prompted to change) because of their modular nature
– evolution can be prompted by modifications in any of the factors that affect development (not only genes are involved!)
– phenotypic novelties are largely reorganizational rather than a product of innovative genes

Such a way of conceiving of biolinguistics, which I personally find very much in line with the early conception of the field articulated by Eric Lenneberg before his tragic death, will give pride of place to Darwin (as opposed to Descartes): Darwinian (bio)linguistics, as opposed to Cartesian linguistics. It will move beyond “Poverty-of-Stimulus” arguments, because while it is true that numerous aspects of mature linguistic knowledge cannot be directly derived from the environment, they cannot be directly encoded in the genes either. Instead, the central argument of the field will be one revolving around the granularity issue discussed above.
I anticipate that this new way of articulating the problematic of the field will not only bring us closer to the biologists, it will also bring antagonist tendencies in linguistics to be viewed as complementary. Language is such a complex (adaptive) system that no single gene, no single brain area, no single computational operation on its own will explain it all. Failure to appreciate this is what made the initial promises of Biolinguistics vanish in the eyes of many. We now have in our hands a unique opportunity to revive those interdisciplinary hopes, and shed light on what makes us human.

References


SORESSI, M.; BRESSON, F.; MAUREILLE, B.; NOWELL, A.; LAKARRA, J.; BACKWELL,
L.; JULIEN, M. “Archeological evidence for the emergence of language, symbolism,
and music – An alternative multidisciplinary perspective”. J. World Prehist., 17

vande MOORTELE, P. F.; LEHÉRICY, S.; Le BIHAN, D. “Anatomical variability in
the cortical representation of first and second language”. Neuroreport, 8 (1997),
p. 3809-3815.


EMMOREY, K.; MCCULLOUGH, S. “The bimodal bilingual brain: effects of sign

ENARD, W. “FOXP2 and the role of cortico-basal ganglia circuits in speech and

FALCARO, M.; PICKLES, A.; NEWBURY, D. F.; ADDIS, L.; BANFIELD, E.; FISHER,
S. E.; MONACO, A. P.; SIMKIN, Z.; CONTI-RAMSDEN, G.; SLI Consortium. “Genetic
and phenotypic effects of phonological short-term memory and grammatical
morphology in specific language impairment”. Genes Brain Behav., 7 (2008),
p. 393-402.

FEDORENKO, E.; KANWISHER, N. “Neuroimaging of language: Why hasn’t a clearer

FEDORENKO, E.; BEHR, M. K.; KANWISHER, N. “Functional specificity for high-level
p. 16428-16433.

variability in early child language: Don’t shoot the messenger”. Child Dev., 71 (2000),
p. 323-328.

FISHER, S. E.; SCHARFF, C. “FOXP2 as a molecular window into speech and

GIBSON, C. J.; GRUEN, J. R. “The human lexinome: genes of language and reading”.

GRAHAM, S. A.; FISHER, S. E. “Decoding the genetics of speech and language”. Curr

GRAYBIEL, A. M. “Building action repertoires: memory and learning functions of the

GRIFFITHS, P. E. “Evo-Devo meets the mind: towards a developmental evolutionary
psychology”. In: BRANDON, R.; SANSOM, R. (Eds.). Integrating Evolution and

GRIFFITHS, P. E.; GRAY, R. D. “The developmental systems perspective: organism-
environment systems as units of evolution”. In: PRESTON, K.; PIGLIUCCI, M. (Eds.).

GRODZINSKY, Y. “The picture of the linguistic brain: how sharp can it be? Reply to


VARGHA-KHADEM, F.; WATKINS, K. E.; ALCOCK, K. J.; FLETCHER, P.; PASSINGHAM, R. E. “Praxic and nonverbal cognitive deficits in a large family with a genetically


Postal address:
Department of Linguistics
Universitat de Barcelona
Gran Via de les Corts Catalanes, 585
08007 Barcelona, Spain

Date received: 30/09/15
Date accepted: 07/10/15