Nature and Culture in Language = Syntax and Lexicon in Languages
Trying to reconcile an old controversy in the theory of language*

José-Luis Mendívil Giró
jlmendi@unizar.es
University of Zaragoza
2018

Abstract
This contribution identifies a clue that might lead us towards a better understanding of how the weight of natural and cultural factors is distributed in the design of human languages. First, I propose a specific model of the relationship between lexicon and syntax in human language. This model allows us to use the strategy of correlating the general question of what is natural and what is cultural in language with the nature of the various components that make up a language. Thus, I propose the following correlation: the cultural dimension of language would be expressed essentially in the lexical component of languages, while the biological/natural dimension is expressed in the syntactic component. Such a strategy also makes it possible to better understand why different traditions in contemporary linguistic theory arrive at such different assessments of the role of natural factors in language design.

Keywords
Biology of language, Biolinguistics, Lexicon, Syntax, Nature/Nurture Debate

1 Introduction

It is commonly accepted that human language has both a natural and a cultural dimension. The theoretical controversy that characterizes modern linguistics tends to focus, more than on the absolute denial of one of these two dimensions, on the greater or lesser emphasis that is put on each of them when developing language research programs. This contribution proposes an alternative—and somewhat conciliatory—strategy to address the delimitation between the natural and the cultural dimensions of human language. Such an alternative is based on the intuition that there could be a correlation between, on the one hand, the natural and cultural factors in human language and, on the other, the various components that have been recognized in human languages.

Thus, starting from a vision of human language as a specifically human mosaic of cognitive capacities (supposedly) present in other species and/or in other human cognitive spheres, I propose an explicit model of the relationship between lexicon and syntax in languages according to which syntax is universal (invariable in time and space) and, therefore, a solid candidate for representing the natural conditioning for language, while the lexicon would reflect the historical and cultural dimension of human languages.

---

* This paper is a translated and expanded version of a talk in Spanish, collected in the proceedings of the *XIII Congreso Internacional de Lingüística Xeral* (available at http://cilx2018.uvigo.gal/actas/). If you are an editor and you think it might be interesting to publish this text in English, please contact the author.

To the extent that this model of the relationship between lexicon and syntax in human languages is empirically correct, it can be considered an argument in favor of the conception of language from which it derives and, crucially, a way of clarifying old and noxious controversies in linguistic theory.

2 Descartes' error

Denying a natural, biological dimension to human language would be the same as repeating Descartes' dualistic error. This means that it does not make sense to think that language is not an attribute of the human brain such as memory, vision or emotions. As Antonio Damasio (who will undoubtedly have come to the reader's mind when I mentioned *Descartes' error*) pointed out, the human brain, as an organ of the body, is a subtle mixture of innate dispositions and development through experience:

"At birth, the human brain comes to development endowed with drives and instincts that include not just a physiological kit to regulate metabolism but, in addition, basic devices to cope with social cognition and behavior. […] Yet there is another role for these innate circuits which I must emphasize because it usually is ignored in the conceptualization of the neural structures supporting mind and behavior: *Innate circuits intervene not just in bodily regulation but also in the development and adult activity of the evolutionarily modern structures of the brain*." (Damasio 1994: 126, 110, original emphasis).

Modern neuroscience has shown that memory, vision and emotion are capacities that cannot be explained without an innate bias in the development of the brain tissues that make them possible. I think that in the context of modern cognitive science, assuming that language is an exception might well be considered surprising, and indeed suspicious. As Damasio himself has pointed out, "It is not only the separation between mind and brain that is mythical: the separation between mind and body is probably just as fictional" (1994: 118).

However, the traditional claim that language is a social institution (Saussure 1916), rekindled in recent decades by the insistence on the essentially external, cultural nature of languages, seems to point in that direction. Thus, for example, Christiansen and Chater, in a relatively recent paper entitled *The language faculty that wasn't*, state: "It is time to return to viewing language as a cultural, and not a biological, phenomenon" (Christiansen & Chater 2015: 14). This conclusion is based on an empiricist and, in a certain way, dualistic vision of human cognition. Such is the case of Deacon's (1997) influential theory of the coevolution between language and brain. So, from this point of view "in some ways it is helpful to imagine language as an independent life form that colonizes and parasitizes human brains, using them to reproduce" (Deacon 1997: 111). But note that it is then being assumed that language is external to the brain (precisely what the expression *coevolution of language and brain* implies). Deacon's position, widely followed in contemporary functional and so-called cognitive linguistics, reflects an externalist vision of languages as cultural or social objects that "settle" in brains, thus contributing to their organization. This model corresponds to an empiricist vision of the mind and brain, as opposed to the rationalist point of view, which posits that the essential structure of languages comes from the structure of the brain and the mind, and

not the other way around. This also explains why authors of this orientation insist so greatly on the depth of the diversity of languages and reject the unifying effect that a biologically conditioned language faculty would have in languages (see Evans and Levinson 2009 for a programmatic vision in this regard).

3 A brief dissection of a false controversy

However, the faculty of language, whose initial state (prior to experience) is called Universal Grammar (UG) in the Chomskyan generative tradition, exists by definition. As Chomsky has pointed out:

"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're all learn English" (Chomsky 2000: 50).

Note that such a statement, despite how it may appear, is not a simplification (or a provocation). It simply insists on a crucial idea that has been misinterpreted: that the capacity of language is innate does not imply that there must be language genes or linguistic neurons, but rather that human beings have a unique ability to acquire the languages of their environment. As Fitch (2009) points out in developing this argument, since immersion in a linguistic environment is not enough for language to develop in nonhuman (natural or artificial) organisms, then there must be *something* in human children that differentiates them from other organisms. That *something* is precisely the object of study of this linguistic tradition. The relevant question, then, is not whether UG exists, but what its properties are, and from where they are derived (see Mendívil-Giró 2018).

Those who reject the existence of UG argue that the development and use of language can be explained by adducing general principles of human cognition that are not specific to language. Thus, for example, Michael Tomasello, one of the champions of this position, rejects the existence of a human faculty of language and stipulates that the restrictions that the human brain can impose on the structure and nature of languages would be general and not specifically linguistic:

"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).

But here we are faced with a false problem. Note that Tomasello invokes "universal aspects of human cognition", that is, principles of cognition that are common to all human beings and specific to them. But, since human beings are *the only organisms* that develop knowledge of language, then it is very difficult to differentiate those "universal principles of human cognition" from the Chomskyan notion of UG, since it is defined as the set of principles that they make possible the development of language, that is, as a part of human nature.

What is truly relevant here is the unquestionable fact that some sort of biological conditioning determines the course of language development and the subsequent structure of the knowledge systems that we call human languages. Of course, it is debatable whether the principles that form UG are language-specific or if they are the same ones that underlie other human cognitive systems (and, of course, it is debatable whether the term *universal grammar* is adequate), but this discussion soon becomes sterile in the absence of a detailed specification of what those principles are, and also in the absence of a definition of what language is, and what languages are. And it is precisely here, in the conception of what a language is, that we really find a discrepancy, more than as to whether or not there is a biological conditioning for language.

According to the Chomskyan point of view, a specific language is a particular state of the faculty of language (FL), that is, it is a system of knowledge (usually known as I-language); for the opposite point of view, a language is an external system, a cultural object that the brain is able to assimilate and represent internally. This difference is crucial and is at the root of the different assessment that both traditions make of the relative weight that nature and culture have in explaining the structure of languages.

In my opinion, the error of the externalist approach consists in artificially separating the universal aspects of human cognition from the faculty of language, that is, assuming that there can be no 'universal linguistic aspects of human cognition'. But why should there not be universal linguistic aspects of human cognition, since only humans (and all humans) can learn languages? I think that the rejection of this possibility is related to the fact that the externalist approach operates with an inductive conception of language. In such a conception, language is not a cognitive capacity, but a theoretical construct derived from the comparison of languages and the establishment of generalizations, in a Greenbergian manner. If language cannot be separated from languages, then, paradoxically, the species-specific ability to learn languages has to be considered non-linguistic.

Note that in the previous quote from Tomasello this is clearly expressed: languages, he says, emerged after the existence of these general cognitive conditioning factors, that is, languages are considered as independent cultural phenomena that (more or less) adapt to the format required by human brains and their universal general cognitive principles. Yet from a cognitive (internist) perspective such a view is unsustainable. It would be tantamount to saying that the skin of an animal, for example an elephant, is an external object that adapts to the shape of the elephant's body. Obviously it is true that the shape of the skin depends on the shape of the elephant's body, but this does not permit us to ignore the fact that the elephant's skin is *part* of its body, and not an external object that has adhered and adapted to it.

Thus, for the remainder of this contribution I will try to show that the notion of language used in the externalist tradition is incomplete, and that it is this incomplete vision that causes the inadequate 'dualistic' perception of the relations between the brain (biology) and language (culture).

4. A diagnosis and a proposal

The following table summarizes the main discrepancies regarding the nature of language presented by the two models (identified here as internist versus an externalist views):

Language	Internist view	Externalist view
Origin	Natural / biological	Cultural
Development	Innate	Learned
Location	Internal / Individual	Interiorized / Collective
Variation	Superficial / Universality	Deep / Relativism

I have suggested in the previous section that the source of this discrepancy is an incomplete conception of what language is from the externalist point of view. The vision is incomplete because the possibility of the existence of 'general linguistic principles of human cognition' is excluded. The idea that I want to introduce now is that this biased or incomplete view of language, in addition to being related to an empiricist conception of the mind and the brain, could actually be due to a misunderstanding regarding what the term *language* (both in the mass and in the count meanings) refers to. In my opinion, the externalist tradition inappropriately identifies a language with what is actually a *part* of a language, more specifically, with the 'lexicon' of a language. Of course, for this statement to make sense, the term *lexicon* would have to be defined more precisely (see section 5). But before doing so it is worth considering that if this diagnosis were reasonable, then the same table that serves to illustrate deep discrepancies in the nature of language between these two traditions could also show differences as to the nature of the two main components of languages, as can be seen in the following version, in which only the column titles have been changed:

	Syntax	Lexicon
Origin	Natural / biological	Cultural
Development	Innate	Learned
Location	Internal / Individual	Interiorized / Collective
Variation	Superficial / Universality	Deep / Relativism

In what follows I will present a model of the architecture of the faculty of language which is consistent with what is reflected in the second table, that is, a model according to which the division of labor between the natural/biological and cultural factors in human language matches two different components of the language faculty: syntax and lexicon. One of these, syntax, will be considered to be mainly natural, innate, internal to the mind/brain, common in languages and invariable (universal), while the other, the lexicon, will be considered to be mainly cultural, learned from the environment, internalized but collective, and variable in diverse linguistic communities.

If this model is plausible, it would not only offer a useful guideline to explore the relative weight of biology and culture in human language, but also an explanation for the already long and unsuccessful confrontation between the internist and externalist models themselves. We might say, in a certain way, that once we have clarified what we are talking about when we speak of language, then the two models can be considered to a great extent as complementary, rather than contradictory.

5 The lexicon as an interface for language externalization

According to the influential model proposed by Hauser, Chomsky and Fitch (2002), the human faculty of language (FL) could be conceived of as a complex system minimally composed by three independent components: a conceptual-intentional (CI) system, related to meaning and interpretation, a sensorimotor (SM) system, related to the perception and production of linguistic signals, and a computational system (CS), the syntax in the narrow sense, responsible for the creation of the recursive and productive syntactic structures that underlie linguistic expressions.

The model leaves some interesting open questions, such as how the various components of FL are related to each other and, if they are the result of the species' biological evolution, whether they emerged simultaneously or at different times. In later work (Chomsky 2007, Berwick and Chomsky 2011, 2016) Chomsky has suggested that the relationship between the computational system and the two other components (CI and SM systems) is asymmetric, in the sense that the computational system would have evolved adapting itself to the CI system, forming a kind of 'internal language of thought' (ILoT) aimed essentially at the representation of reality and the creation of thought: "the earliest stage of language would have been just that: a language of thought, used internally" (Chomsky 2007: 13). This ILoT, common in essence to the species, would have subsequently been connected to the SM system for externalization and, therefore, for communication. According to this vision, externalization would be ancillary and secondary, that is, a process exposed to fluctuations in the environment and, therefore, susceptible to change and diversification:

"Parameterization and diversity, then, would be mostly –possibly entirely– restricted to externalization. That is pretty much what we seem to find: a computational system efficiently generating expressions interpretable at the semantic/pragmatic interface, with diversity resulting from complex and highly varied modes of externalization, which, furthermore, are readily susceptible to historical change." (Berwick and Chomsky 2011: 37-38)

What this scenario implies, then, is that the FL must also include a component derived from the environment (that is, internalized) whose mission would be to systematically connect the derivations generated by the ILoT (resulting from the interaction between the conceptual and the computational systems) with sensorimotor systems. The crucial idea is that this component, which is the result of the internalization of the environment, is what really differentiates languages (understood as I-languages) from each other and constitutes the genuinely cultural component of every human language. For expository convenience I will call this component *lexical interface* (see Figure 1). The use of the

expression *lexical interface* is based on the traditional idea that the lexicon of a language is the component where sounds and meanings are systematically matched. However, the reading in which the lexicon is the set of words or morphemes that the syntax combines to create sentences must be avoided. In the use of the term that interests me now, the lexical interface should be interpreted as an area of long-term memory that provides a stable connection between, on the one hand, the syntactic derivations produced by the computational system in interaction with the conceptual-intentional system and, on the other hand, the sensorimotor systems that process and produce the material linguistic signals (sounds or visual signs) that human beings perceive and produce when they use language for communication. More specifically, the lexical interface should be understood as a set of systematic pairings between certain syntactic structures and phonological words (see Mendívil-Giró In press for a more detailed justification of this way of understanding the lexicon and full references).

Like any organic system, the FL of each person (their I-language) is conditioned by two types of factors: internal (derived from biology and other natural factors) and external (derived from environmental information). According to the scheme of Figure 1, any I-language, insofar as it is a person's FL, is formed by the four components. Three of these (the conceptual-intentional system, the computational system and the sensorimotor system) are essentially universal because they are organism-internal and are naturally conditioned, while the fourth, the lexical interface (highlighted in a darker tone), is culturally variable since it is the result of internalization from environmental stimuli.

According to this model, the acquisition of language does not imply the internalization of the entire system of knowledge (the I-language), but only of one of its components (the lexical interface). The development of language in the individual is then equivalent to the development in the mind/brain of the lexical interface, which can be glossed as the process of learning to externalize the ILoT in the same way as the members of our community do.

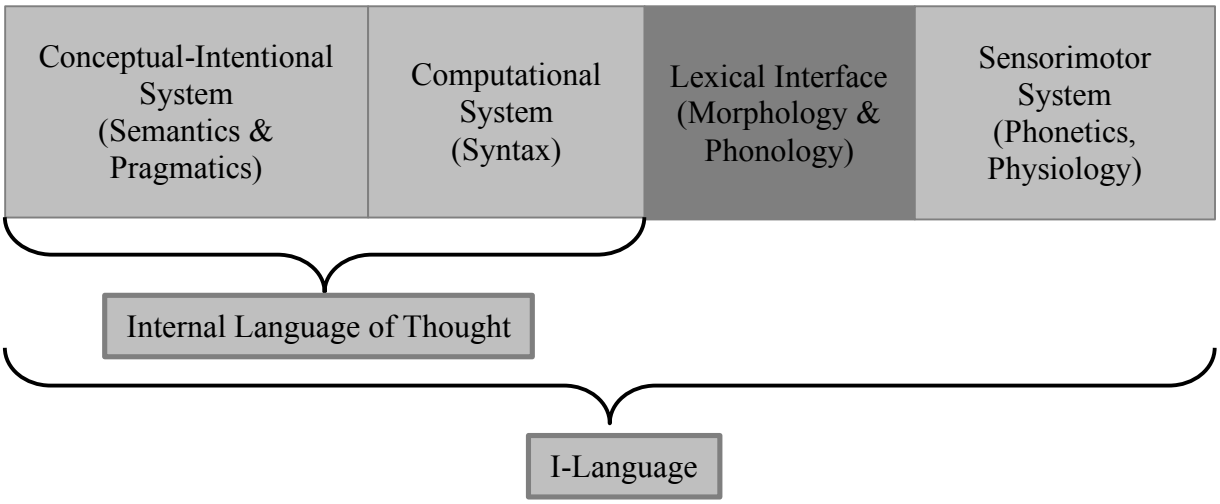

Figure 1 The anatomy of an I-language

Note that in each component I have indicated the scope of traditional grammar with which it would be centrally associated. In this way, the CI system is related to semantic and pragmatic interpretation. I do not mean to imply that there is no linguistic and

cultural variation in this respect, but that there is an underlying uniformity. Note that we can, for example, ask someone in what language they speak, in what language they think, or in what language they dream, but it is strange to ask in what language they mean (see Jackendoff 2012 for a development of this argument). The very fact that we can consider whether two linguistic expressions (of the same or of different languages) have the same meaning shows that there is a layer of meaning deeper than the linguistic forms that externalize it.

I have also assumed that syntax is uniform, but in this case I refer to the basic computational mechanisms and the formal principles governing the syntactic derivation (merge, binariety, endocentricity, etc.) and not to the fact that the apparent syntax of languages is diverse (as with basic word order or argument marking patterns). In fact, the hypothesis of much of modern formal linguistics is precisely that these differences in the 'visible' syntax are the result of differences in the repertoire of linguistic formants that each language uses to externalize (in fact, to materialize) the syntactic derivations produced by the internal computational system. In traditional terms, the underlying hypothesis is that any difference between the structure of languages is of a morphological and phonological nature (see, for example, Richards 2016).

6 Explaining discrepancy

The model of the minimum internal structure of an I-language presented in Figure 1 allows us a better understanding of the disparity of opinions on the balance of forces between culture and nature in language design that we observe in current linguistics. The central (simple) idea is that not all linguists use the word *language* in the same sense. More specifically, I want to suggest that a crucial difference in this use comes from the fact that the externalist tradition inadequately identifies an I-language with one of its parts, the lexical interface. For this reason, generativist authors tend to reject many of the statements about languages that externalist authors make, such as the claim that languages are external to the mind, that they can vary profoundly, that they are learned using general mechanisms of statistical learning, or that they owe their structure to the historical processes of change and not to a faculty of language. But note that these same claims would be more acceptable if they were interpreted as referring to the lexical interface rather than to the whole I-language.

Thus I conclude that the notorious divergences regarding the nature of language and languages currently seen in theoretical linguistics are largely a consequence of an incomplete vision of what a natural human language really is, a vision based on the misidentification of languages with their learned and historically modified components. From the generativist point of view it is reasonable to think that to a large extent the darker cell in Figure 1 is something external and cultural; however, the assertion that it is a language is not admissible. Of course, the opposite might be argued (by Tomasello, for example), that this component is the object of study if one wants to study language and not cognition in general. But this, in my opinion, is a crucial error. It is an error which derives from an externalist conception of language and an empiricist conception of the mind. In fact, any I-language is also part of *general cognition* (if this expression makes sense). To affirm otherwise would be to pretend (to return to our former analogy) that the study of the shape of the elephant's skin is independent of the study of

the elephant's body. Yet the skin is part of the elephant's body, just as a language is the whole thing (the whole body) not just its most superficial part (the skin).

Contrary to what Tomasello suggests in the above quotation, a language is not a cultural system represented in the brain. If anything, this is the definition of the lexical interface, that is, of the cultural component of languages that connects their most internal components with the sensorimotor system of externalization. A language, rather, is a system of knowledge that includes a variable cultural component, but also universal linguistic aspects of human cognition.

7 Conclusion: What does *language* mean?

The conclusion I propose here is two-fold. On the one hand, the 'division of labor' between nature and culture in the design of human language could be explained as a consequence of the different nature of the various components of any human language. Internal systems, such as syntax, would be essentially conditioned by biology and the laws of nature (factors 1 and 3 of Chomsky 2005), while the externalization system (the internalized lexical interface) would be essentially conditioned by environmental and cultural factors (factor 2 of Chomsky 2005). In addition, although this is not the place to develop such a central issue, there are hints that the proposal to identify nature and culture in language with, respectively, syntax and lexicon in languages has empirical support. The evidence here can be grouped into two basic categories, related to the asymmetry in acquisition and to the asymmetry in the degree of variation of each component:

(i) The argument of poverty of the stimulus is especially strong when applied to the development of syntactic knowledge, as opposed to the acquisition of the lexicon. In the same way, connectionist and statistical learning models work better with certain aspects of phonology and morphology than with syntax (see Lidz and Gagliardi 2015, Yang 2016).

(ii) It is easier to make descriptive generalizations about syntax than about morphology, in the sense that all languages have syntax, but not all languages have (the same types of) morphology. Syntactic structures (phrases, sentences) and syntactic relationships (subject of, object of, argument, modifier, etc.) are much more general in languages than morphological categories (agreement markers, pronominal clitics, etc.). For example, it is conceivable to question that in a language there are nouns or verbs, but it is more difficult to argue that there are no noun phrases or verbal phrases.

On the other hand, this proposal could help to clarify the long and harmful misunderstandings about the nature of language and languages that have played a role in recent decades of linguistic theory, offering a fragmented and incoherent view of the science of language and negatively contributing to the development and the adequate social perception of our discipline.

I have shown that it is possible to explain many theoretical disagreements in linguistics as the result of a misunderstanding about the meaning of the word *language* (in both the mass and count readings of the noun). For a generativist linguist, a language is a

biologically determined system of knowledge with a variable cultural component (the lexical interface that serves for the externalization of language). That is why generativist linguists cannot accept the assertion that languages are learned by imitation, that they vary without limit, that they adapt to the communicative needs of speakers, etc. But all these remarks would be much more acceptable if we were to talk about the lexical interface. From the externalist perspective, a language is a cultural object at the service of communication. But note that this definition coincides with *a part* of what a language is from the generativist point of view. Indeed, the lexical interface that is part of each language is a cultural object, which is learned inductively, varies profoundly, etc. Thus, if generativists, when reading a functionalist paper, were to interpreted the word *language* as 'lexical interface', they would find it far more interesting, and if externalist authors, when reading a generativist paper, were to understand the word *language* as 'something that integrates a cognitive system deeper than the surface forms', they would also find the approach more acceptable. For this reason, I would say that internist and externalist research programmes are more complementary than they are incompatible, despite the fact that most practitioners appear to ignore this.

References

Berwick, Robert C. and N. Chomsky (2011). "The Biolinguistic Program: The Current State of its Evolution and Development". In A.M. Di Sciullo y C. Boeckx (eds.) (2011). *The Biolinguistic Enterprise.* Oxford: Oxford University Press, 19-41.

Berwick, Robert C. and Noam Chomsky (2016). *Why Only Us*. Cambridge (MA): The MIT Press.

Chomsky, Noam (2000). *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.

Chomsky, Noam (2005). "Three Factors in Language Design". *Linguistic Inquiry*, 36, 1-22.

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

Christiansen, Morten H. and Nick Chater (2015). "The language faculty that wasn't: a usage-based account of natural language recursion". *Frontiers in Psychology*, 6, 1182.

Damasio, Antonio (1994). *Descartes' Error*. New York: G.P. Putnam Sons.

Deacon, Terrence W. (1997). *The Symbolic Species: the Co-Evolution of Language and the Brain*, New York: W.W. Norton.

Evans, Nicholas and Stephen C. Levinson (2009): "The Myth of Language Universals: Language diversity and its importance for cognitive science". *Behavioral and Brain Sciences*, 32, 429-448.

Fitch, W. Tecumseh (2009). "Prolegomena to a Future Science of Biolinguistics". *Biolinguistics*, 3, 283-320.

Hauser, Mark D., Noam Chomsky and W. Tecumseh Fitch (2002). "The Faculty of Language: What Is It, Who Has It, and How It Evolved?". *Science*, 298, 1569-1579.

Jackendoff, Ray (2012). *A User's Guide to Thought and Meaning*. Oxford: Oxford University Press.

Lidz, Jeffrey and Annie Gagliardi (2015). "How Nature Meets Nurture: Universal Grammar and Statistical Learning". *Annual Review of Linguistics*, 1, 333-53.

Mendívil-Giró, José-Luis (2018). "Is Universal Grammar ready for retirement? A short review of a longstanding misinterpretation". *Journal of Linguistics*, 54/4: 859-888.

Mendívil-Giró, José-Luis (in press). "If everything is syntax, why are words so important?", *Linguistics* (Accepted).

Richards, Norvin (2016). *Contiguity Theory*. Cambridge MA: The MIT Press.

Saussure, Ferdinand de (1916). *Cours de linguistique générale*. Paris: Payot, 1975.

Tomasello, Michael (2009). "Universal grammar is dead". *Behavioral and Brain Sciences*, 32, 470-471.

Yang, Charles (2016): *The Prize of Linguistic Productivity. How children learn to break the rules of language*. Cambridge MA: The MIT Press.