biolinguistics

EDITORIAL

Biolinguistics 14.SI: 51–58, 2020
www.biolinguistics.eu

Why Don't Languages Grammaticalize [±poisonous]?

Evelina Leivada[1,*] & Lluís Barceló-Coblijn[2]

[1] Department of German and English Studies, Universitat Rovira i Virgili, Spain
[2] Laboratori d'Investigació en Complexitat i de Lingüística Experimental (LICLE), Universitat de les Illes Balears, Spain

[*] Corresponding author: evelina.leivada@urv.cat

EL: https://orcid.org/0000-0003-3181-1917
LBC: https://orcid.org/0000-0002-8765-6314

1. Introduction

Biolinguistics is the study of language from a biological point of view. As the *Special Issue—Biolinguistic Research in the 21st Century of Biolinguistics* shows, this area of study can be approached from different perspectives: aspects of language evolution (Wacewicz et al. 2020; Mendívil-Giró 2020), ethology and cross-species comparisons (Schalz & Dickins 2020), and neural network analysis (Collins 2020) are some of them. Instead of presenting a brief overview of these works, we will use this editorial article to illustrate a more valuable and often overlooked point, namely, that another perspective to the study of language from a biological point of view is the one that goes through theoretical linguistics. Any study in theoretical linguistics that strives for explanatory adequacy in terms of the innate primitives it assumes falls within the domain of biolinguistics. Although some theoretical linguists explicitly distance themselves from biolinguistics, the question of 'What is innate in language behavior?' as well as any discussion that evokes innate characteristics is contributing towards a biological theory of language. This is not an original claim; it is how Eric Lenneberg chose to close the last chapter of his ground-breaking book *Biological Foundations of Language* (Lenneberg 1967).

Some of the questions that Lenneberg addressed regarding innate features and cross-linguistic variation are still looking for answers today. One example comes from the first two live debates in a series of video conferences called "Linguistics Flash-Mobs. Epic Battles in History", which started in May 2020. In each meeting, two scholars are invited to discuss longstanding theoretical issues

Published: 25 March 2021 ISSN 1450–3417

in linguistics. In the first debate, two theoretical linguists, Giuseppe Longobardi and Ian Roberts, addressed the topic of parameters and parametric variation. Roberts posed a critical question that summarizes perfectly (i) our current understanding of the set of features languages grammaticalize and (ii) why these features are selected for grammaticalization instead of others. In his words, "which (sub)set of formal features is underspecified is a question that is very difficult to answer in part because we don't really know what a set of formal features is" (Roberts 2020). Addressing this matter, Luigi Rizzi argued that "clearly U[niversal] G[rammar] must say something about the set of formal features. A rather elementary empirical observation is that of all the properties that are cognitively salient or perceptually salient only a very small subset is actually grammatical[ized]" (Rizzi 2020).

To sum up, the argument is that languages grammaticalize only a subset of features and the link that Rizzi established is between UG, as the locus of the features that have been selected for grammaticalization, and parameterization, as the process that accounts for the cross-linguistically variable values of these features. Rizzi also gave a useful example to illustrate the difference between grammaticalized and non-grammaticalized features. In his words, "something like physical shape can be grammaticalized, as many systems of nominal categories may refer to shape, but no language, to the best of our knowledge, may refer to color. There is no agreement for yellow objects" (Rizzi 2020). As Ian Roberts observed, one then faces the following question when addressing this topic: Why doesn't UG encode a color agreement feature?

Before addressing this question—a question which according to Roberts has not received an answer yet—, it is useful to bring into the picture the second live debate of the "Linguistics Flash-Mobs. Epic Battles in History" series. This debate featured again two theoretical linguists, Peter Svenonius and Guglielmo Cinque, who discussed an issue that is conceptually related to the topic of the first debate: functional hierarchies. One of the very first points that Cinque raised related to why we find that only some functional elements are encoded grammatically in all languages. He further argued that our cognitive system is quite rich and involves many notions that are not coded grammatically (Cinque 2020). His conclusion, formed on the basis of these observations, echoed Rizzi's conclusion in the first debate: UG must encode the toolkit of features that languages grammaticalize.

Indeed, studies in crosslinguistic variation attest to the fact that the vast majority of languages grammatically represent certain notions (for example, Tense or Aspect), but not others (for example, whether something is poisonous). The research questions that emerge in this context are the following two:

RQ 1: Does the observation that certain notions are grammaticalized, while others are not, offer an argument in favor of a rich UG that encodes the relevant features?

RQ 2: Why does UG (or language in general, if one does not link the presence of this feature toolkit with UG) grammatically encode specific features, but not others that are equally salient from a cognitive point of view (e.g., color)?

2. A Rich UG?

Undoubtedly, there are restrictions in the pool of features that languages grammaticalize. However, the conclusion that several scholars, including Rizzi and Cinque, reach on the basis of this premise, namely, that UG dictates the set of grammaticalized features, does not follow as a logical conclusion. The observation that certain notions are consistently grammaticalized across languages only entails that languages tend to grammaticalize certain markers, but not others. The reasons for this may have nothing to do with UG. How *Homo sapiens* evolved towards a phenotype that is able to process and encode specific grammatical markers is an empirical question. A rich and highly detailed UG entails a rigid system, and this is less desirable for a developmental process like language development, that follows a nonlinear trajectory and shows abrupt transitions (Ruhland & van Geert 1998; Ninio 2006; Corominas-Murtra et al. 2009; Bassano et al. 2011; Barceló-Coblijn et al. 2012). A system that is flexible enough can reach the mature state more easily than a very detailed, rigid system. From this perspective, it would be surprising if brains show ontogenetic plasticity, but develop rigid cognitive systems. An innate large set of highly detailed features or patterns would also be problematic for language acquisition. When humans have to learn artificial, highly detailed and specific languages, the chosen strategy by the learners tends to be an adaptation and reduction of the features towards under-specification of meanings (Kirby et al. 2008).

Another challenge for the proponents of the position that UG encodes the set of features that languages grammaticalize boils down to the fact that, unlike what RQ 1 may imply, languages do vary as to which features they grammaticalize. For example, if the Tense category is understood as a verb inflection that specifies TIME, spoken languages largely conform with this pattern, but verbs in sign languages often do not inflect for tense. Rather, temporal information is conveyed by time adverbials and/or is inferred from the context (Aarons et al. 1995; Pfau & Steinbach 2006). Clearly, languages vary in terms of the grammatical markers they employ, and they also vary in relation to the strategies they use for expressing the same notion (i.e. grammaticalization or lexicalization). If TIME can be variably expressed across languages through an adverbial, an inflectional marker, or background information, can this notion be removed from the hypothetical UG toolkit that specifies the subset of formal features that are grammaticalized? Put another way, if this notion is expressed in some languages in the *absence* of any dedicated grammatical marker, this feature is potentially grammaticalizable, but not necessarily grammaticalized across all the languages in which it is expressed. It thus seems that RQ 1 must be amended as follows, in order to make reference to notions that are grammaticalizable, but not necessarily grammaticalized: Does the observation that certain notions are grammaticalizable, while others are not, offer an argument in favor of a rich UG that encodes the relevant features? This reformulation of the question puts the matter in its right dimension but brings forward a major problem. It has *not* been shown that features like COLOR or BRIGHTNESS cannot be grammaticalized. What has been observed is that formal grammars do not develop grammatical markers for such notions (Cinque 2013; Adger 2018; Sigurðsson 2020). This suffices to give a

negative answer to RQ 1, because (i) the possibility that COLOR or POISON can be grammatically expressed and the relevant markers successfully acquired has not been discarded and (ii) even if (i) is established, it is not necessary that it is UG that precludes the grammaticalization of certain markers. As the next section will show, several other explanations can be given as to why some features are not grammaticalized, such as adaptation pressures towards input simplification as well as notion-specific semantic restrictions that may favor lexicalization over grammaticalization.

3. Why Not [±poisonous]?

The question that emerges next is why languages don't grammaticalize [±poisonous] or [±bright] if in principle these notions are grammaticalizable. RQ 2 phrases this in a more general way by asking why languages grammatically encode specific features, but not others that may be equally salient from a cognitive point of view. Recall that versions of this question have been voiced as lending support to the idea that if [±poisonous] and other notions are not grammaticalized, it must be the case that UG encodes a toolkit of features from which languages variably choose what to grammaticalize (i.e., not all languages have a grammatical marker for evidentiality, although potentially they could in the sense that this is a grammaticalizable, learnable marker).

One answer to RQ 2 is that the semantics of [±poisonous] is not a good candidate for an atomic semantic primitive in the first place. It is an endpoint, not a building block.[1] A second answer is that any evolved feature must serve a purpose (like any selected trait for an organism) such as facilitating successful reference in some respect. Having a grammatical marker for poison would not serve any such purpose, so from a biolinguistic point of view it would be superfluous and unlikely to ever develop. Of course, describing whether something is poisonous or not is extremely useful in any community, but languages do not need to express this through a grammatical marker, whereas they often need to grammaticalize past and present tense for obvious reasons of time reference any time an event is described. A third answer has to do with learnability considerations. A grammatical marker that encodes [±poisonous] would probably fail to be re-transmitted by learners even in cases of iterated learning in artificial language learning tasks, precisely because it would be useless in most contexts.

A fourth answer relates the existence of formal features with the evolutionary origins of our species. A number of scholars have argued that only concepts that predate the emergence of *Homo sapiens* are candidates for formal features (Emonds 2011; Golston 2018; Panagiotidis 2021). Emonds (2011) presents this hypothesis in the following way.

(1) *Emonds' anti-correlation hypothesis*
Almost all concepts F of human syntax are among those plausibly associated with the cognition of highly developed non-human primates, precisely those who *lack* syntax.

[1] This observation is due to Gillian Ramchand, offered during a Twitter exchange.

This answer seems to suffer from two problems. First, as Emonds (2011) argues, many grammatical features of human language are indeed present in the communication systems of other primates (e.g., ANIMATE). Although this is largely true for animacy, the same cannot be argued for all features that human language can grammaticalize. One example is evidentiality. Defined as the linguistic designation of source of information for a belief, evidentiality has been described as the linguistic coding of epistemology (Chafe & Nichols 1986). Languages differ in the way they grammatically encode their evidential markers across many dimensions. The first and most fundamental difference is that, although all languages have a way to indicate source of evidence for an utterance, only a quarter of them have a grammatical system of specific markers for encoding it (Aikhenvald 2004). These markers can denote source of evidence (i.e., hearsay, inferential, dubitative), modality of source (i.e., auditory, visual), and degree of certainty. According to Emonds' anti-correlation hypothesis (1), the prediction is that such notions *are expressed* in the repertoire of other species. However, the use of the "hearsay" evidential, for example, would entail the transition from a direct predator alarm call (i.e., 'predator X is close') to a "reported speech" predator alarm call (i.e., 'I heard/they say/it is reported that predator X is close'), something that contradicts our current understanding of animal cognition, which has been argued to not provide evidence of hierarchically organized structure learning (Petkov & ten Cate 2020).

The second problem with Emonds' anti-correlation hypothesis is that it predicts that notions that are not grammaticalized in human language must be absent from the cognition of other primates. For example, he argues that is highly doubtful that other species classify the activities of others as respectful, generous, or selfish. Again, results from recent animal studies do not seem to borne out this prediction. For instance, some works suggest that domestic dogs could be able to recognize generous and selfish people and later express this as a systematic preference in choice tests (Carballo et al. 2015). Other works report that chimpanzees could also be able to infer the reputation of humans as selfish or generous agents, perhaps also expressing systematic preferences for generous donors over selfish ones (Subiaul et al. 2008).

Going back to RQ 2, [±poisonous] is not a good candidate for a grammatical marker in human language for at least three reasons that are not related to the anti-correlation hypothesis in (1). At the same time, POISONOUS is a good candidate for a lexical item in the repertoire of both other species (e.g., chimpanzees produce different type of grunts when encountering different types of food; Slocombe & Zuberbühler 2006) and humans.

4. Outlook

This work focused on two questions that are often addressed within the realm of theoretical linguistics: 1. Why languages consistently (do not) grammaticalize a subset of cognitive notions, and 2. Whether this observation justifies the conclusion that Universal Grammar must encode a toolkit of grammaticalizable features. The first question received several explanations. Learning considerations, adaptation constraints, and semantic restrictions that may make a notion

a better candidate for lexicalization instead of grammaticalization are some of the reasons that explain why TENSE is a good candidate for the development of a grammatical marker, but COLOR or POISON are not.

The second question was answered negatively: Contrary to popular claims within cartography, it has not been established that it is UG that precludes the grammaticalization of certain markers. Providing the bigger picture, the take-home message is that these are empirical questions that show how theoretical linguistics and biolinguistics are conceptually connected, such that addressing long-standing debates in theoretical linguistics inevitably goes through discussing certain key topics in biolinguistics such as evolution, adaptation, and development.

Author contributions

EL and LBC conceptualized and wrote the paper. Both authors jointly revised the manuscript and approved it for publication.

Acknowledgements

We thank all the authors who participated in the *Special Issue—Biolinguistic Research in the 21st Century of Biolinguistics* and all the reviewers who provided their invaluable service. We also thank Associate Editor Patrick Trettenbrein, who provided help in technical matters, and Editor-in-Chief Kleanthes Grohmann, who assisted in the smooth completion of the special issue. For this editorial, EL acknowledges financial support from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement n° 746652 and from the Spanish Ministry of Science, Innovation and Universities under the Ramón y Cajal grant agreement n° RYC2018-025456-I. The funders had no role in the writing of the study and in the decision to submit the article for publication.

References

Aarons, Debra, Benjamin J. Bahan, Judy Kegl, & Carol Neidle. 1995. Lexical tense markers in American Sign Language. In Karen Emmorey & Judy Reilly (eds.), *Language, Gesture, and Space*, 225–253. Hillsdale: Lawrence Erlbaum.

Aikhenvald, Alexandra. 2004. *Evidentiality*. New York: Oxford University Press.

Barceló-Coblijn, Lluís, Bernat Corominas-Murtra, & Antoni Gomila. 2012. Syntactic trees and small-world networks: Syntactic development as a dynamical process. *Adaptive Behavior* 20, 427–442.

Bassano, Dominique, Isabelle Maillochon, Katharina Korecky-Kröll, Marijn van Dijk, Sabine Laaha, Wolfgang U. Dressler, & Paul van Geert. 2011. A comparative and dynamic approach to the development of determiner use in three children acquiring different languages. *First Language* 31, 253–279.

Carballo, Fabricio, Esteban Freidin, Natalia Putrino, Carolina Shimabukuro, Emma Casanave, & Mariana Bentosela. 2015. Dog's discrimination of human selfish and generous attitudes: the role of individual recognition, experience, and experimenters' gender. *PLoS ONE* 10(2): e0116314.

Chafe, Wallace L. & Johanna Nichols (eds.). 1986. *Evidentiality: The Linguistic Coding of Epistemology*. Norwood, NJ: Ablex.

Cinque, Guglielmo. 2020. Debate 2: Functional hierarchies (Guglielmo Cinque and Peter Svenonius). *Linguistics Flash-Mobs. Epic Battles in History*, Università degli Studi di Padova, https://www.youtube.com/watch?v=r_6DRMgxI44&t=4453s.

Cinque, Guglielmo. 2013. Cognition, Universal Grammar, and typological generalizations. *Lingua* 130, 50–65.

Collins, Joe. 2020. The phonological latching network. *Biolinguistics* 14.SI, 102–129.

Corominas-Murtra, Bernat, Sergi Valverde, & Ricard V. Solé. 2009. The ontogeny of scale-free syntax networks: Phase transitions in early language acquisition. *Advances in Complex Systems* 12, 371–392.

Emonds, Joseph E. 2011. From primate to human in two easy steps. In Nicholas Saul & Simon James (eds.), *The Evolution of Literature: Legacies of Darwin in European Cultures*, 43–69. Amsterdam: Rodopi.

Golston, Chris. 2018. Phi-features in animal cognition. *Biolinguistics* 12, 55–98.

Kirby, Simon, Hannah Cornish, & Kenny Smith. 2008. Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences* 105(31), 10681–10686.

Lenneberg, Eric H. 1967. *Biological Foundations of Language*. New York: John Wiley & Sons.

Mendívil-Giró, 2020. In defence of FLB/FLN: A reply to Wacewicz et al. (2020). *Biolinguistics* 14.SI, 145–153.

Ninio, Anat. 2006. *Language and the Learning curve: A New Theory of Syntactic Development*. Oxford: Oxford University Press.

Panagiotidis, Phoevos. 2021. Towards a (minimalist) theory of features: Preliminary notes. *Lingbuzz*, https://ling.auf.net/lingbuzz/005615.

Petkov, Christopher I. & Carel ten Cate. 2020. Structured sequence learning: Animal abilities, cognitive operations, and language evolution. *Topics in Cognitive Science* 12(3), 828–842.

Pfau, Roland & Markus Steinbach. 2006. Modality-independent and modality-specific aspects of grammaticalization in sign languages. *Linguistics in Potsdam* 24, 5–98.

Rizzi, Luigi. 2020. Debate 1: Three questions on parameters (Ian Roberts and Giuseppe Longobardi). *Linguistics Flash-Mobs. Epic Battles in History*, Università degli Studi di Padova, https://www.youtube.com/watch?v=uGCHNcI4kGg.

Roberts, Ian. 2020. Debate 1: Three questions on parameters (Ian Roberts and Giuseppe Longobardi). *Linguistics Flash-Mobs. Epic Battles in History*, Università degli Studi di Padova, https://www.youtube.com/watch?v=uGCHNcI4kGg.

Ruhland, Rick & Paul van Geert. 1998. Jumping into syntax: Transitions in the

development of closed class words. *British Journal of Developmental Psychology* 16, 65–95.

Schalz, Sabrina & Thomas E. Dickins. 2020. Humans discriminate individual zebra finches by their song. *Biolinguistics* 14.SI, 130–144.

Sigurðsson, Halldór Ármann. 2020. Universality and variation in language: The fundamental issues. *Evolutionary Linguistic Theory* 2(1), 5–29.

Slocombe, Katie E. & Klaus Zuberbühler. 2006. Food-associated calls in chimpanzees: Responses to food types or food preferences? *Animal Behaviour* 72(5), 989–999.

Subiaul, Francys, Jennifer Vonk, Sanae Okamoto-Barth, & Jochen Barth. 2008. Do chimpanzees learn reputation by observation? Evidence from direct and indirect experience with generous and selfish strangers. *Animal Cognition* 11, 611–623.

Wacewicz, Slawomir, Przemyslaw Zywiczynski, Stefan Hartmann, Michael Pleyer, & Antonio Benítez-Burraco. 2020. *Language* in language evolution research: In defense of a pluralistic view. *Biolinguistics* 14.SI, 59–101.