Nordic Journal of Linguistics (2020), **43**, 127–132
doi:10.1017/S0332586519000258

CAMBRIDGE
UNIVERSITY PRESS

REVIEW

Morten H. Christiansen & Nick Chater, *Creating Language: Integrating Evolution, Acquisition, and Processing*. Cambridge, MA: MIT Press, 2016. Pp. xiv + 330.
doi:10.1017/S0332586519000258

Reviewed by Giosuè Baggio

Language Acquisition and Language Processing Lab, Department of Language and Literature, Norwegian University of Science and Technology, 7491 Trondheim, Norway. giosue.baggio@ntnu.no

An extraterrestrial looking down on us may or may not be struck by the (apparent) uniformity of human languages (Chomsky 2000), but they would likely notice that, by some measures, there is greater diversity among THEORIES of language than there is among languages. It may not take long for them to discover the cause of this. We are grappling with a 'hard problem' in language science: on the one hand, linguistic structures afford multiple descriptions, formalizations, etc.; on the other, whatever formalism one adopts, it is difficult to see how such structures could have emerged gradually as the result of processes of evolution, historical change, acquisition, and neural computation. Over almost three decades of joint work, Morten Christiansen and Nick Chater (C&C) have developed an original and productive approach to this 'hard problem', synthesized in a recent book, *Creating Language*, published by the MIT Press. This book occupies a unique position in the vast and diverse landscape of theories of human language. It is not only one of the very few proposals that aim to explicitly reconcile accounts of linguistic structure and process: it is also serving as an 'aggregator' of theoretical views and empirical results that have originated in different communities (cognitive linguistics, psycholinguistics, AI, etc.) in response to mainstream generative linguistics and that have remained scattered or in search of a unifying framework for many years. *Creating Language* is mandatory reading for anyone interested in understanding ongoing tectonic shifts in language science.

The book's core idea is presented in Chapter 1. Human language is 'created' across multiple time scales: biological evolution fixes the 'hard priors' (i.e. the brain) that then shape, during cultural transmission, acquisition, and processing, the linguistic structures that emerge, and eventually stabilize, in communities of language users. Chapter 1 also states the philosophy and methodology of C&C's book: questions of processing, acquisition, and evolution should be addressed jointly (e.g. acquiring a language is learning to process that language), and theories of linguistic structures should take into account various processing constraints (more on this later). Thus, language science is primarily driven by experimental

and computational research, and much less so by formal accounts of language structure – those are expected to follow from the former kind of enterprise. The rationalism implicit in 'traditional' approaches in theoretical linguistics, such as generative syntax, formal semantics, and others, is briskly replaced with a new empiricism, more in line with advances in cognitive linguistics, psychology, neuroscience, and connectionist modeling.

Chapter 2 presents a piercing critique of the idea of universal grammar (UG) and of adaptationist and non-adaptationist arguments for it. The 'moving target' scenario developed by C&C against adaptationist views of UG is now well known: languages typically change too rapidly to lead to adaptations that can be fixed in the genome; the slower pace of biological adaptation may even lead to decreasing fitness, if the target language has changed again by the time those adaptations have caught on. C&C offer two new arguments to that effect. The non-adaptationist picture is ruled out by a simple argument, designed to show that the probability that UG arises just by chance would require implausibly large populations of speakers. C&C conclude that the brain did not 'evolve' UG to accommodate language. Rather, it is languages that change over generations of learners and users to fit the given brain structures and processing resources. This small 'Copernican revolution' proposed by C&C and others (Dehaene & Cohen 2007) is becoming highly influential in cognitive science. It still allows linguists to ask questions such as whether and how the brain would constrain the range of variation of languages, but it also enables them to abandon the controversial concept of UG (or any prior structure which is language specific, as opposed to domain general) in attempts to answer such questions. However, in the model proposed by C&C, language acquires a problematic ontological status: it is no longer only an internal state of human minds and brains, but 'something' that is under pressure to evolve and adjust in order to survive, and can survive only if it is easily learnable and processable (pp. 42–43). Philosophically, the question arises whether it is possible to marry classical internalism about linguistic structures and processes with C&C's thesis that language is 'shaped by the brain'. Perhaps one can assume that what is shaped by the brain is not language as such (which is rather a theoretical construct; see Stokhof & van Lambalgen 2011) but linguistic behavior: over successive generations, linguistic behavior (primarily that of adult speakers) gradually adapts to the biases or constraints imposed by the brain (primarily that of learners). Nevertheless, C&C's own proposal is that language is itself a complex 'organism' (p. 43; Box 2.2) which adapts to human biological prior structures, yet one need not accept this idea to appreciate the book's many solid contributions.

In Chapter 3, C&C attempt to reconnect language evolution (historical change) and acquisition. The structure of language gradually adapts to the brain during cultural transmission: languages will evolve (change) to fit learning biases and constraints. The problem of learning a language is not one that can be solved via what C&C call 'N-induction' (i.e. discovering stable properties of the natural world), but requires 'C-induction' (coordinating with other language users). In C-induction, the learner acquires a very specific skill (do what other speakers do), while in N-induction one has to learn to do the objectively right thing (to respond appropriately to events in the environment). This 'dramatically simplifies' the learning problem: because we all share the same learning biases and

https://doi.org/10.1017/S0332586519000258 Published online by Cambridge University Press

constraints, it is likely that learners will find shared solutions quickly (e.g. 'focal points' in game theory). Although many details of this proposal are not fully clarified (e.g. how much syntax can really be acquired in this way? See Table 3.2 on p. 72), and although formal models of coordination as applied to language are not mentioned and discussed (e.g. signaling games; Jäger 2008), C&C's proposal is generally coherent and plausible.

Chapter 4 presents another core idea of the book. The 'now-or-never' bottleneck is the main constraint to which language must adapt: linguistic information is lost for good if it is not processed rapidly. This is argued to explain facts such as language's multi-level organization and the increasingly large temporal frames for processing afforded by each level, the prevalence of local dependencies, the incrementality of processing, predictive mechanisms, aspects of language acquisition, 'and even the structure of language itself' (p. 94). The key to understanding these phenomena is how the brain responds to the bottleneck – by adopting a chunk-and-pass strategy, where the input is broken down into chunks that can be easily encoded and passed on to the next level of representation. This encoding involves lossy compression, so that lower level representations (e.g. phonological or lexical representations) may not be recoverable from higher level representations (e.g. sentence meaning). The psycholinguistic evidence for these processes is overwhelming, but one wonders if C&C's case is overstated. First, it is not clear that the now-or-never bottleneck is as 'tight' as C&C argue: speech production systems in the brain are tuned to language comprehension systems and vice versa, such that neither system may be expected to lag behind the other in terms of information processing speed, resolution, etc. In addition, the brain has various resources, beyond chunk-and-pass, for dealing with the volatility of the input: inference and the ability to control interactions with the source of information are two examples (for discussion, see Baggio & Vicario 2016 and other commentaries to Christiansen & Chater 2016). C&C paint a rather bleak picture of language processing: learners and hearers are exposed to a 'deluge' of material, human working memory has 'severe limitations', etc. (pp. 132–133). But emphasis on constraints and limitations should be tempered with emphasis on the CAPACITIES that human brains display (e.g. for syntactic and semantic composition, reference, inference, etc.), which likely 'shape' language as much as the former do.

Part II of the book – Chapters 5–8 – considers the implications of these ideas for theories of linguistic structure and language processing. Chapter 5 presents a view of language processing and acquisition as multiple-cue integration. It is individual lexical items, and not the entire vocabulary, that are optimized for processing and acquisition, and 'optimization' is different at different learning stages. For example, C&C argue that systematic (non-arbitrary) relations between words and meanings facilitate acquisition of early-learned words, whereas arbitrary relations facilitate acquisition of larger vocabularies. Moreover, information is available in the speech input beyond well-understood cues such as length or frequency: e.g. phonological cues help learners to distinguish verbs from nouns. Chapter 6 examines the role of individual experience in fine-tuning processing skills, including chunk-and-pass. In Chapter 7, C&C present an account of recursion as a usage-based skill. Because our ability to process recursive structures is

https://doi.org/10.1017/S0332586519000258 Published online by Cambridge University Press

bounded, a theory of sentence processing does not require formal recursive processing mechanisms of the kind envisaged by most generative theories. Such mechanisms would overgenerate, i.e. would license sentences that are never produced and understood. Instead, C&C suggest, bounded recursive abilities rest on domain-general sequence learning and processing skills. Connectionist models – simple recurrent networks – capture human behavior and capacity limitations in processing recursive structures, without positing externally specified 'performance' limitations.

In Chapter 8, C&C argue for a reintegration of language science. The book provides a compelling case for a unification of empirical perspectives on language evolution, acquisition, and processing, and C&C largely succeed in showing how this could be done and what is to be gained from an integrated science of linguistic processes. It is however much less clear that theories of linguistic STRUCTURE can be incorporated into the resulting picture. For example, a key issue concerns the status of linguistic structures that humans cannot, on average, process, e.g. deep recursive structures. The traditional answer is that those sentences are part of the language, but beyond human processing limitations – hence various methodological distinctions, such as competence *vs* performance. For C&C, any linguistic structure that humans cannot process is not 'part of the domain of linguistic theory in the first place' (p. 234). It may be part of an EXTENSION of language, created artificially through meta-linguistic reasoning, but it would not be part of language as such. Does this thesis apply to ALL non-processable strings? One KNOWS that a given sentence of embedding depth *n* is grammatical, if it results from the application of the appropriate structure-building operations, irrespective of whether or not one can also process it. Even if one is not using one's linguistic competence to reach this conclusion, but rather some form of meta-linguistic reasoning, as C&C suggest, the claim that such sentences are not in the language seems very difficult to sustain. After all, we would never suggest that very large natural numbers are not the same kind of mathematical objects as small natural numbers just because we cannot process them: processability, learnability, etc. just do not seem to provide appropriate criteria for what counts as a sentence, a natural number, and other formal objects. This is one of the points of Chomsky's (1959) critique of behaviorism that should be retained. There IS a 'hard problem': it is hard to see how our understanding of linguistic structure could be reconciled with knowledge of linguistic processes. But C&C's proposed solution is a bitter pill to swallow for many linguists, regardless of their theoretical orientation.

Some psycholinguists may want to keep full-blown accounts of linguistic structure as a guide in empirical inquiry (Jackendoff 2017). C&C correctly consider linguistic processes (biological or cultural evolution, acquisition, neural computation, etc.) as ontologically prior, whereas linguistic structure is inferred or derived as an aspect of 'processing history': e.g. a syntactic tree of a sentence is not psychologically and neurally real; it is a theoretical construct that recapitulates the number and type of chunk-and-pass events at one particular level of linguistic representation. But such theoretical constructs should be seen as EPISTEMOLOGICALLY prior. That is, knowledge of the formal structures of language is required in order to generate knowledge of linguistic processes, e.g. when a set of stimuli, with specific structural features, is constructed for an experiment

https://doi.org/10.1017/S0332586519000258 Published online by Cambridge University Press

on language acquisition or language processing, or when structural features of sentences are specified in order to state the objective function or the supervised learning procedure in artificial neural network models of language (for discussion, see Baggio, van Lambalgen & Hagoort 2012, 2015; Baggio 2018). This point is well illustrated even by C&C's own work, e.g. on recursive structures. The new empiricism advocated by C&C may then be balanced by a new rationalism, which recommends preparatory theoretical research into relevant linguistic structures before observational, experimental, and computational investigations are carried out. The formal core of 'traditional' linguistic theory, including selected aspects of generative syntax and formal semantics, remains potentially highly relevant, even in the context of theories of language evolution, acquisition, and processing. Taking seriously the distinction between ontological priority (process precedes structure) and epistemological priority (structure precedes process) may allow us to address more effectively the 'hard problem' of language science.

Another important point brought up by C&C's discussion of learning or processing constraints is the role of systems of thought (semantics) in shaping language. C&C mention (endorse?) the view that 'compositionality, function-argument structure, quantification, aspect, and modality' are properties of the thoughts that language may express (p. 51). If that is so, then these semantic categories or structures are not culturally evolved/evolvable, but are part of the neurocognitive prior structures against which language (i.e. the system of phonology, morphology, and syntax) evolves by adaptation. The hard, open question is precisely how these semantic categories or structures have arisen, while the question C&C focus on is how they were encoded in language. If we are rejecting UG – and perhaps we should – we may be shifting the problem of universal prior structure for language from syntax to semantics. So what language (or linguistic behavior; see above) is adapting TO would include, as part of innate neural prior structures, a set of universal semantic capacities, such as core ontological categories, computational resources for meaning composition, inference, etc. For example, one can show that compositional structure IN LANGUAGE is a consequence or a response to a bottleneck on cultural transmission (Smith, Brighton & Kirby 2003). But this presupposes that brains CAN compute meaning compositionally. It seems more accurate to say that there is a CAPACITY for compositional processing and representation which is recruited and expressed in language as a result of the effects of a learning bottleneck. The cultural transmission approach also adopted by C&C may not have much to say about this deeper problem, even if it does shed new light on how this capacity (and others) got expressed in languages.

'Creating Language' is an excellent overview of several areas of research at the frontier of language science. It succeeds in weaving together disparate lines of inquiry on language evolution, language acquisition, and language processing. Although it leaves open important questions on the role of formal theories of linguistic structure in a new unified science of language, it contributes several ideas and insights that may help framing those questions correctly. The road to a truly integrative science of language is long and tortuous, and we are barely at the beginning of the journey. Christiansen and Chater's book is an important milestone in that direction.

https://doi.org/10.1017/S0332586519000258 Published online by Cambridge University Press

References

Baggio, Giosuè. 2018. *Meaning in the Brain*. Cambridge, MA: MIT Press.

Baggio, Giosuè, Michiel van Lambalgen & Peter Hagoort. 2012. Language, linguistics and cognition. In Ruth Kempson, Tim Fernando & Nicholas Asher (eds.), *Philosophy of Linguistics*, 325–355. Amsterdam: Elsevier.

Baggio, Giosuè, Michiel van Lambalgen & Peter Hagoort. 2015. Logic as Marr's computational level: Four case studies. *Topics in Cognitive Science* **7**(2), 287–298.

Baggio, Giosuè & Carmelo M. Vicario. 2016. Language processing is not a race against time. *Behavioral and Brain Sciences* **39**, 20–21.

Chomsky, Noam. 1959. A review of B. F. Skinner's *Verbal Behavior*. *Language* **35**(1), 26–58.

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

Christiansen, Morten H. & Nick Chater. 2016. The now-or-never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences* **39**, doi:10.1017/S0140525X1500031X.

Dehaene, Stanislas & Laurent Cohen. 2007. Cultural recycling of cortical maps. *Neuron* **56**(2), 384–398.

Jackendoff, Ray. 2017. In defense of theory. *Cognitive Science* **41**, 185–212.

Jäger, Gerhard. 2008. Applications of game theory in linguistics. *Language and Linguistics Compass* **2**(3), 406–421. doi.org/10.1111/j.1749-818X.2008.00053.x.

Smith, Kenny, Henry Brighton & Simon Kirby. 2003. Complex systems in language evolution: The cultural emergence of compositional structure. *Advances in Complex Systems* **6**(4), 537–558.

Stokhof, Martin & Michiel van Lambalgen. 2011. Abstractions and idealisations: The construction of modern linguistics. *Theoretical Linguistics* **37**(1–2), 1–26.