LOOKING BACK, MOVING FORWARD

Is Universal Grammar ready for retirement?
A short review of a longstanding misinterpretation[1]

JOSÉ-LUIS MENDÍVIL-GIRÓ

University of Zaragoza

In this paper I consider recent studies that deny the existence of Universal Grammar (UG), and I show how the concept of UG that is attacked in these works is quite different from Chomsky's, and thus that such criticisms are not valid. My principal focus is on the notions of 'linguistic specificity' and of 'innateness', and I conclude that, since the controversy about UG is based on misinterpretations, it is rendered sterile and thus does unnecessary harm to linguistic science. I also address the underlying reasons for these misunderstandings and suggest that, once they have been clarified, there is much scope for complementary approaches that embrace different research traditions within current theoretical linguistics.

KEYWORDS: faculty of language, linguistic specificity, nativism, recursion, Universal Grammar

1. INTRODUCTION: UNIVERSAL GRAMMAR UNDER ATTACK

Each year the Internet-based scientific magazine *Edge* publishes a question for the scientific community. In 2014 this question was: What scientific idea is ready for retirement? Not surprisingly, one of the responses (by Benjamin Berger) proposed the Chomskyan notion of UNIVERSAL GRAMMAR (UG); other answers included the concept of race, Moore's law, that there is no reality in the quantum world, and that the Universe began in a state of extraordinary low entropy. We might justifiably use the phrase 'not surprisingly' here in that a notable amount of recent work has had as its main (almost exclusive) objective the denial of the existence of UG. Publications include a book by Evans (2014), and papers in *Frontiers in Psychology* by Dąbrowska (2015), Christiansen & Chater (2015), and Everett (2016). Another common (and symptomatic) trait of these contributions is that they have appeared in media with a broad audience or in general cognitive science journals, and not in technical journals of linguistics.

859

My main concern here is not to present arguments against these contributions and in support of UG, but to show that the concept of UG which these works attack is quite different from that found in Chomsky's work itself, and that, consequently, the criticisms are not valid.

I will set out what Chomsky means by UG, how it has evolved within Generative Grammar (GG), and why we can say that the concept of UG, as dealt with by its opponents, is the result of a misinterpretation. The controversy, being so vitiated, becomes sterile and does unnecessary harm to linguistics as a discipline, especially when it is presented in the general context of the cognitive sciences.

In the final sections I show that the underlying main point of conflict between UG supporters and detractors is the different appreciation of the role of syntax in human language and cognition, and I suggest that a clarification of this misunderstanding can reveal that there is more complementarity between current linguistic traditions than usually admitted.

2. WHAT WAS, AND STILL IS, UNIVERSAL GRAMMAR?

The concept of Universal Grammar commonly associated with Chomskyan linguistics is in fact a very conservative concept. It dates back to at least the 13th century, when speculative grammar flourished, and is characterized in Aristotelian terms by Roger Bacon (1214–1299) in a very similar way to how it was inherited by Chomsky: 'Grammatica una et eadem est secundum substantiam in omnibus linguis, licet accidentaliter varietur' (R. Bacon c. 1250, *apud* Malmberg 1991: 128).[2]

In the tradition of rationalist grammar (to avoid the controversial expression *Cartesian Linguistics*, Chomsky 1966) this notion is taken up in a very similar sense, as illustrated in the following quotations from influential French grammarians of the period:

> Dans toutes les langues du monde, il n'y a qu'une même manière nécessaire pour former un sens avec des mots (Du Marsais 1729, *apud* Joly & Stéfanini (1977): 11)[3]

> La science grammaticale est antérieur à toutes les langues, parce que ses principes sont d'une vérité éternelle, et qu'ils ne supposent que la possibilité des langues (Beauzée 1767, *apud* Joly & Stéfanini (1977): 11)[4]

What these selected references have in common is the idea of a set of properties or principles underlying the structure of all languages that restrict what a possible human language can be. The term GRAMMAR is clearly related to the study of

---

[2] 'Grammar is substantially the same in all languages, even though it may vary accidentally' (translation *apud* Lyons 1991: 172).

[3] 'In all the languages of the world, there is only a necessary way to build a sense with words' (my translation).

[4] 'Grammatical knowledge is prior to all languages, because its principles are eternal truths, and they only determine the potential of languages' (my translation).

the structure of languages, and UNIVERSAL has to do with the strict sense of the term: that which is invariable in space and time. Thus, as suggested by F. Bacon, languages change, but their nature as languages remains.

As early as 1965 (in *Aspects of the Theory of Syntax*) Chomsky echoes this perspective and argues that, contrary to the then current trend of structural linguistics, 'the grammar of a particular language' '[needs] to be supplemented by a universal grammar that ... expresses the deep-seated regularities which, being universal, are omitted from the grammar itself' (Chomsky 1965: 6). Indeed, in its original use, the expression UNIVERSAL GRAMMAR also referred to those aspects that traditional descriptive grammars do not analyse because they are considered common. Thus, in a French grammar it might be worth mentioning that there are articles (compared to Latin, for example), but it would make no sense to say that French has words or sentences.

In the same work, Chomsky claims that 'the main task of linguistic theory must be to develop an account of linguistic universals' that reveals 'the properties of any generative grammar' and is not 'falsified by the diversity of languages' (ibid.: 27–28). It is important to note that Chomsky does not refer to properties common to all languages (in the style of Greenberg's universals), but to properties common to the grammars that generate languages: so, regularities 'need not be stated in the grammar' of a language, but 'only in general linguistic theory as part of the definition of the notion "human language"' (ibid.: 117).

Given the naturalistic conception of language adopted by Chomsky (in contrast to the understanding of language as a cultural construct which was prevalent when GG arose, and on which the authors who now oppose UG also rely), his proposal, which has gone unchanged throughout his vast body of work, is that UG is the theory of the requirements that human nature – human biology – imposes on the systems of linguistic knowledge which are eventually developed, that is, languages (in the sense of I-languages).

The model known as Standard Theory (Chomsky 1965) barely developed the theory of UG beyond the foundational proposal (which has remained constant to this day) that human language is a natural object, and that it is characterized by recursive enumerability, resulting in its digital infinity.

As we know, the transition from the Standard Theory to the model of Principles and Parameters (P&P, Chomsky 1981, Chomsky & Lasnik 1993) involves the aim of developing a theory that maintains the descriptive adequacy of the generative grammars developed in the previous period, but which at the same time addresses so-called explanatory adequacy, understood as the model's ability to explain language acquisition. It is during this period that a specific theory about the content of UG really flourishes within generative grammar, and the theory of UG developed within this model is, to a large extent, the source of the erroneous (or caricatured) conceptions of UG mentioned above. Nevertheless, the notion of UG continues unchanged in this model:

> To a good first approximation, the initial state appears to be uniform for the species. Adapting traditional terms to a special usage, we call the theory of the state attained its *grammar* and the theory of the initial state *Universal Grammar* (Chomsky & Lasnik 1993: 507)

What has changed in the development of GG is the explicit content that is attributed to UG at a given time, but not its definition. UG is therefore the theory of the initial state of the language faculty: 'it is the prespecification in the brain that permits the learning of language to take place' (Jackendoff 2002: 72). It is very important to note that, as it is defined, UG exists by definition. We can only deny the existence of the initial state of the language faculty if we deny that the language faculty exists. But the language faculty is an objective fact that cannot be denied. Indeed, even the most radical critics of Chomsky's naturalistic view are compelled to accept that humans, compared to other (natural or artificial) organisms, are endowed with a capacity for language. Let us consider just two examples:

> We would urge our colleagues simply to talk about the human capacity for language (including language-special aspects of cognition, if any) and the ingredients or elements that may contribute to that – the rest carries too much baggage, as the field advances rapidly beyond the creaky old baggage train. (Evans & Levinson 2010: 2742)

> Nativism, again, is the idea not only that we are innately capable of language (everyone surely believes this), but that our capabilities are specific to particular domains, e.g., grammar. (Everett 2016: 1)

In what follows I will consider why for many detractors of UG an innate capacity for language is acceptable, but UG is not, despite it being specifically defined as the initial state of the faculty of language, that is, again in Chomsky's terms, 'a characterization of the child's pre-linguistic initial state' (Chomsky 1981: 7).

3. THE COMPLEX (BUT FALSE) PROBLEM OF SPECIFICITY[5]

As we can clearly perceive in Everett's words, above, the main problem in accepting the nativist complexion of GG seems not to involve the rejection of the idea that the human body affects the structure of the systems of knowledge we call languages (which is what Chomsky's hypothesis asserts), 'but that our capabilities are specific to particular domains, e.g., grammar'. That is, UG's detractors admit that there is a biological bias for language, but they reject that this bias is specifically linguistic. And it is in this notion of LINGUISTIC SPECIFICITY where all the problems on the definition and interpretation of UG in fact arise.

The idea that there is a biological bias for language seems impossible to be distinguished from the idea that this biological bias is specifically linguistic.

---

[5] Some paragraphs of this section are adapted from Mendívil-Giró (2012).

862

Hence, the debate is in itself sterile. Of course, such an argument does not imply that those who fail to see the two statements as indistinguishable lack basic rationality, but that they operate with a different concept of what a human language is (basically an object external to the mind).

If one wants to argue that human languages are cultural systems it is clear that one will not be comfortable with the idea that brain structure determines the essential structure of languages, though perhaps one will be more comfortable with the (more diffuse) idea that such restrictions imposed by the brain are of a general nature, that is, similar to those that can be imposed on any other system of knowledge internalized from the environment. It should be added that the conception of languages as cultural objects is also associated with a partial (or simplified) vision of what is actually the structure of languages, sometimes implying the denial that languages have the syntactic structure generative linguists have discovered (e.g. Everett 2005, Christiansen & Chater 2015). These significant issues will be addressed in Sections 6 and 7 below.

I turn now to focus on the claim, perhaps a surprising one, that accepting a biological bias for language and simultaneously rejecting UG is inconsistent. Let us consider the view of another prominent detractor of UG, in which this difference is stated explicitly:

> For sure, all of the world's languages have things in common . . . But these commonalities come not from any universal grammar, but rather from universal aspects of human cognition, social interaction, and information processing – most of which were in existence in humans before anything like modern languages arose. (Tomasello 2009: 471)

Here we find both of the issues mentioned: the problem of specificity, and the conception of languages as external objects (to be addressed in Section 7 below). The former clearly has to do with the literal meaning of the expression UNIVERSAL GRAMMAR. Note that Tomasello does not reject language universals, but rather that the properties that explain them are specifically linguistic:

> Why don't we just call this universal grammar? The reason is because historically, universal grammar referred to specific linguistic content, not general cognitive principles, and so it would be a misuse of the term. It is not the idea of universals of language that is dead, but rather, it is the idea that there is a biological adaptation with specific linguistic content that is dead. (Tomasello 2009: 471)

But we can ask what is meant by 'a biological adaptation with specific linguistic content': Does it mean that there would be genes that specify grammatical categories? Or that there would be parts of the brain dedicated to language and whose removal would leave intact all other cognitive and motor faculties? Or that there would be linguistic neurons and nonlinguistic neurons, just as there are pyramidal and spherical neurons? No one seems to have defended such a position. The idea of a 'biological adaptation with specific linguistic content' is thus a kind of straw man in the argument. Innatism, in Chomky's definition of UG, refers to the bias that the human body and the laws of nature underlying its anatomy and physiology impose on the systems of linguistic knowledge that

863

are developed. Chomsky's main hypothesis is that such a bias exists, not what its ultimate nature is, something that only can be discovered empirically, and not only through linguistics.

Thus, if everyone agrees that this bias exists, the dispute seems to be meaningless.

Of course, one can object to such a conclusion on the grounds that it is indeed relevant to state whether or not this biological conditioning is specifically linguistic, but then we are forced to define more precisely what we mean by 'specifically linguistic', and at this point problems and misunderstandings arise.

In fact, if we stick to the literal meaning of the expression 'specifically linguistic', we might say that there cannot be anything that is at once biological and specifically linguistic, just as that there cannot be anything that is at the same time chemical and specifically biological. These are contradictory expressions: if something is biological, it cannot not be specifically linguistic (in that it will also be biological), just as if something is chemical, it cannot be specifically biological (because it will also be chemical).

When, in a naturalistic context, we say that a principle, or an entity (for example the notion of C-COMMAND), is specifically linguistic, what we are saying is that we do not YET know from what non-linguistic aspects that principle, or that entity, is derived. But, crucially, we are not saying that this principle, or entity, must be 'linguistically irreducible', as seems to be erroneously attributed to the theory of UG. Consider, as an example, the following passage:

> The extent of crosslinguistic diversity and the considerable individual differences in the rate, style and outcome of acquisition suggest that it is more promising to think in terms of A LANGUAGE-MAKING CAPACITY, i.e., a set of domain-general abilities, rather than an innate body of knowledge about the structural properties of the target system. (Dąbrowska 2015: 1, my emphasis)

But what really is the difference between 'a language-making capacity' and 'an innate body of knowledge about the structural properties of the target system'? It is very difficult to see any difference between the two descriptions once we have established that 'linguistic specificity' can only emerge from what is non-linguistic, something never denied by Chomsky, as we will see.

In fact, as illustrated below, the overall goal of GG has always been to integrate the study of language within the natural sciences and, therefore, as in any scientific endeavour, the ultimate goal is unification, that is, the reduction of what is specifically linguistic to elementary entities which are not, by definition, specifically linguistic.

In general terms, functionalist and cognitive linguists criticize Chomsky's postulation of a specifically linguistic and specifically human Faculty of Language (FL), whose initial state is UG, arguing that it is more economical, more biologically plausible, and more consistent with standard scientific reasoning to assume that no such thing exists, and that languages can be explained as the result of general limitations of learnability, the recruitment of other cognitive systems, and functional pressures arising from their use for communication and thought.

864

And this is certainly a possibility, one which the Minimalist Program (MP) developed by Chomsky and followers has in fact been exploring over the last twenty years.

It could be said, then, that what has confronted the traditions in linguistic research during the last half century is the problem of the innate or emergent character of the FL. However, this is again a false problem, a false controversy, because FL, like any other human faculty, is simultaneously innate and emergent.

On the one hand, there is no doubt that human beings have an ability that enables them to learn and use a human language. Given that the rest of known organisms (whether natural or artificial) lack such an ability, it is fair to say that this ability is specifically human and, therefore, innate in any normal human being. As Chomsky points out (using a *reductio ad absurdum*),

> To say that "language is not innate" is to say that there is no difference between my granddaughter, a rock and a rabbit. In other words, if you take a rock, a rabbit and my granddaughter and put them in a community where people are talking English, they'll all learn English. (Chomsky 2000: 50)

Fitch follows this same elementary line of reasoning, in a more temperate mode:

> Clearly, immersion in a linguistic environment is not enough for spoken language to develop in most organisms. There must therefore be something about human children which differentiates them from other species, and this something provides one of our core explananda in biolinguistics. We might gloss this neutrally as 'the human capacity to acquire language'. In generative linguistics this capacity is traditionally called the 'Language Acquisition Device', and a characterization of its properties termed 'Universal Grammar'. (Fitch 2009: 288)

On the other hand, it cannot be false that human language (as a human ability) is emergent, that is, the result of the combination or the arrangement of elements that in themselves are not 'linguistic'. Let's consider an analogy between life and language: life is certainly an emergent phenomenon, but by no means less real and specific because of this. As Stuart Kauffman puts it:

> Life is not located in the property of any single molecule – in the details – but is a collective property of systems of interacting molecules. Life, in this view, emerged whole, . . . not to be located in its parts, but in the collective emergent properties of the whole they create . . . The collective system is alive. Its parts are just chemicals. (Kauffman 1993: 18, 24)

On the linguistic side, it is a given configuration of non-linguistic elements that can produce 'linguistic' entities. An arrangement of 'dead' elements gives rise to 'life', and an arrangement of 'biological' elements gives rise to 'language'. In fact, Chomsky himself, when describing FL as a subcomponent of the brain that is 'specifically dedicated to language', notes: 'as a system, that is; its elements might be recruited from, or used to, other functions' (Chomsky 2004a: 124, fn. 1). As noted by Rooryck et al., the so-called STRONG MINIMALIST THESIS (Chomsky 2007) explores precisely this direction, but this does not negate the notion of UG:

865

> In terms of Hauser et al.'s contrast between FLB and FLN, the latter would then be empty and the uniqueness of UG would reside in its particular combination of traits each of which is shared with some other domain or organism. As human language differs from all other systems *something* (but not necessarily *some thing*) must be unique. (Rooryck et al. 2010: 2655)

If we relax the literal interpretation of the expression 'specifically linguistic', it would not be contradictory to state, for example, that there is a biological adaptation that is specific to language. But even in such a case, it would not be true that it is a 'biological adaptation with specific linguistic content', for at least two reasons: (i) such an entity, if it is biological, should be made of molecules, proteins and cells, none of which are 'specifically linguistic', nor can they be; (ii) such a biological entity could and should continue to have the structure and function that it had prior to the 'specifically linguistic' adaptation.

Let us take a concrete example to make this clearer. Assume, ignoring possible objections, that certain aspects of the human vocal tract have been selected through evolution to favour speech (e.g. Lieberman 1984). Obviously we would not say that this organ is specifically linguistic or represents 'a biological adaptation with specific linguistic content'. First, because it is made of biological material (cells, tissues), not of linguistic material; second, because it has other functions besides allowing the articulation of sounds for speech. The same should apply to less concrete (and less isolatable) entities than the larynx, that is, the properties of the anatomy and physiology of the brain that allow, for example, syntactic computation: even if they had been selected for language through evolution, in no way would it be correct to claim that they were thus specifically linguistic adaptations, for the same two reasons.

What may be, and in fact must be, specifically linguistic is the initial formulation of the properties of such entities (again, take c-command). And it must be so because, unlike what happens with the larynx, linguists have to formulate theories of UG based on the analysis of languages, not based on the analysis of the brain or the genome.

The question of which of the principles governing the formation of the FL in each person are specifically human (and specifically linguistic) is an empirical matter that cannot be resolved prior to the determination of what such principles are. Distinction between a Narrow and a Broad Faculty of Language (FLN & FLB) of Hauser, Chomsky & Fitch (2002), and the 'factorization' formulated by Chomsky (2005) (which is common in developmental biology) are attempts to make this more explicit in the investigation of FL.

However, as I noted above, the question is not whether linguistic principles and structures may be reduced to more basic principles and structures (and ultimately to neuronal structures), but when, how and at what level the reduction can be done without ceasing to account for the structure of languages. As will be discussed in Section 6 below, the model which UG's detractors support skips a stage, so to speak, the stage of linguistic theory itself. Among the detailed (albeit often superficial) descriptions of the structure of languages and the general cognitive

866

principles that they tend to mention, there is an absence of explanation, and this produces an incomplete view of language.

4. REDUCING SPECIFICITY IS A SIGN OF PROGRESS (IF YOU MAINTAIN DESCRIPTIVE ADEQUACY)

Since one of the essential objectives of GG is to integrate the study of language within the natural sciences, it should come as no surprise that the natural tendency, as the discipline has developed, has been to try to reduce as much as possible what is postulated as specifically linguistic in the theory of UG.

As I have already noted, the shift to the P&P model was driven by the need to make compatible the process of language acquisition with the complex systems of rules of the initial models. In the jargon of the time, there was a transition from descriptive adequacy to 'explanatory adequacy'. A natural consequence of this step was the postulation of a richer and more complex UG, and hence a higher degree of linguistic specificity in the principles and entities postulated.

Not surprisingly, attributing to UG much of the structural complexity of human languages greatly facilitated the explanation of the acquisition of the mechanisms necessary to explain the linguistic knowledge of speakers, but simultaneously compromised the biological and evolutionary plausibility of the theory of UG itself. And as a result of this tension the so-called Minimalist Program (Chomsky 1995 and following) emerged. Chomsky himself has characterized this transition:

> At the time, it seemed that FL must be rich, highly structured, and substantially unique.... Throughout the modern history of generative grammar, the problem of determining the character of FL has been approached 'from top down': How much must be attributed to UG to account for language acquisition? The MP seeks to approach the problem 'from bottom up': How little can be attributed to UG while still accounting for the variety of I-languages attained? (Chomsky 2007: 2, 4)

For any organ of the body, Chomsky (2005) observed that we can identify three factors that determine its development: (i) genetic endowment, (ii) environment, and (iii) general principles of growth and development that are not specific to the organism, and perhaps not even to the organic world. The Minimalist Program emphasizes the investigation of the 'third factor' with the objective of finding what Chomsky calls 'principled' explanations of the properties of the FL finally obtained in each person. As Chomsky explains unambiguously,

> to the extent you can answer the 'why' question, you will have shown that there's a principled reason why things are the way they are. It's not just abstracting principles from the final state and putting them into the initial state, which is a major step forward, but actually extracting them from the language altogether and showing that they really belong somewhere else. (Chomsky 2004b: 151)

It is very difficult to imagine a less 'language-specific' perspective, unless one reads Chomsky again:

867

> To the extent that you can make some progress in the minimalist program, domain-specificity of language is reduced to some special arrangement of elements that are not language-specific. (Chomsky 2004b: 163)

However, note that this conceptual and methodological change has not altered the definition of UG:

> UG is what remains when the gap has been reduced to the minimum, when all third factor effects have been identified. UG consists of the mechanisms specific to FL, arising somehow in the course of evolution of language. (Chomsky 2007: 5)

Crucially, Chomsky's 'negative' definition of UG does not imply 'irreducibility' either. He has suggested (Chomsky 2007 and later works) that the Computational System (CS), the supposed last bastion of 'linguistic specificity', would be an essential part of human thought, developed through evolution prior to the connection with the sensorimotor systems that allow us to speak of 'language' proper, and that CS would also underlie, for example, the human ability for arithmetic (i.e. the successor function) (see Section 6 for discussion).

We have seen, then, that the definition of UG is constant throughout the history of GG as a scientific discipline. We have also seen that, as a logical part of its development, UG reaches its maximum theoretical development and the greatest degree of linguistic specificity in the P&P model. The result is a complex theory that, although favouring the explanation of the process of language development (maintaining descriptive adequacy), complicates its connection with the life sciences, within which a naturalistic linguistics (a biolinguistics) must necessarily be integrated. And we have also seen that the central leitmotif of the Minimalist Program is precisely to facilitate the integration of the theory of UG into more basic disciplines, with the obvious consequence that its 'linguistic specificity' thins considerably.

However, many authors question the theory of UG as if its development had stopped in the 1980s. This is surprising; it would be like criticizing current physics because Newton thought that gravity was an instantaneous force, ignoring Einstein's theory of relativity. Yet many very recent criticisms of the notion of UG rely on a model which has already been superseded in GG itself. The controversial issue of the alleged genetic coding of UG is a clear illustration of this.

5. CONFUSING 'INNATE' WITH 'GENETICALLY SPECIFIED'

Identifying or confusing 'innate' with 'genetically specified' is a common error of both proponents and detractors of UG; yet only the former (Chomsky at least) have corrected this mistake.

As we know, following the development of genetics Darwin's theory of evolution was associated with a geneticist conception of the development of organisms, a vision in which the environment is of scant import during development, the process being entirely governed by the genetic program. Such was the prevailing view in biology in the 1950s and 1960s, precisely when GG emerged.

868

The geneticist model of development was especially attractive to the Chomskyan (naturalist) approach to human language, in that it faced a similar problem: how to explain the robustness and consistency of language development in an unstable and confusing environment which provided very poor evidence on the systems of knowledge finally obtained. Consequently, the P&P model, as we have seen, focused on an innate, rich and specifically human linguistic component as an explanation of what was known as Plato's Problem.

Of course, Chomskyan linguistics was not focused on the study of the possible 'language genes', but rather on opposing the view that continues to defend a conception of language in which its development is a process based on imitation, induction and generalization performed by general systems of learning. If the Chomskyan model was essentially geneticist until the nineties, it was simply because it was a naturalistic model, a model which argued that language is a natural phenomenon common to our species, and that linguistics should be understood as part of biology, the science that studies the form and structure of organic beings. Since the prevailing biology was geneticist, generative grammar was geneticist by simple inheritance.

If we review the literature of the time, we can find statements by Chomsky and other distinguished generativists attributing UG to the human genome (see Benítez-Burraco & Longa 2010: 314, fn. 9, for some examples), but not because they attribute certain properties of UG to certain genes (an absurd claim when working with syntactic categories and agreement features and not with proteins and codons), but because it was assumed that what is naturally specified, what is 'innate' or not learned, corresponded to what is genetically specified. Nevertheless, as Uriagereka pointed out, from a minimalist point of view, 'UG isn't about genes, but about formal properties of human language. UG could equally be true if human beings were made from musical notes instead of genes' (Uriagereka 1998: 47).

Although geneticism in developmental biology is still prevalent, things have changed in recent decades. So-called Evo-Devo models (from evolutionary developmental biology) have transformed current evolutionary biology into a much more pluralistic approach, and GG as accordingly changed.

Longa & Lorenzo (2012) note that the gene-centric model based on the notion of a genetic program has being questioned in recent decades by the so-called developmentalist challenge. The gene-centered model implies that genes are the only possessors of the essential information that guides the growth and maturation of organic structures. Genes, then, are considered as constituting a stand-alone program that includes information on the patterns of structural organization and the instructions for the deployment of these structures in time and space. This view can be considered 'preformationist' because it implies that the sources of organic development are already contained in information in the nuclear DNA, and that the process of development will involve the eclosion of what is already contained in the genes. According to the new biology of development, phenotypic traits (whether anatomical, physiological or cognitive) cannot be contained or

869

specified in genes. This implies that the notion of 'genetic program' as the only source of information for developmental processes is a distorted vision of how such processes occur, as it ignores the significant contribution of other factors and resources located between genotype and phenotype and without which the development process simply cannot occur.

As we have seen in the previous section, the Minimalist Program is a research program clearly connected with this development of biology, and is specifically aimed at trying to clarify which aspects of FL are (i) a consequence of the biological endowment of the species (which can therefore have evolved adaptively and could be genetically encoded) and (ii) which are due to principles of simplicity, computational elegance, or processes of cerebral development, factors which, therefore, are not the result of evolutionary adaptation, but a consequence of the evolution of the human brain, or a result of deeper formal or physical principles governing systems of a certain complexity.

Chomsky (2005) admits that a key issue to solve from the point of view of the biological research of language is to what extent the principles that determine human language are unique to this cognitive system, or if similar formal arrangements can be found in other human cognitive domains or in other organisms; but that 'an even more basic question from the biological point of view is how much of language can be given a principled explanation, whether or not homologous elements can be found in other domains or organisms' (Chomsky 2005: 1). And this is precisely the goal of the Minimalist Program, to look for a 'principled' explanation of the principles attributed to UG.

Hence, it is clearly misguided to present UG as a proposal of some 'genetically codified grammatical content', yet this is just what modern UG detractors do. For example, Bergen (2014) claims that the hypothesis of UG implies that 'core commonalities across languages exist because they are part of our genetic endowment'. But how could a linguistic universal (albeit undisputed) be part of our genetic makeup? Genes do not build systems of knowledge or ideas or concepts; what they build are proteins, cells and tissues, which, depending on how they are organized and operate, restrict or condition the emergence and development of possible systems of knowledge. The real question is: how could Chomsky claim that there are grammatical genes or syntactic neurons if he has never worn a white coat?

Similarly, Goldberg (2015 – in a blog entry) concludes that a study on binding theory which she discusses shows that there is no UG, i.e. a kind of 'innate grammatical knowledge', as if the types of anaphoric expressions that exist in languages would have been selected by natural selection, as is the case with an upright posture or opposable thumbs.

An extreme case of confusion appears in Everett (2016), who quotes an argument by Lieberman (2013). Everett clarifies, rightly, that the hypothesis of UG predicts that any child can learn any language, but argues (following Lieberman's argument) that the claim that UG is genetically encoded predicts the opposite: 'if language were actually specified on the genes, it would be

870

subject to mutations, presenting a non-trivial problem for UG' (Everett 2016: 2). As an example, Everett cites the well-known null subject parameter. It is worth considering his own formulation of the problem:

> To take a concrete example, consider one commonly assumed "parameter" of UG, the so-called "pro-drop" parameter – intended to account for the ability of speakers of a language to omit overt subjects from sentences. Thus in Portuguese, a pro-drop language via a single gene (unlikely) or relationships among multiple genes, one can utter *Está chovendo* while in English the literal translation *Is raining* is ungrammatical. (Everett 2016: 2)

As can be seen, Everett is assuming that a parametric choice would be associated with one or several (more likely, he concedes) genes. But this is a clear misrepresentation of what the principles and parameters in the P&P model in fact are. If it were conceivable at all that there is a specifically grammatical gene, such a gene would determine the underlying common principle (for example, 'every sentence must have a subject'), but the genome would remain 'silent' on the parametric option, that is, whether, in the language developed by an individual in contact with linguistic stimuli in his or her environment, the subject must always be realized phonologically (as in *It's raining*) or not (as in Portuguese *Está chovendo* 'It's raining'). The reason as to why a language, in this idealized context, belongs to one type or the other relates solely to external linguistic stimuli from the environment, and never to aspects of the human genome, even if there were grammatical genes. Therefore, it makes no sense to ask 'whether it is possible for there to be a mutation that would prevent a particular person from learning a pro-drop language' (Everett 2016: 2). To take this to an extreme, we might then ask whether a genetic mutation might result in the inability to speak languages with subjects, whether expressed phonologically or not, although such reasoning is clearly not worth our attention. Thus in Everett's conclusion, both the protasis and the apodosis are meaningless, and a symptom of misunderstanding:

> In other words, and quite ironically, if grammar is carried on the genes, then the strongest evidence for UG would be the discovery that *not all people may be able to learn every language*. (Everett 2016: 2)

6. DENYING/IGNORING LANGUAGE STRUCTURE IS NOT A SOLUTION

As I have already argued, although it may seem paradoxical, in the study of language we must be both nativists and emergentists; in this there is no choice. The difference between those who claim one or the other is a matter of their optimism regarding our capacity to understand emergent systems and of the degree to which we address the formal and abstract structure of human languages as systems of knowledge, which is not uniform among different traditions.

In fact, a real problem behind this false controversy is that a large number of linguists and cognitive scientists ignore or simplify the formal structure of languages. One might ask why some linguists assume that the structure of languages is learnable from environmental data while others say that it is not.

871

It cannot be because we do not know what is and what is not learnable. We know that there are things that cannot be learned by induction from a sample of input, and others that can. The difference of opinion emerges from the fact that we see the structure of languages (and the very nature of language) in different ways. The shallowness of functionalist/cognitive models of the structure of languages is perhaps at the root of the confidence in the idea that induction and analogy are adequate means of extracting the information needed to develop such systems of knowledge.

But as I have argued, the problem is not only terminological, nor does it only involve degrees of optimism and professional rivalries. It may be instructive to return now to the final claim in Tomasello's objection cited above, that is, that many of the factors that determine the unity of language (including universal aspects of human cognition, social interaction, and information processing) 'were in existence in humans before anything like modern languages arose'. Note, however, that Tomasello is here implicitly assuming that human languages emerged externally from such 'previous factors', that is, that human languages are cultural collective 'inventions', objects external to the mind and the brain.

In this externalist context it is conceivable, then, that there are profound and substantial differences among languages, within the loose external frame determined by such factors, which are considered extralinguistic by definition. If we deny the existence of UG (i.e. the initial state of FL), then we are denying that humans have a bias towards language, that is, a specific capacity to learn the language from the environment, and that this capacity inevitably determines its structure. This position would then involve the theoretical possibility that different languages can activate different general cognitive systems or use cognitive systems differently. In such a case, it would not even be justified, as Evans & Levinson (2010: 2742) suggest, 'simply to talk about the human capacity for language', as there may be various human capacities for language. Such a possibility is suggested, in fact, by McMurray & Wasserman (2009) in their laudatory review of Evans & Levinson (2009):

> Yet, just as there is no universal structure, there may be no universal learning mechanism responsible for language. Language is a behavior assembled by many processes, an assembly guided by the language being learned. (McMurray & Wasserman 2009: 459)

I have referred elsewhere (see Mendívil-Giró 2012, Moreno & Mendívil-Giró 2014: Chapter 3) to the logical and empirical problems of such an externalist view of languages as 'the paradox of languages without a faculty of language'. Curiously, in a recent paper, Christiansen & Chater (2015) present their conclusions under the title 'Language without a Faculty of Language'. Actually, their objective is to demonstrate that there is not an FL (and hence, no UG), that is, that language does not have a biological origin but is a purely cultural object: 'It is time to return to viewing language as a cultural, and not a biological, phenomenon' (Christiansen & Chater 2015: 14).

872

The strategy used by Christiansen & Chater (C&C hereafter) to liquidate UG is in fact to 'join' Chomsky's minimalist drift and focus on attacking the last stronghold of the FL: the recursive human computational system (CS).

In contrast to previous attempts to argue against CS claiming that some human languages do not have recursive structures (as in Everett 2005), C&C propose a more radical alternative, that no language has a recursive computational system:

> [T]he recursive character of aspects of natural language need not be explained by the operation of a dedicated recursive processing mechanism at all, but, rather, as emerging from domain-general sequence learning abilities. (Christiansen & Chater 2015: 2)

According to C&C, therefore, the structural complexity of the syntax of human languages (which, according to GG, would be the result of the application of a recursive mechanism of structure generation) would be, rather, 'apparent' complexity. Note their claim that the recursive character of linguistic expressions would emerge from a general ability to learn sequences, which amounts to saying that the syntactic structure underlying linguistic expressions would be linear, i.e. sequential. This is like ignoring the syntactic theory of the last sixty years, as if it simply did not exist. Again, relatively speaking, it would be the same as claiming that modern physics is too abstract and that we should ignore quantum mechanics. In fact, C&C suggest that this complexity comes from the researcher, not from the object of study: 'What needs to be explained is the observable human ability to process RECURSIVE STRUCTURE, and not recursion as a hypothesized part of some grammar formalism' (Christiansen & Chater 2015: 3).

But note the potential problem: if humans are not endowed with a recursive computational system, but (as their proposal explicitly states) we can only use general mechanisms for learning complex sequences to process recursive structures, then it is not clear what is the origin of these 'recursive structures' that, as they claim, we process so clumsily.

However, let us consider first what we mean when we talk about recursion, and then return to C&C's specific proposals, and their implications. Like so many before them, including many generativists, C&C confuse recursion as a property of a computational system (which is Chomsky's use of the concept, from 1955 onwards, as clearly shown by Lobina 2014) with so-called recursive subordinate structures, that is, when a sentence contains a sentence, or when a noun phrase contains a noun phrase. The latter are routine consequences of the first, but by no means are they the same thing.[6]

Since the capacity of human memory is clearly overwhelmed by the creativity of human language, in the sense that we cannot embrace a complete listing of all grammatical sentences (or of all ungrammatical sentences) of a language, Chomsky proposed in the 1950s (a key step in the foundation of linguistics

---

[6] To avoid confusion, I will use the expression SELF-EMBEDDED STRUCTURES for so-called 'recursive structures'.

873

as a cognitive science) that we must attribute to the speaker's mind/brain a computational system (a generative grammar) that recursively produces the set of grammatical sentences, a set which in principle is potentially infinite or indefinite.

The use of the term RECURSIVELY in the preceding paragraph deserves explanation, for it is the source of the errors reported. This notion of recursion (as a property of a computational system) is directly linked to the mathematical logic of the 1930s and 1940s (associated with pioneers in the theory of computability such as Gödel, Post, Church and Turing) that had a direct and decisive influence on Chomsky in the 1950s. In the handbooks on mathematical logic with were part of Chomsky's formal education (see specific citations in Lobina 2014), the mathematical concept of recursion was quasi-synonymous with computability, so that RECURSIVE was considered equivalent to COMPUTABLE. In fact, in Chomsky's original formulation, a generative grammar is nothing but a recursive definition of a specific set (the sentences which it generates). The notion of recursive definition was related to the fact that Post had shown with his production systems how infinite sets could be generated/computed from finite axioms (see Lobina 2014 for a discussion and references). And this is precisely what Chomsky was interested in to capture the well-known property of discrete infinity, or as Humboldt expressed it, the infinite use of finite means.

Therefore, what Chomsky (both before and after his influential 2002 paper with Hauser and Fitch) postulates as the central characteristic of human language is recursion in the computational sense, not the existence of sentences within sentences or the existence of noun phrases inside noun phrases. What is predicted is that in all languages an unlimited number of grammatical expressions can be generated, that is, expressions in which there is a systematic relation between sound and sense, hence, which are computable.

So, when linguists such as Everett (among others) try to prove that there is not a recursive computational system underlying the structure of languages, using the argument that there are no self-embedded structures (in a particular language or in any language) they are shooting at the wrong target.[7] Thus it matters not whether they hit their target, that is, whether there are indeed languages without self-embedded subordination or not. The question may be a debatable one, but it is irrelevant in the present context.

The phrase structure and transformational rules of the Standard Model were replaced by the MERGE operation in the minimalist model. And it is recursion in the sense I have described above that Merge preserves. Thus, Merge is an operation (function) that takes two syntactic objects (W and U) and joins them

---

[7] Things do not improve in later approaches (e.g. Everett 2010), where recursion goes on being interpreted as embedding. Furthermore, Everett's approach shows he considers grammar as external to the mind: 'My conclusion is revealed in my title ["You drink. You drive. You go to jail"]. The recursive interpretation of the first three clauses is imposed on them by the mind, not the grammar' (Everett 2010: 1).

forming W (W, U), which in turn is the input of Merge, which produces another syntactic object (say Z), as shown below (M represents the Merge function):

(1)   X = M (X, Y), where Y = M (Y, Z), where Z = M (Z, W), where W = M (W, U), etc.

What we see, roughly, is that Merge is repeatedly applied from right to left to generate the object X, and that the object X is defined recursively by the application of Merge to the lower syntactic objects, which have been also formed by Merge. The structure created by this process would be that which is represented in Figure 1.
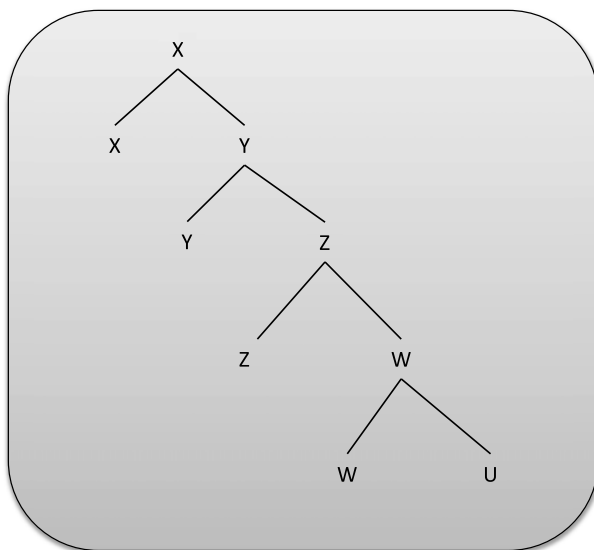

Figure 1
Recursive iteration of Merge produces the hierarchical syntactic object X.

What any claim which seeks to definitely liquidate the minimalist FL needs to show is that the syntactic structure of human languages do not have the hierarchical configuration of constituents represented in the scheme, but that it is purely linear or sequential, i.e. something like this: X = Y, Z, W, U, etc.

As I have already noted, C&C's attack on the minimalist FL is more radical than that of Everett because they do not signal a possible counter-example to the alleged universality of self-embedded subordination, but suggest that the syntax of human languages is sequential. (And, from a generativist point of view, this is even worse.)

C&C propose that 'our limited ability to deal with recursive structure in natural language is an acquired skill, relying on non-linguistic abilities for sequence

875

learning' (Christiansen & Chater 2015: 2). C&C imply that if our ability to process syntactic structure is based on our ability to learn sequences, then syntactic structure is essentially linear. In fact, they argue (using genetic and neurological data) that humans have developed a specific capacity to learn complex sequences, and they suggest that this capacity is what would be behind our language ability, making it unnecessary to postulate a UG, even a minimal one:

> Hence, both comparative and genetic evidence suggests that humans have evolved complex sequence learning abilities, which, in turn, appear to have been pressed into service to support the emergence of our linguistic skills. (Christiansen & Chater 2015: 4)

It is surprising to note in this case how easily a specific evolutionary adaptation to learn sequential order is accepted, although this is another matter. Anyway, there seems to be no objection to the idea that humans have developed more complex capacities to learn sequences than other organisms, and it would be unwise to deny that such a capability is particularly useful in learning and using languages. This is so because human languages have an obvious sequential dimension: we speak producing linear chains of words and learn to speak hearing (and processing) linear chains of sounds. What we can object to (and which can even be demonstrated to be false) is the assumption that the syntax of human languages is linear.

Before turning to this, it is also important to note that by confusing the syntactic structure with the linear sequence that represents it, C&S (and many others) also confuse grammar with processing. In fact, it is not really a confusion, as C&C propose that grammar is unnecessary, and that processing is enough to explain the structure of languages.[8] Their argument is based essentially on very solid experimentation, both their own and that of others, which in large part involves neural network simulations and behavioural contrasts with humans, and which shows that artificial neural networks are able to evolve to basically replicate the limitations that humans have to process 'recursive structures'. Because they assume they have shown that artificial neural networks mimic human capabilities, and given that neural networks do not develop an internal recursive system, they then conclude that there is not an internal recursive system in humans.

The problem with this conclusion is that the processing of linear sequence of words is only part of language processing, precisely because syntax is not linear, but hierarchical.

Let us consider a typical example of what would be a 'recursive structure' in which both humans and trained neural networks show the same pattern of inefficiency, such as so-called central subordination constructions (Christiansen & Chater 2015: 10):

---

[8] But this implies that if we consider Marr's (1982) classic tri-level hypothesis for the analysis of any cognitive system (implementation, representation/algorithm, and computation), we could simply eliminate the computational level, as if the mathematical formula of the function that a computer is calculating is irrelevant to understanding how software and hardware work.

876

(2)   The chef who the waiter who the busboy offended appreciated admired the musicians.

On a first reading any native English speaker would consider sentence (2) unintelligible and would judge it as ungrammatical in an evaluation experiment. However, it is a perfectly grammatical sentence (read it slowly a few times, set yourself some 'mental commas', and its meaning will emerge like a faint glow in the darkness).

Miller & Chomsky (1963) considered these cases as vivid examples of the difference between competence and performance. Although the grammar generates that sentence, people would not use it because it requires a lot of processing effort. The existence of grammatical sentences that are difficult to process is, in fact, a direct proof of the existence of a linguistic knowledge that is independent of language use in real time (and is also an argument that the internal syntax did not arise for communication, a point that I will discuss later). C&C object that if there is an innate recursive competence, there should be no processing difficulties, and that to allege problems of working memory overload is inadequate, since there are people who do better than others, and since in some languages such expressions are more used than in others. This is a reasonable attitude if one considers that human syntax consists of putting words in an essentially flat, linear order, and that the human syntactic ability is a refinement of the capacity of learning sequences. And, indeed, processing problems in sentence (2) are related to linear order. If we now alter the sentence, moderating argument movements, we obtain version (3), which is far more reader-friendly (yet with the same underlying hierarchical structure and therefore the same meaning):

(3)   The chef who appreciated the waiter who the busboy offended admired the musicians.

But this allows for two important conclusions. First, the fact that (2) and (3) mean the same while having different order shows that order does not determine the meaning. Second, that in effect the processor has a preference for adjacent grammatical relations (say, linearly 'local'). But note that this is to be expected if we are processing a structure that has been 'flattened', that is, that when 'externalized' it has become a linear sequence of words. If we use our ability to recognize sequences as part of our language processing, it is expected that (2) involves more processing cost than (3), precisely because in (2) many of the relevant syntactic relations are 'interrupted' by other elements that do not allow us to 'close' the derivation and, plausibly, working memory is overloaded. Therefore, the fact that syntactic processing is sensitive to factors related to linear order does not imply that the syntax is linear.

In addition, C&C themselves report an interesting fact in this context of the distinction between, on one hand, the formal properties of expressions generated by the computational system and, on the other hand, the linear processor employed in its use: both trained neural networks and humans (speakers of English) tend to

consider example (4), below, to be more acceptable (although it is ungrammatical) than (2), although (4) is the result of removing a verb (*appreciated*) from (2), forming a non-sentence (from Christiansen & Chater 2015: 10):

(4)   *The chef who the waiter who the busboy offended frequently admired the musicians.

It is understandable that the absence of a verb in (4) relieves processing overload, but what then becomes clear is that what such experiments are measuring is the 'linear processability' and not grammar; after all, (2) has coherent structure and sense, while (4) lacks both. Interestingly, C&C report that German speakers do not show such a preference for (4) (against (2)), since they are used to processing sentences with systematic alterations in linear adjacency (as, for example, in German *dass Ingrid Peter Hans schwimmen lassen sah*, lit.: '(that) Ingrid Peter Hans saw let swim', i.e. *that Ingrid saw Peter let Hans swim*). C&C use this crosslinguistic variability in processing capacity to strengthen their hypothesis that the ability to process 'recursive structures' depends on practice and habit, and that therefore it is not innate, thus there is no FL.

However, such a conclusion does not hold. What, if anything, these facts do show is that the sequential processing ability can be trained, which in no way rules out the existence of a recursive computational system. The recursive computational system (syntax) generates a potentially infinite number of hierarchical grammatical structures, as shown in Figure 2. Such structures, again shown in Figure 2, are converted into linear sequences during the externalization of language for communication. And once externalized, some of these structures are more or less difficult to process than others by our linear processing capacity.
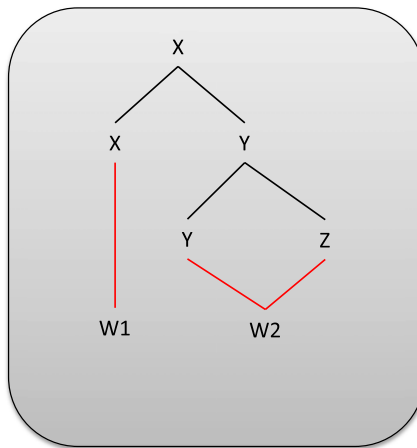

Figure 2
A syntactic hierarchical structure flattened as a sequential string of words.

878

According to our previous experience and habits, we will have a greater or lesser ability to do so, as any reader can see by returning to example (2).

So, the important question is: can the processing of linear sequences account for the syntactic structure of human languages? We already know C&C's answer, which leads them to believe that the structural complexity of the syntax of human languages is a stipulation, not an empirical fact:

> From our *usage-based* perspective, the answer does not necessarily require the postulation of recursive mechanisms as long as the proposed mechanisms can deal with the level of complex recursive structure that humans can actually process. In other words, what needs to be accounted for is the *empirical evidence* regarding human processing of complex recursive structures, and not *theoretical presuppositions* about recursion as a stipulated property of our language system. (Christiansen & Chater 2015: 7)

The crucial error in this approach is, as I anticipated, the lack of distinction between, on the one hand, the linguistic expressions we issue (typically in sequential strings of words, such as those the reader has in front) and, on the other hand, the (inaudible and invisible) syntactic structure that determines why the expressions mean what they mean and nothing else.

It may seem unnecessary to highlight these facts, but clearly it is. For much of the profession, the subtle and complex 'subatomic structure' that underlies linguistic expressions is simply going unnoticed, and this has the potential to be a serious obstacle for the integration of linguistics into the main body of cognitive science.

Perhaps as a reaction to the perception of this situation, at almost the same time that the paper by C&C appeared, an article by Everaert, Huybregts, Chomsky, Berwick and Bolhuis was published with the descriptive title 'Structures, not strings' (Everaert et al. 2015). The reasons for a group of authors very much representative of current generative grammar (including Chomsky himself) to write an almost trivial summary of the research in their field are clear:[9]

> [T]aking language as a computational cognitive mechanism seriously, allow us to address issues left unexplained in the increasingly popular surface-oriented approaches to language. (Everaert et al. 2015: 729)

Let us consider just one example of what, according to the authors, remains unexplained in 'surface-oriented approaches to language':

(5)   The book I bought did not appeal to anybody.

In (5) we can see that the negative polarity item *anybody* has to be preceded by *not*. Otherwise, the sentence is ill-formed:

(6)   *The book I bought appeared to anybody.

---

[9] Everaert et al. (2015) is published in the journal *Trends in Cognitive Science*, which shows that here in some sense the medium is the message.

We could state that the condition to use *anybody* in this sentence is that before *anybody* there must be a *not*. But then the following, similar example from Japanese (from Everaert et al. 2015: 732) would be a counter-example:

(7) Taroo-wa nani-mo tabe-nakat-ta.
     'Taro didn't eat anything.'

(8) *Taroo-wa nani-mo tabe-ta.
     *'Taro ate anything.'

Example (8), parallel to (6), shows that if there is no negation (-*nakat*-), the negative concord item (*nani-mo*, in this case 'anything') cannot be used. But (7) is right, and there it is *nani-mo* 'anything' which precedes the negation (-*nakat*-), so we cannot generalize the explanation. Of course, we could suggest that in English *not* must precede *anything*, but that in Japanese this is reversed. However, what matters is that linear order is insufficient to explain the contrast:

(9) *The book I did not buy appealed to anybody.

In (9) we note that although *not* precedes *anybody*, the sentence is ungrammatical. We could say then that the condition is that *not* has to modify the verb whose complement is *anybody*, but then we are no longer talking about linear order, but about hierarchy. The condition for using *anybody* has to do with the structural position (in a tree like that in Figure 2) in which *not* is found, and not with precedence. The basic idea is that *not* cannot license *anything* if it is included in a constituent from which it is not a 'structural brother' of the constituent including *anybody* (what in the jargon is called C-COMMAND). Thus, the *not* in (9) is 'stuck' in the relative clause *I did not buy*, and does not have the necessary structural relationship with *anybody*.

What is important is that the interpretation of linguistic expressions, although they are materialized linearly, is sensitive to structure, and not to linear order (i.e. that rules are 'structure dependent', in the classic formulation). The proof of this is that in Japanese (a language that orders constituents differently than in English) the conditions are the same, precisely because they depend on the hierarchical structure of human syntax, and not on how that structure is 'flattened' to be issued or processed by the motor system that produces sounds chains.

In my opinion, one of the key achievements of GG (and especially of the MP) has been to show that the connection between the computational system (syntax) and the conceptual-intentional system (CI) is different from the connection between the computational system and the sensorimotor system (SM), responsible for the externalization of linguistic expressions in physical signs. More specifically, Chomsky has argued (since 2007 at least) that it is possible that the human computational system evolved/is optimized for the conceptual-intentional system. Thus, the first use of language would have been an internal system of thought, that is, an 'internal language of mind' capable of combining

conceptual elements in new ways to create more complex concepts (and hence thoughts). The later connection to the sensorimotor system for externalization would therefore be secondary or ancillary. The relationship between the essential components of FL is therefore asymmetric. In the words of Everaert et al. (2015: 741):

> The asymmetry is: the mapping to meaning is primary and is blind to order (language as a system for thought), the mapping to sound/sign is secondary and needs order (imposed by externalization of language). The empirical claim is, therefore, that linear order is available for the mapping to sound/sign, but not for the mapping to meaning.

What this implies is that the mental computations in the creation and understanding of linguistic expressions (and of thought in general) are blind to the linear arrangement of the words that are articulated or perceived by the input and output systems of the interface with the sensorimotor system; put more simply, that linear order is irrelevant to semantics and syntax. Linear order is a secondary feature, imposed through language externalization when used for communication. In terms of Everaert et al. (2015: 740), for the mapping to the CI interface, hierarchical structure is necessary and sufficient, and linear structure is irrelevant; that is, order is inaccessible. But for the mapping to the SM interface, hierarchical structure is necessary, but not sufficient, and linear structure is relevant; that is, order is needed for externalization. In the form of a take-home message, they conclude: 'What reaches the mind is unordered, what reaches the ear is ordered' (Everaert et al. 2015: 740).

Undoubtedly, many aspects of this view may be wrong or may need adjustment and more experimentation, but it is unquestionable that the identification of syntax with linear order is simply the fruit of ignorance of a branch of science, the one we call syntactic theory.

The publication of the article by Everaert et al. (2015) in a high-impact journal may seem strange, in the sense that any undergraduate student of linguistics may have been familiar with its content. Yet as these authors say:

> These somewhat elementary but important insights have been recognized since the very origins of generative grammar, but seem to have been forgotten, ignored, or even denied without serious argument in recent times. (Everaert et al. 2015: 742)

There are many ways to do linguistics, and all of them have something important to contribute. One may be interested in syntax or not, but it is not wise to pretend it does not exist.

7. AN INCOMPLETE VIEW OF LANGUAGE(S) IS THE SOURCE OF OUR PROBLEMS

The most important consequence of denying or ignoring the structure of languages is the emergence of an incomplete and biased view of what language is, and of what languages are. In fact, only a deep disagreement on what is a human

881

language can explain the current situation of our field. It is a real (and somewhat surprising) fact that a scientist who comes to current theoretical linguistics from outside the discipline will encounter a fragmented and often contradictory field. On reading Chomsky or Pinker she will be led to conclude that human language is a mental organ that is characteristic of the species, and that therefore it should be addressed scientifically like any other organ of the body. She will also be told that all human languages are superficially different variations of a single system of knowledge, which is uniform within the species, since it has a narrow natural conditioning. In addition, she will read that language is essentially innate, and that the development of a language in the child's mind and brain is more comparable to the growth of tissue or the development of the visual system than to the learning of rules and cultural conventions. However, if the same reader approaches the study of language from the perspective of the functionalist-cognitive tradition (if she reads, for example, Levinson or Tomasello), she will be led to believe that human languages are profoundly different, given that each human language is a different cultural solution to the same set of communicative and cognitive challenges. She will learn that the structure of each language is extracted from the environment by children, along with knowledge about the rest of their culture, as a result of exposure to stimuli, and that in doing so children use the same general purpose learning systems that are used for the development of other systems of belief and knowledge. She will even read that each language tends to structure thought differently.

My conclusion is that the main source of such a radical discrepancy is a fundamental disagreement about what we mean when we say LANGUAGE. To understand this, let us consider again the general model of the FL by Hauser et al. (2002) in which it is formed by the computational system, the conceptual-intentional system (CI), and the sensorimotor system (SM) (see Figure 3).
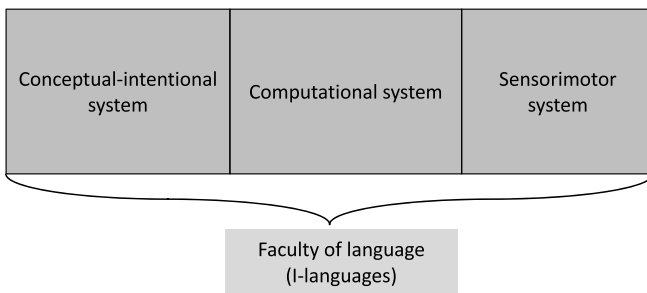

Figure 3
The essential components of the Faculty of Language according to Hauser et al. (2002).

It is clear that this system is too general to serve as a representation of I-languages (recall that an I-language is the FL of each person), in that it does not capture their obvious historical nature and, therefore, their diversity. To put

882

it in clearer terms: the scheme provides an overview of the essential architecture of FL, but does not allow us to grasp what distinguishes each natural language, and that arises from the linguistic and cultural environment in which each person develops.

We saw at the end of the previous section that Chomsky and others have suggested that the relationship between CS and CI and SM systems is asymmetrical, in the sense that CS 'may be optimized relative to the CI interface, with mapping to SM an ancillary procedure, and complex to the extent that SM has no prior adaptation to these needs' (Chomsky 2007: 14). The connection between SC and the CI system would imply that

> the earliest stage of language would have been just that: a language of thought, used internally . . . while phonology, morphology, and whatever else is involved in externalization might be variable and complex and subject to large-scale historical accident, satisfying the linking condition in ways that are as good as possible. (Chomsky 2007: 13, 14–15)

What this scenario implies, then, is that FL (i.e. every human I-language) should also include a component derived from the environment (that is, internalized), whose mission would be to systematically connect the derivations generated by the 'language of thought' (resulting from the interaction between the computational and the CI systems) with sensorimotor systems. The central idea is that this component, which is exposed to change (as it is internalized from the environment), is the one that really reflects the differences between languages. For expository convenience I will call that component the LEXICAL INTERFACE. The use of the expression 'lexical interface' is based on the traditional idea that the lexicon of a language systematically and arbitrarily pairs senses and meanings, but we must be clear to avoid a reading in which the 'lexicon' is the set of words or units that syntax combines to create sentences.[10]

For our present purposes we can imagine this lexical interface as a set of morphological and phonological exponents (words) that are internalized from the environment during the process of language acquisition. The speaker uses these exponents or words to linearize/process parts of the hierarchical structures generated by the computational system (see again Figure 2). The lexical interface can be conceived of as a component of long term memory that provides a stable connection between, on one side, the conceptual elements and structures constructed by the computational system interacting with CI, and, on the other, the sensory-motor systems that process and produce the physical signals human beings use when they use language for communication (and perhaps also for some modalities of thought). The previously mentioned asymmetry can then be represented as in Figure 4.

---

[10] In general, this approach is concomitant with insights from Distributed Morphology (Halle & Marantz 1993, Embick 2015) and Nanosyntax (Starke 2009).
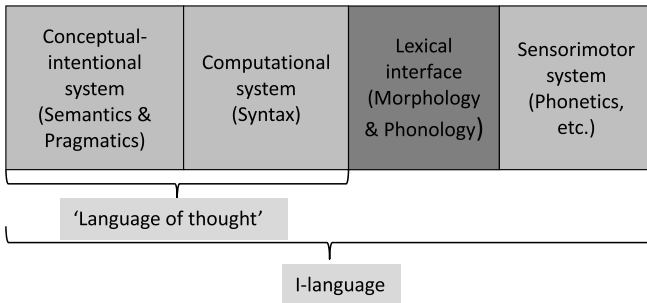
883

Conceptual-intentional system (Semantics & Pragmatics)	Computational system (Syntax)	Lexical interface (Morphology & Phonology)	Sensorimotor system (Phonetics, etc.)

'Language of thought'

I-language

Figure 4
The essential anatomy of an I-language.

According to the scheme of Figure 4, any I-language, as long as it is a person's FL, is made up of all four components (for the sake of clarity I have indicated in each component the field of traditional grammar with which it would be roughly associated). Three of them are internal to the organism and would be very highly conditioned by our biology and by other (3rd type) factors (and, thus, are essentially uniform), while the lexical interface (highlighted in a darker grey), having been internalized from the environment, is subject to variation and to historical change.

Note that, of course, the possible outcomes ('lexical interfaces') of the connexion between the computational system and the sensorimotor system (externalization) are highly restricted by the computational system itself, and by the common format (to all humans) of the sensorimotor system, which severely limits the range of variation in this domain.[11]

From this general perspective, it is tempting to consider that the points of deep disagreement we have seen have their roots in a misunderstanding of what a language is. According to the point of view reflected in Figure 4, every I-language includes all three components of FLB (CIS, CS and SMS), and also the 'lexical interface' that singles out each one because of its historical nature. But it seems that from the externalist point of view that underlies current functionalist and cognitivist approaches, there is a misidentification of human languages with their 'lexical interfaces' (the dark grey box of Figure 4). My point (see also Mendívil-Giró 2014) is that if we (mis)identify an I-language with its 'lexical interface',

---

[11] As an anonymous *JL* referee suggests, some of the best-motivated universals have to do with the system of morpho-phonological realization. He/she mentions Zwicky's (1977) generalization (according to which the inclusive is frequently syncretic with the exclusive, but rarely if ever with the second person plural), the absence of so-called ABA patterns in various morphological paradigms (Bobaljik 2012), and the strong asymmetry between 1–2 person and 2–3 person syncretism on the one hand, and 1–3 person syncretism on the other (see Harbour 2016). Caha's (2009) approach to case syncretism is another influential example.

884

then the disparity of opinions we have on UG, language, and even linguistics itself, follows rather naturally.

Note that a 'lexical interface' is indeed a set of words and constructions, it lacks recursive structure, and is learned (at least in part statistically) from the environment. Thus, if one identifies languages with their 'lexical interfaces' it follows naturally that linguistic structure is sequential (Christiansen & Chater 2015), and that universals are a myth, so that 'structural differences should instead be accepted for what they are, and integrated into a new approach to language and cognition that places diversity at centre stage' (Evans & Levinson 2009: 429). From this incomplete view of languages it also follows quite reasonably that there cannot be a universal theory of grammatical categories (Haspelmath 2007), or that functional categories and even syntax are the historical consequence of linguistic change, that is, grammaticalization theory (Heine & Kuteva 2007). If we return now to Tomasello's (2009: 471) claim that common features of languages 'come not from any universal grammar, but rather from universal aspects of human cognition, social interaction, and information processing – most of which were in existence in humans before anything like modern languages arose', we can note that he is explicitly subtracting 'universal aspects of human cognition' from languages, in what seems to be a clear misidentification of I-languages with their external surfaces (i.e. 'lexical interfaces').

It is important to recognize that whereas the dark grey section of Figure 4 is something external and cultural (although internalized), it is not A LANGUAGE. A language, as a system of knowledge, includes all of the components, not only the component derived from the linguistic environment. It might perhaps be argued that such a separation is legitimate if one wants to study 'language' and not 'general cognition', but its legitimacy is easily called into question. Let us look once more at biology in order to understand why: humans and chimpanzees share over 99% of DNA, but we are not tempted to say that this 1% of specifically human DNA is a human being or can give rise to a person; we require all 100%.

8. CONCLUSION: WE NEED COMPLEMENTARITY, NOT RIVALRY

If my diagnosis of why the division within theoretical linguistics persists is correct, then there is a glimmer of hope. If we think of the controversy in terms of the point of view raised in the previous section, we can conclude that there is actually much more complementarity than incompatibility between the findings and results of the two major research frameworks.

For example, interpreting the word *language* as 'lexical interface' would render a good part of the cognitivist literature much more acceptable (and interesting) for a generative linguist. Similarly, interpreting the word *language* as 'something that integrates a cognitive system deeper than the surface forms of languages' could perhaps make generative literature more attractive to cognitivist linguists.

885

Anyway, the recent trend, illustrated in this contribution, of presenting to the scientific community the approaches of another research tradition by misinterpreting its statements and denying or ignoring its results, although perhaps in some way rewarding for those who do it, is very negative for the science of language, and does nothing to improve its potential impact within the rest of cognitive science and indeed within science in general.

It is neither healthy nor productive for the science of language that we, as a profession, engage in false polemics and become distracted from developing our many different specific areas of linguistic research, all of which are worthy of serious investigation. As noted by Saussure a century ago (Saussure 1916), human language is a complex phenomenon involving biology, neurology, philosophy, computation, culture, communication, social and economical relations, and much more. Generative Grammar is a branch of cognitive science specialising in the study of the formal structure of human languages. One can decide to cultivate this discipline or not, but it is not necessarily inconsistent or incompatible with other possible approaches.

So, is UG ready for retirement? Well, yes and not. The term is very old and can lead to confusion. Yet in itself it is no more than a technical term, with no literal sense, like atoms or sunsets. I get the feeling, though, that its rejection is more ideological than scientific.

REFERENCES

Benítez-Burraco, Antonio & Víctor M. Longa. 2010. Evo-Devo – Of course, but which one? Some comments on Chomsky's analogies between the biolinguistic approach and Evo-Devo. *Biolinguistics* 4, 308–323.

Bergen, Benjamin K. 2014. Universal Grammar. http://edge.org/response-detail/25539 (accessed 10 January 2016).

Bobaljik, Jonathan David. 2012. *Universals in comparative morphology*. Cambridge, MA: MIT Press.

Boeckx, Cedric, María Carmen Horno & José-Luis Mendívil-Giró (eds.). 2012. *Language, from a biological point of view*. Newcastle: Cambridge Scholars Publishing.

Caha, Pavel. 2009. *The nanosyntax of case*. Ph.D. dissertation, University of Tromsø.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1966. *Cartesian linguistics*. New York: Harper & Row.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. *The architecture of language*. New Delhi: Oxford University Press.

Chomsky, Noam. 2004a. Beyond explanatory adequacy. In Adriana Belletti (ed.), *Structures and beyond*, 104–131. Oxford: Oxford University Press.

Chomsky, Noam. 2004b. *The generative enterprise revisited*. Berlin: Mouton de Gruyter.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36, 1–22.

Chomsky, Noam. 2007. Approaching UG from below. In Uli Sauerland & Hans-Martin Gärtner (eds.), *Interfaces + recursion = language? Chomsky's minimalism and the view from semantics*, 1–30. Berlin: Mouton de Gruyter.

Chomsky, Noam & Howard Lasnik. 1993. The theory of Principles and Parameters. In Joachim Jacobs, Armin von Stechow, Wolfgang Sternefeld & Theo Vennemann (eds.), *Syntax: An international handbook of contemporary research*, 506–569. Berlin: Walter de Gruyter.

Christiansen, Morten H. & Nick Chater. 2015. The language faculty that wasn't: A usage-based account of natural language recursion. *Frontiers in Psychology* 6, 1182. doi:10.3389/fpsyg.2015.01182.

Dąbrowska, Ewa. 2015. What exactly is Universal Grammar, and has anyone seen it? *Frontiers in Psychology* 6, 852. doi:10.3389/fpsyg.2015.00852.

886

Embick, David. 2015. *The morpheme: A theoretical introduction*. Boston, MA & Berlin: De Gruyter Mouton.

Evans, Nicholas & Stephen C. Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32, 429–448.

Evans, Nicholas & Stephen C. Levinson. 2010. Time for a sea-change in linguistics: Response to comments on 'The myth of language universals'. *Lingua* 120, 2733–2758.

Evans, Vyvyan. 2014. *The language myth*. Cambridge: Cambridge University Press.

Everaert, Martin B. H., Marinus A. C. Huybregts, Noam Chomsky, Robert C. Berwick & Johan J. Bolhuis. 2015. Structures, not strings: Linguistics as part of the cognitive sciencies. *Trends in Cognitive Sciences* 19.12, 729–743.

Everett, Daniel L. 2005. Cultural constraints on grammar and cognition in Pirahã. *Current Anthropology* 46, 621–646.

Everett, Daniel L. 2010. You drink. You drive. You go to jail. Where's recursion? http://ling.auf.net/lingbuzz/001141 (accessed 16 September 2016).

Everett, Daniel L. 2016. An evaluation of Universal Grammar and the Phonological Mind. *Frontiers in Psychology* 7, 15. doi:10.3389/fpsyg.2016.00015.

Fitch, W. Tecumseh. 2009. Prolegomena to a future science of biolinguistics. *Biolinguistics* 3, 283–320.

Goldberg, Adele. 2015. Strong evidence that the roots of binding constraints are pragmatic from Cole et al. (2015). http://dlc.hypotheses.org/865 (accessed 10 January 2016).

Halle, Morris & Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In Kenneth Hale & Samuel J. Keyser (eds.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, 111–176. Cambridge, MA: MIT Press.

Harbour, Daniel. 2016. *Impossible persons*. Cambridge, MA: MIT Press.

Haspelmath, Martin. 2007. Pre-established categories don't exist: Consequences for language description and typology. *Linguistic Typology* 11, 119–132.

Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The Faculty of Language: What is it, who has it, and how it evolved? *Science* 298, 1569–1579.

Heine, Bernd & Tania Kuteva. 2007. *The genesis of grammar*. Oxford: Oxford University Press.

Jackendoff, Ray. 2002. *Foundations of language*. Oxford: Oxford University Press.

Joly, André & Jean Stéfanini (eds.). 1977. *La Grammaire Générale. Des modistes aux idéologues*. Lille: Publications de l'Université de Lille.

Kauffman, Stuart A. 1993. *The origins of order: Self-organization and selection in evolution*. London: Oxford University Press.

Lieberman, Philip. 1984. *The biology and evolution of language*. Cambridge, MA: Harvard University Press.

Lieberman, Philip. 2013. *The unpredictable species: What makes Humans unique*. Princeton, NJ: Princeton University Press.

Lobina, David J. 2014. What linguists are talking about when talking about... *Language Sciences* 45, 56–70.

Longa, Víctor M. & Guillermo Lorenzo. 2012. Theoretical linguistics meets development: Explaining FL from an epigeneticist point of view. In Boeckx et al. (eds.), 52–84.

Lyons, John. 1991. *Chomsky*. London: Fontana Press.

Malmberg, Bertil. 1991. *Histoire de la linguistique. De Sumer à Saussure*. Paris: P.U.F.

Marr, David. 1982. *Vision: A computational investigation into the human representation and processing of visual information*. New York: Freeman.

McMurray, Bob & Edward Wasserman. 2009. Variability in languages, variability in learning? *Behavioral and Brain Sciences* 32, 458–459.

Mendívil-Giró, José-Luis. 2012. The myth of language diversity. In Boeckx et al. (eds.), 85–133.

Mendívil-Giró, José-Luis. 2014. What are languages? A biolinguistic perspective. *Open Linguistics* 1, 71–95. doi:10.2478/opli-2014-0005.

Miller, G. A. & Noam Chomsky. 1963. Finitary models of language users. In R. D. Luce, R. R. Bush & E. Galanter (eds.), *Handbook of mathematical psychology*, vol. 2, 419–492. New York: Wiley.

Moreno, Juan-Carlos & José-Luis Mendívil-Giró. 2014. *On biology, history and culture in human language: A critical overview*. Sheffield: Equinox.

Rooryck, Johan, Neil V. Smith, Anikó Liptak & Diane Blakemore. 2010. Editorial introduction to the special issue of *Lingua* on Evans & Levinson's 'The myth of language universals' In Johan Rooryck, Neil V. Smith, Anikó Liptak & Diane Blakemore (eds.), *The myth of language universals*: Special issue of *Lingua* 120, 2651–2656.

Saussure, Ferdinand de. 1916. *Cours de linguistique générale*. Paris: Payot.

Starke, Michal. 2009. Nanosyntax: A short primer to a new approach to language. *Nordlyd* 36, 1–6.

Tomasello, Michael. 2009. Universal Grammar is dead. *Behavioral and Brain Sciences* 32, 470–471.

Uriagereka, Juan. 1998. *Rhyme and reason: An introduction to minimalist syntax*. Cambridge, MA: MIT Press.

Zwicky, Arnold. 1977. Hierarchies of person. In Woodford Beach, Samuel Fox & Shulamith Philosph (eds.), *Papers from the Thirteenth Annual Meeting of the Chicago Linguistic Society*, 714–733. Chicago, IL: Chicago Linguistic Society.

Author's address: *Department of General and Hispanic Linguistics, University of Zaragoza,*
C/ Pedro Cerbuna 12, E-50009 Zaragoza, Spain
jlmendi@unizar.es