



Complexity and Relative Complexity in Generative Grammar

Frederick J. Newmeyer*

University of Washington, Seattle, WA, United States

The notions of “complexity” and its antonym “simplicity” have played an important role in the history of generative grammar. However, these terms have been used in different ways. There have been discussions about whether the raw data is complex (or not), about whether a particular theory is complex (or not), and about whether a particular analysis is complex (or not). This article both sorts out the various uses of these terms in the history of generative grammar and demonstrates that motivations have changed over time for whether a complex theory or a simple theory is more desirable. The article concludes with a discussion of the issue of relative complexity in generative grammar, that is, whether the theory embodies the possibility that a grammar of one language can be more or less complex than the grammar of another.

Keywords: Chomsky, complexity, complexity (relative), generative grammar, minimalist program, parameters, simplicity metric, universal grammar

OPEN ACCESS

Edited by:

Kilu Von Prince,
Heinrich Heine University of
Düsseldorf, Germany

Reviewed by:

Terje Londahl,
Norwegian University of Science and
Technology, Norway
Theresa Biberauer,
University of Cambridge, Cambridge,
United Kingdom

*Correspondence:

Frederick J. Newmeyer
fjn@uw.edu

Specialty section:

This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Communication

Received: 05 October 2020

Accepted: 12 February 2021

Published: 19 March 2021

Citation:

Newmeyer FJ (2021) Complexity and
Relative Complexity in
Generative Grammar.
Front. Commun. 6:614352.
doi: 10.3389/fcomm.2021.614352

INTRODUCTION

The notions of “complexity”, and its antonym “simplicity”, have played a major role in the development of generative grammar¹. From the earliest work in the approach to the present day, features of the theory have been evaluated with respect to how “complex” they are with respect to the (uncontestably complex) data that we find in natural language. But as we see in what follows, attitudes have changed with respect to the relationship of the theory to the data, as far as complexity is concerned. In the first 2 decades of the theory, that is until the late 1970’s, the complexity of the theory was extolled. For the next couple of decades (roughly from the mid-1970’s to the mid-1990’s), the theory itself was no longer characterized as “complex”. Rather it was considered to be composed of a set of relatively simple principles, each allowing a number of parameter settings (normally just two). From the interaction of these parameterized principles, the complexity of the observed data was to be derived. For the last twenty-five years or so, we have found universal grammar (UG) described as maximally non-complex, consisting of just the operation Merge (simple recursion) and perhaps some principles relating the output of Merge to the systems that interface with its output. But this gross simplification of UG comes with a price: Much of the data whose analysis was once considered the responsibility of UG is now attributed to these interface systems. UG is less complex, but its explanatory domain is correspondingly reduced.

Each of these stages in the development of the theory was explicitly motivated, though the nature of the motivations changed over time. In the very earliest work, Chomsky and others argued that the complexity of transformational-generative grammar (TGG) was a necessity: Simpler theories were not up to the task of accounting for the full range of grammatical

¹I would like to thank Terje Londahl and Theresa Biberauer for their extremely useful comments.

phenomena in natural language. The theory by the mid-1960's was presented as a model of the cognitive representation of language, where humans are endowed with a rich innate linguistic faculty, namely, UG. The complexity of UG in this period was seen as an asset: The richer UG is, the easier it is to explain how the complexities of a language can be acquired by a child. By the early eighties, many cognitive scientists had adopted a modular view of the human mind, where the apparent complexity of the domain under study was derived from the interaction of autonomous systems, each relatively simple in and of itself. The modular structure of government-binding theory of this decade both reflected and helped further motivate the aforementioned view current among cognitive scientists. The drastically pared down structure of UG in the minimalist program (MP) of today has, in part, an external motivation: The simpler UG is, the more plausible it is that it could have been encoded in the human genome in the process of evolution.

A parallel issue is whether languages (or, more correctly, their grammars) can differ from *each other* in terms of relative complexity. For the most part, this has not been an issue of much concern for generative grammarians. In fact, most generativists would probably argue that the notion of “grammatical complexity” is too obscure to allow languages to be “ranked” along a complexity scale. Nevertheless, a popular view, though one not often argued explicitly, is that a UG perspective entails that all languages be of equal complexity. Such an entailment would follow, it might seem, from the fact that all normal human beings possess the same UG. However, the theory itself allows, in principle, for differential complexity in a variety of ways: There are aspects of language external to UG per se that would seem to require inductive learning, such as peripheral constructions in the syntax, as well as many features of the morphology and phonology. Even the parameterized principles of UG have at times been considered to form part of a hierarchy, where a particular position on the hierarchy might reflect the relative complexity of the phenomenon derived by these principles. Finally, a number of generative grammarians have taken part in the debate on the status of creole languages, some arguing that their (putative) simplicity endows them with a special status with respect to UG, with others arguing that there are no grammatical properties at all that distinguish them from non-creoles.

The paper is organized as follows. *The Three Dimensions of Complexity: the Data to be Explained, the Architecture of the Theory, and the Properties of the Analysis* Section reviews the different types of complexity that have been discussed in the generative literature. *The Changing Attitudes to the Complexity of Universal Grammar in the Development of Generative* Section documents the changing attitudes to the complexity of UG in the development of generative grammar. *The Relative Complexity in Generative Grammar* Section discusses debates among generativists about whether languages can differ in their relative complexity. *The Conclusion* Section is a brief conclusion.

THE THREE DIMENSIONS OF COMPLEXITY: THE DATA TO BE EXPLAINED, THE ARCHITECTURE OF THE THEORY, AND THE PROPERTIES OF THE ANALYSIS

This *Introduction* section discusses the three dimensions of complexity, as discussed in the generative literature: the complexity of the data to be explained (*The Data to be Explained* Section), the complexity of the architecture of the theory (*The Architecture of the Theory* Section), and the complexity of analyses put forth within the theory (*The Adequacy of the Analysis and the Simplicity Metric* Section).

The Data to be Explained

One dimension of complexity in language is that of the data to be explained. No generative grammarian, nor I would assume any other type of grammarian, has denied that the explananda of linguistic theory are complex. References abound in Chomsky's work to “a system as complex as a natural language” (Chomsky, 1965: 192). Indeed, as Chomsky observed several decades later, “As languages were more carefully investigated from the point of view of generative grammar, it became clear that their diversity had been underestimated as radically as their complexity” (Chomsky, 2000: 7). But, “Any complex system will appear to be a hopeless array of confusion before it comes to be understood, and its principles of organization and function discovered” (p. 104). And even more recently, Chomsky and his co-author had no reservations about referring to “the diversity, complexity, and malleability of language” (Berwick and Chomsky, 2016: 107). Nothing more will be said in this article about the undisputed complexity of the raw data that linguists are confronted with.

The Architecture of the Theory

Theories of language can in principle be compared with each other in terms of their relative complexity. But an important caveat is in order. Such comparisons are coherent only if the theories in question have the same ultimate goals. To give a somewhat extreme example, what would it mean to talk about the relative complexity of traditional grammar, as represented by the work of Otto Jespersen, the structuralist grammar of Zellig Harris, and the government-binding theory (GB) proposed within generativism? Given that the underlying assumptions, goals, and methodologies of the three approaches differ in most crucial respects, there is no reasonable way to rank them in terms of their complexity.

The first part of Chomsky's 1957 work *Syntactic Structures* does indeed discuss theories in terms of their relative complexity, in this case finite-state grammars, phrase-structure grammars, and transformational grammars. But in order to carry out this discussion in meaningful way, Chomsky had to reinterpret the assumptions, goals, and methodologies of the advocates of the former two theories as being identical to his own. For example, he began his key chapter of *Syntactic Structures*, “On the goals of linguistic theory”, with the claim that “a grammar of the language L is essentially a theory of L” (Chomsky, 1957: 49). He went on to discuss requirements “that could be placed on the relation between a theory of linguistic structure and particular grammars,” 50). From the

strongest to weakest they comprise a “discovery procedure” for the theory, a “decision procedure”, and an “evaluation procedure”. Chomsky then wrote:

As I *interpret* most of the more careful proposals for the development of linguistic theory, they attempt to meet the strongest of these three requirements. That is, they attempt to state methods of analysis that an investigator might actually use, if he had the time, to construct a grammar of a language directly from the raw data (Chomsky, 1957: 52; emphasis added).

Here we find Chomsky being charitable to his adversaries (if ‘charitable’ is the right word) by attributing to them the same conception—that of regarding a grammar of a language as a theory of that language—that he himself had. Very few linguists at the time would have described their aims in such a manner, a point driven home by the Voegelins, who remarked that “the argumentation employed by transformational-generative grammarians places models of their own making as constructs followed by their predecessors and thereby distorts history” (Voegelin and Voegelin, 1963: 22).

In any event, in later years, we rarely find Chomsky and other generative grammarians comparing their theory with non-generative approaches to language in terms of their relative complexity. We find no shortage of derogatory modifiers used to describe the work of the opponents of generative grammar, ranging from “inadequate” to “incoherent” and everything in between (for an overview, see Newmeyer to appear). However, “overly complex” is not one of them.

The Adequacy of the Analysis and the Simplicity Metric

Virtually every research paper ever written in the generative framework argues that the analysis put forward therein is “less complex” than prior analyses. The complexity comparison might have invoked a major shift in the theoretical apparatus deemed necessary or might merely have referred to a slight tinkering with the formulation of one or another constructs generally agreed to be in the theoretical arsenal. We can see appeals to greater simplicity/less complexity throughout Chomsky’s work. For example, contrasting two possible analyses of the passive within the *Syntactic Structures* framework, Chomsky concluded “that the grammar is much more complex if it contains both actives and passives in the kernel than if the passives are deleted and reintroduced by a transformation that interchanges the subject and object of the active” (Chomsky, 1957: 77). He devoted several pages of *Aspects of The Theory of Syntax* (Chomsky, 1965) to arguing that a theory that allowed for recursion in the base component was less complex than one that handled this phenomenon in the transformational component. Chomsky (1973) provided argument after argument that the single principle of subadjacency was both simpler and more general in its applicative domain than the various individual constraints on movement proposed in Ross (1967). And the MP (Chomsky, 1995) was motivated in great part on simplicity grounds: Among other things, it allowed for the abandonment of the levels of D-structure and S-structure.

The question is how one *knows* that a particular theoretical innovation or technical proposal is less complex than its antecedents. There is no easy answer to this question. In general one appeals to criteria that border on being aesthetic. The simpler, and therefore more desirable, analysis is more elegant and economical in terms what needs to be assumed than its rival. Or perhaps the simpler theory includes data within its explanatory scope that could only have been treated in an ad hoc fashion in the past. Chomsky has always made it clear that in following this path, linguistics is no different from any other science:

Such considerations (involving simplicity, economy, compactness, etc.) are in general not trivial or “merely esthetic” It has been recognized of philosophical systems, and it is, I think, not less true of grammatical systems, that motives behind the demand for economy are in many ways the same as those behind the demand that there be a system at all. Cf. Goodman (1943). (Chomsky, 1979: 1).

In the early days of generative grammar, it was hoped that a formal metric might be devised that would automatically choose the better of two descriptively adequate analyses:

The evaluation metric [also called the “simplicity metric”—FJN] is a procedure that looks at all the possible grammars compatible with the data the child has been exposed to and ranks them. On the basis of some criterion, it says that G_1 is a more highly valued grammar than G_2 , and it picks G_1 , even though G_1 and G_2 are both compatible with the data (Lasnik, 2000: 39).

How could that possibly work? Examples of how the metric might operate usually involved discussion of notational conventions in the formulation of rules. For example, parentheses and brackets in the formulation of phrase structure and transformational rules were chosen so what appeared on intuitive grounds to be the simplest analysis also turned out to be the most compact in its formulation. The metric was referred to in work up to the late 1960s, “but any such measure was more honored in the breach than in the observance” (Aronoff, 2018: 394). Aronoff went on to note that “no useful concrete evaluation metric was ever found” (p. 397). Chomsky himself seemed to abandon the idea of an evaluation metric in his book *Rules and Representations*, writing that the idea that the child tests alternative grammars vis-à-vis an evaluation metric is just a “metaphor” that he doesn’t “think should be taken too seriously” (Chomsky, 1980b: 136). More recently, others have argued that a simplicity metric for syntax is no longer even necessary. Given Chomsky’s speculation that if the parameters of UG relate not to the computational system, but only to the lexicon, “there is only one human language, apart from the lexicon, and language acquisition is in essence a matter of determining lexical idiosyncrasies” (Chomsky, 1991: 419). If so:

[A]cquisition is portrayed not as a construction and comparison procedure, but as merely a procedure of setting “switches” or toggling between fixed options. The child’s mind does not hypothesize alternative grammars, but just grows a single one (McGivray, 2013: 29).

Nevertheless, the difficulties with “switch”-based models operating with “fixed” UG-given parameters are well known (see Fodor and Sakas, 2017 for a useful overview). Even for the now dwindling number of acquisition theorists who operate with a rich UG, the idea that “the child’s mind does not hypothesize alternative grammars, but just grows a single one” no longer holds (see notably Yang 2002 and others adopting variational learner-type models). On a lexical parameters view, it becomes rather implausible to assume parameters to be “fixed” (see the contributions in Biberauer et al., 2014; Picallo, 2014 for some discussion).

Furthermore, the idea of “ranking” alternative grammars in terms of simplicity or similar constructs continues to be popular, and has been developed in different ways for first language acquisition by Roeper (1999) and Fodor (2009) and in constructs such as the “transparency principle”, the “fitness metric” (Clark and Roberts, 1993), a particular “least effort strategy” (Roberts, 1993; Roberts and Anna, 2003; Roberts, 2007) “competing grammars” (Kroch, 2001), and the “tolerance principle” (Yang, 2016).

Given the relative concreteness of phonology as compared to syntax, the simplicity metric had a somewhat longer life in the former subfield than in the latter (for discussion, see Hyman, 1975). But even here serious problems were encountered from the beginning. What should one count in comparing two analyses of the same phenomenon? For example, the number of distinctive features utilized might yield a different complexity result from the number of rules applied. The marking conventions discussed in the Epilogue to Chomsky and Halle (1968) were the last serious attempt to put the simplicity metric into practice. The 1970’s development of different approaches such as lexical phonology, autosegmental phonology, and metrical phonology combined to detract phonologists still further from the goal of developing a formal metric of complexity. Phonologists continue to discuss the idea, however. Durvasula and Liter (2020) offer both a valuable overview of the approaches that have been taken and their own new work on simplicity in phonological learning.

CHANGING ATTITUDES TO THE COMPLEXITY OF UNIVERSAL GRAMMAR IN THE DEVELOPMENT OF GENERATIVE GRAMMAR

This section discusses the changing attitudes to the complexity of UG in the development of generative grammar.² *The Early*

²Chomsky was not to use the term ‘linguistic universal’ until 1962 (Chomsky, 1962a: 536) or refer to ‘universal grammar’ until 1965 (Chomsky, 1965). However, the notion was fully present in *Syntactic Structures*, where he had referred to a ‘condition of generality’ which must be posed by the theory: ‘We require that the grammar of a given language be constructed in accord with a specific theory of linguistic structure in which such terms as “phoneme” and “phrase” are defined independently of any particular language’ (Chomsky, 1957: 50). In what follows I make the simplifying, though strictly speaking incorrect, assumption that Chomsky referred to ‘UG’ in his 1957 book.

Generative Grammar: Universal Grammar is Complex Section discusses why the complexity of UG was considered to be a positive thing in early TGG. By the 1980s, as *The Later Generative Grammar: Universal Grammar is Composed of a Set of Interacting Modules, Each of Which is not Complex* Section points out, UG was considered to be composed of a set of interacting modules, each of which is not complex. And *Current Generative Grammar: Universal Grammar is Simple* Section calls attention to the fact that UG is now considered to be a non-complex faculty and why this is considered to be a good thing.

Early Generative Grammar: Universal Grammar is Complex

In his earliest work, Chomsky never hesitated in describing the theory of TGG as being “complex”, or at least as incorporating more complexity than that of its alternatives. For example, in *Syntactic Structures* he wrote that “The grammar of a language is a complex system with many and varied interconnections between its parts” (Chomsky, 1957: 11). While Chomsky never argued that a complex theory of UG was in and of itself desirable, he did stress that the complexity was necessary to the task of providing adequate grammars of natural languages. As noted above, he contrasted three models of grammatical analysis and opted for the third—the most complex of the three—which allowed for transformational rules. As he went on to remark, “We shall study several different conceptions of linguistic structure in this manner, considering a succession of linguistic levels of increasing complexity which correspond to more and more powerful modes of grammatical description [...]” (Chomsky, 1957: 11). The meat of the book was the demonstration that only the more complex of the three approaches was up to the necessary task. For example:

Once again, as in the case of conjunction, we see that significant simplification of the grammar is possible if we are permitted to formulate rules of a more complex type than those that correspond to a system of immediate constituent analysis.’ (Chomsky, 1957: 41).

What might appear confusing to the modern reader is that at the same time Chomsky also described UG as a “simple” theory:

We must apparently do what any scientist does when faced with the task of constructing a theory to account for a particular subject-matter—namely try various ways and choose the simplest that can be found’ (Chomsky, 1962b: 223)

There is no contradiction here. What Chomsky meant was that TGG was complex compared to finite-state grammars and phrase-structure grammars, but that this necessary complexity allowed for simpler accounts of grammatical phenomena than did its alternatives.

From very early on, Chomsky assumed a “realist” interpretation of linguistic theory, in which “the principles of [a] theory specify the schematism brought to bear by the child in

language acquisition” (Chomsky, 1975: 45). In Chomsky’s opinion, the realist interpretation was “assumed throughout” his mid 1950’s work. But one historiographer of linguistics has asserted in reply that the theory of grammar presented in *Syntactic Structures* was simply “a formal characterization of the distributional structure of a certain set of sentences. It said nothing itself about meaning, or about the psychological basis for the intuitive judgments that speakers make” (Matthews, 1993: 202). That statement appears to be immediately falsified by the following passage from the book, whose realist interpretation seems airtight:

Any grammar of a language will project the finite and somewhat accidental corpus of observed utterances to a set (presumably infinite) of grammatical utterances. In this respect, a grammar mirrors the behavior of the speaker who, on the basis of a finite and accidental experience with language, can produce or understand an indefinite number of new sentences. (1957: 15)

As I have noted in an earlier publication, “If the linguist’s grammar ‘mirrors the behavior of the speaker’, then how could the speaker have failed to internalize the linguist’s grammar?” (Newmeyer, 1996: 208).

Matthews is correct that Chomsky in the above quote did not explicitly refer to the child as a grammar acquirer. That task was taken on by his Ph. D. student Robert B. Lees the same year:

We would not ordinarily suppose that young children are capable of constructing scientific theories. Yet in the case of this typically human and culturally universal phenomenon of speech, the simplest model that we can construct to account for it reveals that a grammar *is* of the same order as a predictive theory. If we are to account adequately for the indubitable fact that a child by the age of five or six has somehow reconstructed for himself the theory of his language, it would seem that our notions of human learning are due for some considerable sophistication (Lees, 1957: 408; emphasis in original).

As I went on to write, “It is true that in 1957, Chomsky considered the grammatical model as a model of ‘behavior’, rather than one of knowledge (Lees had also, on an earlier page, described the grammar as a model of speech behavior). But that is not the issue that concerns us here. Rather, we are addressing the questions of whether Chomsky attributed ‘psychological reality’ (to use a term that he has always despised -- see Chomsky, 1980b: 189–197) to the grammar and whether the child might plausibly be said to have brought to bear the constructs of the theory to the process of language acquisition. The answer appears to be ‘yes’ to both questions” (Newmeyer, 1996: 209).

In the following year, Chomsky’s position with respect to the grammar as a model of internalized competence had become his current one:

[...] it seems to me that to account for the ability to learn a language, we must ascribe a rather complex ‘built-in’ structure to the organism. That is, the [language acquisition device] will have complex properties beyond the ability to match, generalize, abstract, and categorize items in the simple ways that are usually considered to be available to other organisms. In other words, the particular direction that language learning follows may turn out to be determined by genetically determined maturation of complex “information-processing” abilities, to an extent that has not, in the past, been considered at all likely (Chomsky, 1958: 433).

The abovementioned points were important to stress because they bear directly on the issue of complexity. As the previous quote suggests, the complexity of language (and the speed of its acquisition, which he would call attention to in subsequent work) entails that a considerable amount of the properties of language need to be hard-wired into the child. But given the theory as it existed in the first quarter-century of its existence, a complex UG was necessary to account for complex language data. For that reason, the complexity of UG (or its “richness”, to use an alternative term) was seen as a very positive thing. Consider the following quote by way of illustration:

If the system of universal grammar is sufficiently rich, then limited evidence will suffice for the development of rich and complex systems in the mind [. . .]. Endowed with this system and exposed to limited experience, the mind develops a grammar that consists of a rich and highly articulated system of rules, not grounded in experience in the sense of inductive justification, but only in that experience has fixed the parameters of a complex schematism with a number of options (Chomsky, 1980b: 66).

Later Generative Grammar: Universal Grammar is Composed of a Set of Interacting Modules, Each of Which is not Complex

With the advent of the government-binding theory in 1981, UG ceased being described as “complex”. By this point many (though certainly not all) cognitive scientists had begun to regard the human mind as modular in character, that is, composed of relatively simple autonomous subsystems, whose mutual interaction yielded the perceived complexity of the data within its domain (see especially Fodor, 1983). GB was a modular theory par excellence:

The full range of properties of some construction may often result from interaction of several components, its apparent complexity reducible to simple principles of separate subsystems. This modular character of grammar will be repeatedly illustrated as we proceed (Chomsky, 1981: 7).

The GB (or principles-and-parameters) model consisted of the following subsystems of principles: bounding theory, government theory, theta-theory, binding theory, case theory, and control theory. Any grammatical phenomenon, from long-distance movement to anaphora to lexical incorporation typically involved appeal to several of the subsystems, if not all of them. What that meant was that the relationship between theory and data was far more indirect than in early TGG, where grammatical rules often mirrored the phenomena they were designed to account for. And this fact led, in turn, to Chomsky using the term “complexity” in a new sense, namely the complexity of the chain of inference involved in deriving the data from the theory:

Insofar as we succeed in finding unifying principles that are deeper, simpler and more natural, we can expect that the complexity of argument explaining why the facts are such-and-such will increase, as valid (or, in the real world, partially valid) generalizations and observations are reduced to more abstract principles. But this form of complexity is a positive merit of an explanatory theory, one to be valued and not to be regarded as a defect in it (Chomsky, 1981: 15).

In other words, “[in the principles-and-parameters model], argument is much more complex, the reason being that the theory is much simpler; it is based on a fairly small number of general principles that must suffice to derive the consequences of elaborate and language-specific rule systems.” (Chomsky, 1986: 145)

Current Generative Grammar: Universal Grammar is Simple

Chomsky’s current research program is to investigate “how little can be attributed to UG while still accounting for the variety of I-languages attained” (Chomsky, 2007: 3). Indeed, Chomsky now wishes to shift “the burden of explanation from [...] the genetic endowment to [...] language independent principles of data processing, structural architecture, and computational efficiency [...]” (Chomsky, 2005: 9). What has driven this change in Chomsky’s attitude towards a rich UG? In my view, as we have seen, in 1980 his most important goal was to solve the acquisition problem. In that case, one needed to appeal to a rich UG as a way of “easing the burden” on the child. But now, a central goal of Chomsky’s is to solve the evolution problem (see especially Berwick and Chomsky 2016), a problem not on Chomsky’s agenda forty years ago. Clearly, the richer UG is, the more implausible it is that it could have developed by any known processes shaping evolution in general.

In a now classic formulation, “FLN [= the faculty of language in the narrow sense—FJN] comprises only the core computational mechanisms of recursion as they appear in narrow syntax and the mapping to the interfaces” (Hauser et al., 2002: 1,573). What, one might ask, could be simpler than that? The answer depends on what in particular happens in the mapping to the interfaces and to what extent the constructs

appealed to in this mapping form part of our innate endowment for language. While approaches differ, “the mapping to the interfaces” in general encompasses a wide variety of operations. To give one example, “UG makes available a set F of features (linguistic properties) and operations C_{HL} . . . that access F to generate expressions” (Chomsky, 2000: 100). In addition to features and the relevant operations on them, as I noted in earlier work, minimalists have posited principles “governing agreement, labelling, transfer, probes, goals, deletion, and economy principles such as Last Resort, Relativized Minimality (or Minimize Chain Links), and Anti-Locality. None of these fall out from recursion per se, but rather represent conditions that underlie it or that need to be imposed on it. To that we can add the entire set of mechanisms pertaining to phases, including what nodes count for phasehood and the various conditions that need to be imposed on their functioning, like the Phase Impenetrability Condition. And then there is the categorial inventory (lexical and functional), as well as the formal features they manifest” (Newmeyer, 2017: 558). To the extent that these principles are provided by the innate language faculty, that is, UG, UG would appear to be not at all simple.

All of the above principles are syntax-oriented. But there is much more to grammar than syntax, of course. In the claimed drastic reduction of the complexity of UG, where do phonology and morphology, for example, fit in? At first, Chomsky seemed doubtful that the idiosyncrasies of phonology might be amenable to a minimalist treatment, writing that “The whole phonological system looks like an imperfection, it has every bad property that you can think of” (Chomsky, 2002: 118). More recently he has asserted that “If you look at language—one of the things that we know about it is that most of the complexity is in the externalization [the surface manifestation of sound and meaning—FJN]. It is in phonology and morphology, and they’re a mess. They don’t work by simple rules’ (Chomsky 2012: 52). But one should not lose hope:

[T]he mapping to the sound side varies all over the place. It is very complex; it doesn’t seem to have any of the nice computational properties of the rest of the system. And the question is why. Well, again, there is a conceivable snowflake-style answer, namely, that whatever the phonology is, it’s the optimal solution to a problem that came along somewhere in the evolution of language—how to externalize this internal system, and to externalize it through the sensory-motor apparatus.’ (Chomsky, 2012: 40)

I am not sure what to make of the above quote, given the issues that concern us in this article. Chomsky at one and the same time seems to be acknowledging that phonology is complex (because it is filled with irregularity and idiosyncrasy), but asserting that deep-down it is simple (because evolution shaped it snowflake-style). I leave it to the reader to sort out both the interpretation and the implications of his views on the matter.

RELATIVE COMPLEXITY IN GENERATIVE GRAMMAR

We find three different positions in the generative literature on whether languages can differ from each other in terms of their relative complexity: that they are all equally complex (*§Universal Grammar Demands That all Languages be Equally Complex* Section), that they can differ in complexity (*§Universal Grammar Allows for Differences in Complexity Among Languages* Section), and that the notion of “complexity” is so poorly defined that no coherent claims can be made about relative complexity (*§The Notion of “Relative Complexity” of Languages is Incoherent* Section).

Universal Grammar Demands That all Languages be Equally Complex

As early as the 1930’s most structural linguists agreed that the same methods were applicable to languages with a long literary history as to those that had no writing system at all. One could still maintain that position, of course, and accept the idea that the grammars of different languages could be differentially complex. But I know of no mainstream structuralist in the 1950’s who was arguing for differential complexity. Generative grammar, however, with its universalist orientation, made the idea that all languages might be equally complex both intriguing and plausible. As the following quote illustrates for Chomsky in the mid-1950’s was characterizing the grammars of all languages as being “essentially comparable”, despite the “great complexity” of each one:

The fact that all normal children acquire essentially comparable grammars of great complexity with remarkable rapidity suggests that human beings are somehow specially designed to do this, with data-handling or “hypothesis-formulating” ability of unknown character and complexity (Chomsky, 1959: 57).

But if grammars were “essentially comparable”, how might one encode this idea in the theory, while at the same time capturing surface differences? That became possible in 1965 with the introduction of the level of deep structure, as distinct from surface structure:

Modern work has indeed shown a great diversity in the surface structure of languages. However, since the study of deep structure has not been its concern, it has not attempted to show a corresponding diversity of underlying structures, and, in fact, the evidence that has been accumulated in modern study of language does not appear to suggest anything of this sort (Chomsky, 1965: 118).

The above quote leaves open the possibility that surface structures might differ markedly in complexity from language

to language. Fifteen years later, however, Chomsky seemed to dismiss such an idea:

... if, say, a Martian superorganism were looking at us, it might determine that from its point of view the variations of brains, of memories and languages, are rather trivial, just like the variations in the size of hearts, in the way they function, and so on; and it might be amused to discover that the intellectual tradition of its subjects assumes otherwise (Chomsky, 1980a: 77).

A decade later, Chomsky seemed to have taken another step toward embracing the idea that all languages are equally complex:

It has been suggested that the parameters of UG relate, not to the computational system, but only to the lexicon [...]. If this proposal can be maintained in a natural form, there is only one human language, apart from the lexicon, and language acquisition is in essence a matter of determining lexical idiosyncrasies. Properties of the lexicon too are sharply constrained, by UG or other systems of the mind/brain. If substantive elements (verbs, nouns, and so on) are drawn from an invariant universal category, then only functional elements will be parameterized (Chomsky, 1991: 419).

The idea that there is “only one human language” would seem to render absurd the idea that one language might be more complex than another, at least as far as their grammars are concerned. Chomsky did, of course, refer to “lexical idiosyncrasies”. Could they differ in complexity from language to language? Possibly, but it is not clear if Chomsky believes that. In an interview, Chomsky was asked about the “cost” of language-particular lexical peculiarities. It seems to me that one might equate “cost” with “complexity”. When asked if “All languages ought to be equally costly, in this sense?” Chomsky replied: “Yes, they ought to be” (Chomsky, 2004: 165–166).

I have never found any passage where Chomsky has asserted explicitly the idea of universal equal complexity. Nevertheless, several of Chomsky’s intellectual allies have asserted it. The first citation below is from a popular outlining of Chomsky’s ideas, which begins with the following question and assertion: “Why is Chomsky important? He has shown that there is really only one human language: that the immense complexity of the innumerable languages we hear around us must be variations on a single theme” (Smith, 1999: 1). The second citation is from a technical work that contains a glowing Foreword by Chomsky:

Although there are innumerable languages in the world, it is striking that they are all equally complex (or simple) and that a child learns whatever language it is exposed to (Smith, 1999: 168).

Similarly, if we assume biologically determined guidance [in language acquisition], we need to assume that languages do not vary in complexity (Moro, 2008: 112).

Moreover, it has become standard practice for introductory texts with generative orientations to assert equal complexity, as the following three examples show:

There are no “primitive” languages—all languages are equally complex and equally capable of expressing any idea in the universe (Fromkin and Rodman, 1983: 16).

Contrary to popular belief, all languages have grammars that are roughly equal in complexity [...] (O’Grady et al., 1989: 10)

Although it is obvious that specific languages differ from each other on the surface, if we look closer we find that human languages are at a similar level of complexity and detail—there is no such thing as a primitive language (Akmajian et al., 1997: 8).

It is always difficult to put an exact (or even inexact) figure on the percentage of individuals who believe such-and-such, but my impression is that the most generative grammarians would say, if asked, that the theory itself demands that all languages be equally complex³.

Universal Grammar Allows for Differences in Complexity Among Languages

Despite what I have written in §*Universal Grammar Demands That all Languages be Equally Complex* Section, there have been a number of proposals in the generative literature that either allow for or advocate the idea that languages can differ in overall complexity. Let us begin with the issue of parameters and their settings. Chomsky has left no room for doubt that the set of principles and the set of their possible settings are innately provided by UG:

[W]hat we “know innately” are the principles of the various subsystems of S_0 [= the initial state of the language faculty—FJN] and the manner of their interaction, and the parameters associated with these principles. What we learn are the values of these parameters and the elements of the periphery (along with the lexicon, to which similar considerations apply). The language that we then know is a system of principles with parameters fixed, along with a periphery of marked exceptions (Chomsky, 1986: 150–151).

The interesting question is whether parameters can be “ranked” in some sense with respect to each other. Many generative grammarians have replied to this question in the affirmative. As my collaborator John Joseph and I have noted: “the idea that one parameter setting might be more marked than

another has been exploited by a number of generative linguists as a means of characterizing the differential complexity of one grammar vis-à-vis another. Some proposals involving complexity-inducing marked settings have treated preposition-stranding in English and a few other Germanic languages (van Riemsdijk, 1978; Hornstein and Weinberg, 1981), the inconsistent head-complement orderings in Chinese (Huang, 1982; Travis, 1989), and unexpected (i.e., typologically rare) orderings of nouns, determiners, and numerals in a variety of languages (Cinque, 1996). In a pre-parametric version of generative syntax, Emonds (1980) had hypothesized that verb-initial languages are rarer than verb-medial languages because their derivation is ‘more complex’, as it involves a marked movement rule not required for the latter group of languages. Baker (2001) reinterpreted Emonds’ analysis in terms of marked lexical parameters. And Newmeyer (2011) has pointed out that every version of generative syntax has posited syntactic-like rules that apply in the ‘periphery’ or in the mapping from syntax to phonology and are hence exempt from the constraints that might force ‘core grammar’ or the ‘narrow syntactic component’ to manifest equal degrees of complexity in every language” (Joseph and Newmeyer, 2012: 358).

More than a few of Chomsky’s supporters have been troubled by the idea of a plethora of innate parameters in an otherwise “minimalist” approach to language (Newmeyer, 2004; Boeckx, 2011; Newmeyer, 2017). A possible alternative is suggested by Pinker and Bloom:

Parameters of variation, and the learning process that fixes their values for a particular language, as we conceive them, are not individual explicit gadgets in the human mind ... Instead, they should fall out of the interaction between the specific mechanisms that define the basic underlying organization of language (“Universal Grammar”) and the learning mechanisms, some of them predating language, that can be sensitive to surface variation in the entities defined by these language specific mechanisms (Pinker and Bloom, 1990: 183).

An interesting attempt to carry out Pinker and Bloom’s program is Biberauer et al. (2014) and, more recently, Roberts (2019). In their way of looking at things, the child is conservative in the complexity of the formal features that it assumes are needed (what they call “feature economy”) and liberal in its preference for particular features to extend beyond the input (what they call “input generalization”). The idea is that these principles drive acquisition and thus render innately-specified parameters unnecessary, while deriving the same effects. The interest of their work for our purposes is that the “choices” that the child makes in the acquisition process are codified in a set of hierarchies. In their view, it is possible to calculate the grammatical complexity of a language based on the number of choices on the hierarchies needed to fix the grammar of that language. They go so far as to show how complexity indices might be assigned to particular languages (the lower the index, the less complex the language): In their

³As an anonymous reviewer pointed out to me, some generativists would hedge by claiming that the computational systems of all language are equally complex, but not necessarily their grammars taken in their entirety. That might well be Chomsky’s position.

preliminary and admittedly incomplete study, Japanese has a ranking of 1.6, Mohawk 1.8, Mandarin 2, Basque 2, and English 3.

Nowhere has the debate among generative grammarians over whether languages can differ in complexity been as intense as with respect to creoles. Some generativists—I would say a minority—take the position that creoles are simpler than non-creoles, in that they manifest the unmarked parameter settings of UG. The position was argued at length in Bickerton (1984), where he presented his “language bioprogram hypothesis”. Bickerton took as primary evidence for his claim the idea that the (putatively) similar properties of creoles around the world arise from their being “new” languages, which have not had the time to develop marked parameter settings. Bickerton’s hypothesis has been hotly opposed in a number of papers by Michel DeGraff, in particular DeGraff (2001). Among other things, DeGraff argues that the three features that Bickerton claims creoles have in common—verb serialization, a type of complementation, and an approach to tense-modality-aspect marking—are *not* shared by all creoles, and even if they were they would have no relevance to the theory of UG. As DeGraff pointed out, given the data from non-creoles, these particular features bear little relationship to what other have taken to be unmarked features of UG. Nevertheless, other generativists (e.g. Roberts, 1999) have taken creoles to illustrate a stripped down UG, while Jackendoff and Wittenberg (2014) have placed creoles low down on their hierarchy of complexity.

The Notion of “Relative Complexity” of Languages is Incoherent

There is a good reason why only a small number of generative grammarians have taken on the question of whether languages can differ in complexity: Nobody has ever come close to arriving at a metric allowing entire languages to be ranked. Morphology, a relatively concrete component of the grammar, has at times been subject to a complexity metric. The best known example was put forward by Edward Sapir in his book *Language* (Sapir, 1921), which was improved upon in Greenberg (1960). Chomsky and Halle (1968) took on phonological complexity in their book *Sound Pattern of English* (see above, *The Adequacy of the Analysis and the Simplicity Metric* Section). Miller and Chomsky (1963) tried to relate complexity to processing difficulty, as did Hawkins (2004) many years later. Even the index proposed in Biberauer, et al. deals only with morphosyntax. But, in fact, as John Joseph and I pointed out close to a decade ago, “no comprehensive proposal exists to date for measuring the degree of complexity of an entire language, nor is there even agreement on precisely what should be measured” (Joseph and Newmeyer, 2012: 360).

Some linguists, for example (the non-generativist) John McWhorter have correlated degree of complexity of a language with the amount of overspecification, structural elaboration, and irregularity manifested in the language (McWhorter, 2001). The following quote from Aboh and Michel (2017) hits the nail on the head with respect to the

attempts by McWhorter and others to rank languages on a scale of complexity. What they write about creoles would be applicable to any language whatever.

Another fundamental theoretical flaw in the “simplicity” literature on Creoles is the absence of a rigorous and falsifiable theory of “complexity.” Consider, for example, Creole-simplicity claims where complexity amounts to “bit complexity” as defined in DeGraff (2001:265–274). Such overly simplistic metrics consist of counting overt markings for a relatively small and arbitrary set of morphological and syntactic features (see, e.g., McWhorter, 2001; Parkvall, 2008; Bakker et al., 2011; McWhorter, 2011). In effect, any language’s complexity score amounts to the counting of overt distinctions (e.g., for gender, number, person, perfective, evidentiality) and on the cardinality of various sets of signals (e.g., number of vowels and consonants, number of genders), forms (e.g., suppletive ordinals, obligatory numeral classifiers) and “constructions” (e.g., passive, antipassive, applicative, alienability distinction, difference between nominal and verbal conjunction). The problem is that such indices for bit complexity resemble a laundry list without any theoretical justification: “[T]he differences in number of types of morphemes make no sense in terms of morphosyntactic complexity, unless they tell us exactly how overt morphemes and covert morphemes interact at the interfaces, and how they may burden or alleviate syntactic processing by virtue of being overt or covert” (Aboh and Smith, 2009: 7). The problem is worsened when bit-complexity metrics are mostly based on the sort of overt morphological markings that seem relatively rare in the Germanic, Romance, and Niger-Congo languages that were in contact during the formation of Caribbean Creoles (Aboh and Michel, 2017: 417).

In the absence of a scale of complexity that is both theoretically informed and sensitive to all components of the grammar, it seems most prudent to remain agnostic as to whether languages can differ in overall complexity.

CONCLUSION

The notions of “complexity” and its antonym “simplicity” have played an important role in the history of generative grammar. However, these terms have been used in different ways. There have been discussions about whether the raw data is complex (or not), about whether a particular theory is complex (or not), and about whether a particular analysis is complex (or not). Virtually all linguists, including generativists, have agreed that natural language data is complex. Likewise, no generativist would deny that, all other things being equal, a less complex analysis of a particular phenomenon is preferable

to a more complex one. However, the attitude to the complexity of the theory itself has changed over the years. In early TGG, it was stressed that the complex theory presented in the 1950's was superior to its less complex rivals, because only a theory with a particular level of complexity could produce descriptively adequate grammars. By the 1960's it was argued that a complex theory of UG was necessary in order to solve the problem of how a child could master the acquisition of language in such a short period of time. In the 1980's, with the adoption of a modular theory of grammar, UG was conceived as a set of (ideally) simple principles, whose interaction would yield the observed data. Since the 1990's, the theory of UG has been described as "simple". Other systems interacting with UG have taken on much of the burden for accounting for the complexity of the data.

REFERENCES

- Aboh, E. O., and Michel, D. (2017). "A null theory of creole formation based on universal grammar," in *Oxford handbook of universal grammar*. Editor I. Roberts (Oxford, United Kingdom: Oxford University Press), 401–58.
- Aboh E. O., and Smith, N. (Editors) (2009). *Complex processes in new languages*. Amsterdam, Netherlands: John Benjamins.
- Akmajian, A., Demers, R., Farmer, A. K., and Harnish, R. M. (1997). *Linguistics: an introduction to language and communication*. 4th Edn. Cambridge, MA: MIT Press.
- Aronoff, M. (2018). "English verbs in *Syntactic structures*," in *Syntactic structures after 60 years: the impact of the Chomskyan revolution in linguistics*. Editors N. Hornstein, H. Lasnik, P. Patel-Grosz, and C. Yang (Berlin, Germany: DeGruyter Mouton), 381–403.
- Baker, M. C. (2001). *The atoms of language: the mind's hidden rules of grammar*. New York, NY: Basic Books.
- Bakker, P., Daval-Markussen, A., Parkvall, M., and Plag, I. (2011). Creoles are typologically distinct from non-creoles. *JPLCL* 26, 5–42. doi:10.1075/jplcl.26.1.02bak
- Berwick, R. C., and Chomsky, N. (2016). *Why only us: language and evolution*. Cambridge, MA: MIT Press.
- Biberauer, T., Holmberg, A., Roberts, I., and Sheehan, M. (2014). "Complexity in comparative syntax: the view from modern parametric theory," in *Measuring grammatical complexity*. Editors F. J. Newmeyer and L. B. Preston (Oxford, United Kingdom: Oxford University Press), 103–27.
- Bickerton, D. (1984). The language bioprogram hypothesis. *Behav. Brain Sci.* 7, 173–188. doi:10.1017/s0140525x00044149
- Boeckx, C. (2011). "Approaching parameters from below," in *Biolinguistic approaches to language evolution and variation*. Editors A. M. Di Sciullo and C. Boeckx (Oxford, United Kingdom: Oxford University Press), 205–21.
- Chomsky, N., and Halle, M. (1968). *The sound pattern of English*. New York, NY: Harper and Row.
- Chomsky, N. (2007). *Approaching UG from below. Interfaces + recursion = language?* Editors U. Sauerland and H. M. Gärtner (Berlin, Germany: Mouton de Gruyter), 1–29.
- Chomsky, N. (1959). A Review of B. F. Skinner's *Verbal Behavior*. *Language* 35, 26–57.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. (1973). *Conditions on transformations*. A *FESTSCHRIFT for Morris Halle*. Editors S. Anderson and P. Kiparsky (New York, NY: Holt Rinehart and Winston), 232–86.
- Chomsky, N. (1980a). *Discussion. Language and learning: the debate between Jean Piaget and Noam Chomsky*. Editor M. Piattelli-Palmarini (Cambridge, MA: Harvard University Press), 73–83.
- Chomsky, N. (1980b). *Rules and representations*. New York, NY: Columbia University Press.
- Chomsky, N. (1962a). "Explanatory models in linguistics," in *Logic, methodology, and philosophy of science*. Editors E. Nagel, P. Suppes, and A. Tarski (Stanford, CA: Stanford University Press), 528–550.
- Chomsky, N. (1962b). "A transformational approach to syntax. Proceedings of the third Texas conference on problems of linguistic analysis in English," in *The structure of language: readings in the philosophy of language*. Editors J. Fodor and J. Katz 1964 Edn. (Englewood Cliffs, NJ: Prentice-Hall), 211–243. Reprinted in. 1964. Editors A. Hill, 124–58. (Austin: University of Texas Press)
- Chomsky, N. (1986). *Knowledge of language: its nature, origin, and use*. New York, NY: Praeger.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, Netherlands: Foris.
- Chomsky, N. (1958). *Linguistics, logic, psychology, and computers. Computer programming and artificial intelligence; an intensive course*. Editor J. W. Carr, III (Ann Arbor, MI: University of Michigan College of Engineering), 429–54.
- Chomsky, N. (1975). *Reflections on language*. New York: Pantheon.
- Chomsky, N. (1979). *Morphophonemics of modern Hebrew*. New York, NY: Garland.
- Chomsky, N. (2000). *New horizons in the study of language and mind*. Cambridge, MA: Cambridge University Press.
- Chomsky, N. (2002). *On nature and language*. Cambridge, MA: Cambridge University Press.
- Chomsky, N. (1991). "Some notes on economy of derivation and representation," in *Principles and parameters in comparative grammar*. Editor R. Freidin and R. A. Freidin (Cambridge, MA: MIT Press), 417–54. [Reprinted in *The Minimalist Program* by Noam Chomsky Cambridge, MA: MIT Press (1995). [129–166].
- Chomsky, N. (1957). *Syntactic structures*. The Hague, Netherlands: Mouton.
- Chomsky, N. (2004). *The generative enterprise revisited: discussions with Riny Huybregts, Henk van Riemsdijk, Naoki Fukui, and Mihoko Zushi*. Berlin, Germany: Mouton de Gruyter.
- Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.
- Chomsky, N. (2012). *The science of language: interviews with James McGilvray*. Cambridge, MA: Cambridge University Press.
- Chomsky, N. (2005). Three factors in language design. *Linguist. Inq.* 36, 1–22. doi:10.1162/0024389052993655
- Cinque, G. (1996). The 'antisymmetric' program: theoretical and typological implications. *J. Linguist.* 32, 447–65. doi:10.1017/s0022226700015966
- Clark, R., and Ian, Roberts. (1993). 'A computational theory of language learnability and language change. *Linguistic Inquiry.* 24, 299–345.
- DeGraff, M. (2001). On the origin of creoles: a cartesian critique of neo-Darwinian linguistics. *Linguist. Typology* 5, 213–310.
- Durvasula, K., and Liter, A. (2020). There is a simplicity bias when generalizing from ambiguous data. *Phonology* 37, 177–213. doi:10.1017/s0952675720000093
- Emonds, J. E. (1980). Word order in generative grammar. *J. Linguist.* 1, 33–54.
- Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.
- Fodor, J. D. (2009). "Syntax acquisition: an evaluation measure after all?," in *Of minds and language: a dialogue with Noam Chomsky in the Basque country*.

AUTHOR CONTRIBUTIONS

This is a review of the treatment of complexity and relative complexity in generative grammar.

- Editors M. Piattelli-Palmerini, P. Salaburu, and J. Uriagereka (Oxford, United Kingdom: Oxford University Press), 256–277.
- Fodor, J. D., and Sakas, W. G. (2017). *Learnability. Oxford handbook of universal grammar*. Editor I. Roberts (Oxford, United Kingdom: Oxford University Press), 249–269.
- Fromkin, V. A., and Rodman, R. (1983). *An introduction to language*. 3rd Edn. New York, NY: Holt, Rinehart and Winston.
- Goodman, N. (1943). On the simplicity of ideas. *J. Symb. Log.* 8, 107–21. doi:10.2307/2271052
- Greenberg, J. H. (1960). A quantitative approach to the morphological typology of language. *Int. J. Am. Linguist.* 26, 192–220. doi:10.1086/464575
- Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science* 298, 1569–1579. doi:10.1126/science.298.5598.1569
- Hawkins, J. A. (2004). *Efficiency and complexity in grammars*. Oxford, United Kingdom: Oxford University Press.
- Hornstein, N., and Weinberg, A. (1981). Case theory and preposition stranding. *Linguist. Inq.* 12, 55–92.
- Huang, C. T. J. (1982). Logical relations in Chinese and the theory of grammar. Available at: <http://www.ai.mit.edu/projects/dm/theses/huang82.pdf>. MIT unpublished Ph. D. dissertation
- Hyman, L. M. (1975). *Phonology: theory and analysis*. New York, NY: Holt, Rinehart and Winston.
- Jackendoff, R., and Wittenberg, E. (2014). “What you can say without syntax: a hierarchy of grammatical complexity,” in *Measuring grammatical complexity*. Editors F. J. Newmeyer and L. B. Preston (Oxford, United Kingdom: Oxford University Press).
- Joseph, J., and Newmeyer, F. J. (2012). All languages are equally complex: the rise and fall of a consensus. *Historiogr. Linguist.* 39, 341–368. doi:10.1075/hl.39.2-3.08jos
- Kroch, A. (2001). “Syntactic change,” in *The handbook of contemporary syntactic theory*. Editors M. Baltin and C. Collins (Oxford, United Kingdom: Blackwell), 699–729.
- Lasnik, H. (2000). *Syntactic structures revisited: contemporary lectures on classic transformational theory (with Marcela Depiante and Arthur Stepanov)*. Cambridge, MA: MIT Press.
- Lees, R. B. (1957). Review of *Syntactic-Structures* by Noam Chomsky. *Language* 33, 375–408. doi:10.2307/411160
- Lightfoot, D. (1979). *Principles of diachronic syntax*. Cambridge, MA: Cambridge University Press.
- Matthews, P. H. (1993). *Grammatical theory in the United States from BLOOMFIELD to Chomsky*. Cambridge, MA: Cambridge University Press.
- McGilvray, J. (2013). “The philosophical foundations of biolinguistics,” in *Cambridge handbook of biolinguistics*. Editors C. Boeckx and K. K. Grohmann (Cambridge, MA: Cambridge University Press), 22–46.
- McWhorter, J. H. (2001). The world’s simplest grammars are creole grammars. *Linguist. Typology* 5, 125–66. doi:10.1515/lity.2001.001
- McWhorter, J. H. (2011). *Linguistic simplicity and complexity: why do languages undress?* Berlin, Germany: Mouton DeGruyter.
- Miller, G., and Chomsky, N. (1963). “Finitary models of language users,” in *Handbook of mathematical psychology*. Editors P. Luce, R. Bush, and E. Galanter (New York, NY: Wiley) Vol. 2, 419–92.
- Moro, Andrea (2008). *The boundaries of Babel: the brain and the enigma of impossible languages*. Cambridge, MA: MIT Press.
- Newmeyer, F. J. (2004). Against a parameter-setting approach to language variation. *Linguistic Variation Yearbook* 4, 181–234. doi:10.1075/livy.4.06new
- Newmeyer, F. J. (2011). Can one language be ‘more complex’ than another? Unpublished paper, UW, UBC, and SFU. Available at: <https://linguistics.washington.edu/complexity-slides.pdf> (Accessed March 5, 2021).
- Newmeyer, F. J. (1996). Review of grammatical theory in the United States from BLOOMFIELD to Chomsky. (*Historiographia Linguistica*) 23, 200–210.
- Newmeyer, F. J. (2017). “Where, if anywhere, are parameters?: A critical historical overview of parametric theory” in *On looking into words (and beyond): structures, relations, analyses*. Editors C. Bower, L. Horn, and R. Zanuttini (Berlin, Germany: Language Sciences Press), 545–66.
- O’Grady, W., Michael, D., and Aronoff, M. (1989). *Contemporary linguistics: an introduction*. New York, NY: St. Martin’s Press.
- Parkvall, M. (2008). “The simplicity of creoles in a cross-linguistic perspective,” in *Language complexity: typology, contact, change*. Editors M. Miestamo, K. Sinnemäki, and F. Karlsson (Amsterdam, Netherlands: John Benjamins), 265–85.
- Picallo, C. (2014). *Linguistic variation in the minimalist framework*. Oxford, United Kingdom: Oxford University Press.
- Pinker, S., and Paul, B. (1990). Natural language and natural selection. *Behav. Brain Sci.* 13, 707–84.
- Roberts, I., and Roussou, A. (2003). *Syntactic change: a minimalist approach to grammaticalization*. Cambridge, MA: Cambridge University Press.
- Roberts, I. (2007). *Diachronic syntax*. Oxford, United Kingdom: Oxford University Press.
- Roberts, I. (2019). *Parameter hierarchies and universal grammar*. Oxford, United Kingdom: Oxford University Press.
- Roberts, I. (1993). *Verbs in diachronic syntax: comparative history of English and French*. Dordrecht, Netherlands: Kluwer.
- Roberts, I. (1999). “Verb movement and markedness,” in *Language creation and language change: creolization, diachrony, and development*. Editor M. DeGraff (Cambridge, MA: MIT Press), 287–327.
- Roeper, T. (1999). Universal bilingualism. *Biling. Lang. Cogn.* 2, 169–186.
- Ross, J. R. (1967). Constraints on variables in syntax. MIT PhD thesis. Norwood (NJ): MIT. (published in 1985 as ‘Infinite syntax!’).
- Sapir, E. (1921). *Language*. New York, NY: Harcourt, Brace, and World.
- Smith, N. (1999). *Chomsky: ideas and ideals*. Cambridge, MA: Cambridge University Press.
- Travis, L. (1989). “Parameters of phrase structure,” in *Alternative conceptions of phrase structure*. Editors M. R. Baltin and A. S. Kroch (Chicago, IL: University of Chicago Press), 263–79.
- van Riemsdijk, H. (1978). *A case study in syntactic markedness: the binding nature of prepositional phrases*. Dordrecht, Netherlands: Foris.
- Voegelin, C. F., and Voegelin, F. M. (1963). On the history of structuralizing in 20th century America. *Anthropol. Linguistics* 5, 12–35.
- Yang, C. (2002). *Knowledge and learning in natural language*. Oxford, United Kingdom: Oxford University Press.
- Yang, C. (2016). *The price of linguistic productivity: how children learn to break the rules of language*. Cambridge, MA: MIT Press.

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Newmeyer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.