Language complexity: An insight from complex-system theory

Alexander Andrason
Dep. of African Languages, University of Stellenbosch, Stellenbosch, South Africa

Email address: andrason@sun.ac.za

To cite this article:

Abstract: This paper aims to establish a more direct relation between the studies of complexity in the field of typological-evolutionary linguistics and complex-system theory. The article explains what complex-system theory can tell us about language complexity and how insights from the science of complex systems can be important to the analysis of linguistic complexity. By founding his argumentation on the principles of complex-system theory, the author maintains that linguistic complexity cannot be equaled to cardinality. Although a comparison of complexities can be effectuated only in numerical or set-theoretical terms, the cardinality of a series or a set is neither unique nor the most important property of a complex system. Equally or even more relevant features are openness, situatedness, lack of boundaries, individual instability, uncertainty, non-linearity, exponential sensitivity to initial conditions, dynamicity, metastability, path dependency, emergence, regional chaos, non-additivity, non-modularization, irreducibility, organizational intricacy, and models’ incompressibility, incompleteness, provisionality or plurality. The author argues that, following complex-system theory, once a distinction between the complexity of language as a real-world phenomenon and the complexity of its model is made, any numerical comparison of the overall or local complexity of languages becomes either futile or deeply theory conditioned. In realistic complex systems, complexity is always infinite, while in models – where, by means of fictionalized approximations adopted by an observer or explainer, it can be made finite – complexity entirely depends on a scientific frame of reference and approach with which it has been quantified. Accordingly, any quantification of the complexity of natural languages is either identical, as it is infinite (in realistic languages), or relative (in scientific models). In this manner, no measurement may claim to establish the ultimate hierarchy of less or more complex languages.

Keywords: Complexity, Complex-System Theory, Typological-Evolutionary Linguistic Complexity, Modelling

1. Introduction

1.1. TEC Approach

In the last fifteen years, the term ‘complexity’ has become common in linguistics. One faction of linguists among those who are engaged in the issue of complexity has vividly questioned the view, typically acclaimed in the 20th century, whereby all the languages are equally complex and therefore insignificant for the measurement or comparison of their complexity. In contrast, these scholars have begun to compare languages in respect to their complexity and analyze its modifications across periods. As a provisional result of their paradigm-breaking studies, it has been proposed that languages display distinct degrees of complexity and that, with time, natural linguistic systems become more complex. Thus, complexity is viewed as being both a synchronic and diachronic variable: its range changes cross-linguistically and increases from an evolutionary perspective (Trudgill 2001 [1], 2004 [2], Sampson 2001 [3], 2009 [4], Dahl 2004 [5], 2009 [6], 2011 [7], Gil 2006 [8], 2007 [9], 2008 [10], Hawkins 2004 [11], Miestano 2008 [12], 2009 [13], McWorter 2005 [14], 2009 [15], Nichols 2007a [16], 2007b [17], 2009 [18], Miestamo, Simmelkäi & Karlsson 2008 [19], Sampson, Gill & Trudgill 2009 [20], Bisang 2009 [21], Progovac 2009 [22], Givón 2009 [23], Givón & Shibatani 2009 [24], Kortmann & Szmarcianski 2012 [25]). Being aware of individual distinctions that exist among all these linguists, in the remaining portions of the present paper, merely for the sake of simplicity, this group of scholars and their treatment of complexity in living languages will be referred to as a ‘typological and
1.2. The Notion of Complexity in the TEC Approach

In their discussion of the issues of complexity in natural languages, linguists who adhere to the TEC faction usually distinguish two main sub-types: agent-related complexity and absolute complexity. The former is relative and refers to the difficulty (so-called ‘cost’) with which the users or second language speakers acquire a given language (Trudgill 2001:371 [1], Kusters 2003 [26], Miestano 2008 [12] and 2009:81-82 [13], Szmrecsanyi & Kortmann 2012:16 [27]). The latter is claimed to be more objective (and, therefore, more suitable for scientific treatment) and refers to the system’s realistic complexity and especially to the complexity of its grammar and rules (Dahl 2004 [5], 2009:50-54 [6], Miestano 2008 [12], 2009:81-82 [13]). The absolute complexity itself is understood in two manners, following either the idea of Kolmogorov complexity (also known under the notion of algorithmic, descriptive or program-size entropy) or the notion of Gell-Mann complexity (also denominated as effective complexity). 

Kolmogorov complexity measures the computability requirements for specifying a given system (be it individual, phenomenon or feature) and quantifies the disorder or randomness of its descriptive series: the more complex, the longer the necessary series of description is. Accordingly, a series for randomness is the longest and the most complex, while complete regularity is the shortest and the least complex (Li & Vitányi 2008 [28]).

Gell-Mann complexity – treated as entropy that characterizes a system with information concerning its regularity – quantifies the computability of non-random information in the system, i.e. order introduced by rules, as contrary to randomness or disorder. It specifies how complex the rules governing a system are: the more complex, the longer the description of regularities is. In this manner, Gell-Mann complexity enables us to distinguish complex systems, characterized by a high organizational depth, from systems that are composed by a large number of components but in a simple repetitive and/or random manner. By specifying organizational intricacy, Gell-Mann complexity appears to be closer to an intuitive understanding of what complex means and constitutes a type of complexity that is exemplary to models of real-world complex systems, whether they represent ecological structures, organisms or social networks. When represented in models, real-world complex systems are typically complex in the sense of Gell-Mann definition (Gell-Mann 1995 [29], Gell-Mann & Lloyd 1996 [30], 2004 [31], Gell-Mann & Tsallis 2004 [32], Köhler, Altmann & Piotrowski 2005 [33], Cilliers et al. 2013 [34]).

Gell-Mann type of complexity quantification is considered to be more appropriate for linguistics. When measuring language complexity, scholars typically ask the following question: for a given content, how complex are the rules that enable this content to be expressed (Miestano 2009:81-82 [13], Dahl 2011:154-155 [7]; see also McWorter 2005 [14] and 2009 [15])? Accordingly, the complexity of a linguistic system is viewed as ornamental rule intricacy (opposed to rule simplicity), which refers to the number of sumptuously complex traits (McWorter 2001 [35], 2007 [36], Szmrecsanyi & Kortmann 2012:16 [27]).

In order to quantify the complexity of a grammar, linguists usually take into account the amount of information required to describe a rule governing given content. In particular, they analyze the number and variety of elements involved and their interactions (the number of paradigmatic variants, number of syntagmatic dependencies among elements, and number of constraints on elements and their combinations; Trudgill 2004 [2], McWorter 2005 [14], 2009 [15], Szmrecsanyi & Kortmann 2009 [37], Nichols 2009:111-114 [18]; see also Dahl 2011 [7]).

In most cases, when quantifying the global complexity of a linguistic system – i.e. when determining how complex a given language is in its totality – scholars limit themselves to establishing a rather restricted set of distinctions or categories. This limitation – of which all the TEC scholars are well aware – stems from the fact that the measurement of global complexity and, thus, its comparison among languages is extremely difficult. This difficulty involves more two specific problems: a problem of representativity (it is virtually impossible to account for all the features or aspects of grammar so one is compelled to determine a representative sample) and a problem of comparability (it is impossible to quantify different areas and features of grammar in identical terms or digits because such numbers represent entirely different and incomparable realistic properties; Miestano 2006 [38], 2008 [12] and 2009:83 [13], Nichols 2009:113-114 [18], Deutscher 2009:248-250 [39] and Dahl 2011 [7]). Aiming at minimizing the effects of these weaknesses, scholars narrow global complexity to local (fragmentary or domain specific) complexity (Miestano 2009:82 [13]), while global complexity tends to be considered only in closely related idioms, where the

---

5 Sub-types of such quantification usually involve quantitative complexity (structural elaboration or a number of components and rules without taking into account their qualitative aspects), redundancy-induced complexity (extent of overspecification), irregularity-induced complexity (non-transparency), language-external induced complexity (e.g. sensitivity to sociological parameters; for a more detailed description of all these subtypes, see Szmrecsanyi & Kortmann 2012:10-13 [27]).

6 Accordingly, the computation of complexity commonly involves more specific measurements such as an “overt grammatical analyticity index” (frequency of free/analytical markers), “overt grammatical syntheticity index” (frequency of bound/synthetic markers), “overt grammaticity index” (frequency of any grammatical markers) and “irregularity index” (frequency of irregular markers; Szmrecsanyi & Kortmann 2012:17 [27]).

7 When applied to specific linguistic domains, complexity triggers the following local varieties: phonological complexity, morphological complexity, syntactic complexity, lexical complexity and pragmatic complexity (see Szmrecsanyi & Kortmann 2012:9-9 [27] and the references therein).
relation of set-theoretic inclusion can be more successfully established (Deutscher 2009:250 [39], Dahl 2009 [6] and 2011:155-156 [7]).

1.3. Complex-System and TEC Approach – Objective of the Paper

As is evident from the preceding discussion, the TEC scholars usually equate complexity to cardinality, and in particular to the cardinality of components (a number of constituents of the system) and the cardinality of rules (a number of rules that organize components and constraint their combinations). However, according to complex-system theory – which is a modern framework designed specifically to treat and control various aspects of complexity – cardinality of constituents or rules is only one of many features and properties of complexity. To put it differently, cardinality by itself does not render a system complex (see, section 2.1., below).

Generally, despite the steady growth of interest in the issue of linguistic complexity, the TEC scholars seem to have paid less attention to complex-system theory, which, as mentioned above, is the science explicitly concerned with real-world complexity and its modelling. De facto, complex-system theory constitutes a theoretical basis that is typically employed when treating the complexity of the real world in science, in all of its aspects, domains and levels (Schlindwein & Ison 2007:236-238 [40], Auyang 1998 [41], Hooker 2011a [42]). This framework has been extensively applied to the study of physical, chemical and biological systems and well as the analysis of other non-physical aspects of human life, be they economical, sociological or cultural (Auyang 1998 [41], Kauffmann 2000 [43], Lewin 2000 [44], Hooker 2011a [42]). Nevertheless, while complex-system scientists (be they physicists or philosophers) have included language in their examinations, emphasizing its prototypical complex-system behavior and need to be analyzed within the complex-system framework (Ceinarova 2005 [45] and Solé 2010 [46]) and other branches of linguistics have sometimes referred to certain advances and findings of complexity science (especially neuro-linguistics, psycholinguistics, sociolinguistics and of course computational linguistics; cf. for instance Pinker 1994 [47], Li & Vitanyi 1995 [48], Herdina & Jessner 2002[49], Culicover & Nowak 2003 [50], Köhler, Altmann & Piotrowski 2005 [33], Massip-Bonet 2013 [51], Munné 2013 [52], Muñwene 2013 [53]; see also Jenner, van Peursen & Talstra 2006 [54] and Andrason 2012 [55]), the TEC approach – which is specifically concerned with the idea of linguistic complexity – has been much less receptive. Even though it has adapted some important insights from fields related to complexity such as information theory and information processing, the direct influence of complex-system theory on the TEC approach and its definition of complexity has been marginal, limited to a few general statements (see, however, Dahl 2009 [6] and 2011 [7], Kortmann & Szmacien 2012 [25] and especially Sinnemäki 2011 [56]). As a result, TEC linguists have been talking about complexity without giving a necessary prominence to the framework which has been developed in order to talk about complexity.

Without undermining the relevance and ground-breaking work of the TEC researchers, the author of the present article aims at correcting a certain deficiency in the field of linguistic complexity discussed by the TEC approach, evident in the shortage of a direct methodological relationship with complex-system theory. Accordingly, this paper will introduce the principal ideas of complex-system theory that are of high importance for the study of complexity in linguistics. By demonstrating what complex-system theory can tell us about language complexity and which – as well as how – insights of complex thinking can be important to the analysis of linguistic complexity, this article aspires to pave the way for a more systematic use of complex-system theory in the discussion on linguistic complexity. As a result, it may open new ways of methodological treatment and systematic comprehension of certain problems, encountered by the TEC linguistics, and thus bestow scholars with an alternative means of addressing old questions or proposing their solutions in novel – until now, less appreciated – models. It must be emphasized that this article should by no means be understood as an attempt to attack the TEC approach or question its pioneering findings and highly valuable advances. It rather seeks to offer an alternative perspective on the problematic of linguistic complexity by directly linking this area of studies to the core of complexity science, viz. complex-system theory.

In order to accomplish this goal, the study will be organized in the following manner: in the next part of the article, the properties of real-world complex systems, as posited by complexity science, will be presented (section 2.1.) and the issue of their modelling explained (section 2.2.). Afterwards, in an analogical manner, the notion of language as a real-world complex system – with all its prototypical complex characteristics – will be introduced (section 3.1.) and the problem of its representation in models discussed (section 3.2.). Finally, main conclusions will be drawn and a line of further studies on language complexity, emerging from the findings of this paper, suggested (section 4.).

2. Real World

As acknowledged by complex-system theory, the idea of complexity underlies all real-world organizations. Complexity is present everywhere. However, its definition – just like the systems to which it applies – is far from simple 6 and is codified in an accumulative manner as a set of more specific properties. To be precise, a system is

---

6 According to Edmond (1999), there are at least forty definitions of complexity (cf. also Horgan 1995:75 [57], Franco Parellada 2007:154 [58], and Ceinarova 2005:16, 57 [45]; see also Lloyd 2001 [59]). Sometimes it is also claimed that “complexity […] appears as essentially undefinable in any way that allows objective measurements (Ayers 1994:13-14 [60])."
complex if it displays some or all of the following properties: it is open, situated, boundary-free and replete with unstable individual; “infinitively” cardinal, uncontrollable and uncertain; dynamic, metastable and path dependent; non-linear, sensitive to initial conditions, exponentially amplifiable and in regions chaotic; emergent, non-additive, non-modularizable, irreducible and organizationally intricate. It is also self-organizing and adaptive. Where its modelling is concerned, a complex system is typically incompressible, model-specific and model-plural (Cilliers 1998 [61] and 2005 [62], Schlindwein & Ison 2007:232 [40], Wagensberg 2007:12, 27, 56-62 [63], Hooker 2011b:20-21, 40 [64], Bishop 2011:112 [65], Cilliers et al. 2013:2-4 [34]).

In this part, the said exemplary properties of complex systems will be described in detail and their contribution to the overall intricacy of such organizations explained. First, the author will discuss the complexity of realistic natural systems (section 2.1.) and next the characteristics of its modelling (section 2.2.).

2.1. Real-World Complexity

2.1.1. Complexity Means Openness, Situatedness, Lack of Boundaries and Fluidity of Individuality

Real-world complex systems are never isolated. As they exchange material, energy and information with the environment, they are inherently open and relational. The concepts of openness and interaction with the “external” world are so relevant that all natural systems – as well as their parts and constituents – are viewed as essentially situated entities: their behavior depends not only on the parts of which they are composed but also on the whole(s) in which they are embedded. Given that various important properties of the system are dictated by its global situation and that the essence of a constituent derives from non-interiorized relations, the line between the system and its environment becomes fluid and the concept of boundary highly problematic. In fact, a clear distinction between the system and its external background, as well between parts and wholes, is pragmatic rather than real. In nature, rigid and permanent boundaries do not exist – we merely draw them according to our needs. Consequently, the environment in which the system is inserted constitutes this system’s important part. The environment participates in the system’s behaviour and regulates it, being in turn simultaneously influenced by the system, which it frames. It is impossible to determine which fragments of the environment are irrelevant for the system – and, thus, unconnected to it – because even the smallest value in the “external” universe can have a substantial (including catastrophic) impact on the system due to non-linearity and an exponential amplification of the error margin (cf. section 2.1.4). Furthermore, since boundaries are arbitrary and “external” relations may constitute highly relevant characteristics of an entity, the concept of individuality is undermined. Individuals, rather than being stable, form fluid hierarchies of individuality: a lower-level individual is always a part of a higher-level individual and the properties of the latter importantly contribute to the state and behavior of the former (cf. Cilliers 1998:4 [61], Auyang 1998:47, 121 [41], Schneider & Sagan 2009:141-142, 376-377 [66], Prigogine 2009:177 [67], Richardson, Mathieson &Cilliers 2000 [68], Hooker 2011b:23, 31-35, 43 [64], Bickhard 2011:98-101, 112, 115, 127 [69], Cilliers et al. 2013:2 [34]).

2.1.2. Complexity Means Infinite Cardinality, Uncontrollability and Uncertainty

Real-world complex systems contain an immeasurable number of components. Since individuals are fluid (they can be decomposed into more basic constituents and composed into larger singularities) and since the boundaries of the system are arbitrary (in order to satisfy the system, this system should comprise everything, including the environment, in which it is embedded), the cardinality – or the total number of the participating elements – is infinite. Even if we select a finite set of components, the amount of possible configurations is infinite or radically uncontrollable due to the non-linear nature of the relations that exist among them (cf. section 2.1.4). As everything interacts with everything else – every entity somehow affects the state of the remaining entities, being simultaneously affected by all the other components and the system globally – the network of interconnections and possible states that emerge from them is absolutely untreatable. In fact, it is relations – even more than constituents – that render complex systems entirely uncontrollable. Relations constitute the core of complex systems – they cannot be understood as external and exogenous to the system’s constituents because constituents are not mere aggregates of isolated individuals but strongly depend on multi-level (micro- and macroscopic) interactions. Relations in complex systems are typically non-linear and create feedback loops (the results of an action feed back onto itself).

Apart from the infinite number of components, relations and configurations, the infiniteness of complex systems surfaces in yet another manner. When determining the state of a system or even one of its components, it is impossible to provide a complete series by which it could be fully represented. To ultimately satisfy such a description, an infinite amount of information would be needed, which is physically impossible. This stems from the fact that there is no limit to the longitude of an empirical series that represents realistic phenomena – therefore the series can be extended indefinitely. By increasing the longitude of the sequences of data, at a certain point, any two series will always diverge. This is related to the fact that all the realistic contexts are unique and no two phenomena are indistinguishable. Since we must impose limits when describing an object or phenomenon, an infinite portion of data must be put aside in determining a series, which

1Boundaries and individuals are also questioned because of the inherently dynamic nature of complex systems: as everything is a process, constituents cannot be fully individualized.
implies that there will always be an inherent disturbance or uncertainty in defining the state of a system or a component. Due to the phenomena of non-linearity and exponential amplification, this uncertainty will have an unpredictable effect on the behaviour of the system (Auyang 1998:344 [41], Richardson, Cilliers & Lissack 2007:33 [70], Wagensberg 2007:27, 56-60 [63], Schneider & Sagan 2009:55 [66], Bishop 2011:116-117, 121-123 [65], Cilliers et al. 2013:2 [34]).

2.1.3. Complexity Means Metastability and Historicity

The complexity of natural systems is additionally augmented by the fact that such organizations are inherently evolving. Reality is dynamic and time is a central concept in real-world organizations. It is not enough to describe static properties of the system and its components as they appear at a time \( t_0 \). In fact, “[a]ny analysis of a complex system that ignores the dimension of time is incomplete” (Cilliers et al. 2013:2 [34]). One must provide information about the system’s dynamics. This includes all its past states (as well as the states of the “external” environment) and equations regulating this organization’s development. These equations, on the one hand, relate the system’s past to its present and, on the other, predict its possible future behaviors. Although various real-world objects are, for certain reasons, regarded as static “things”, they are in fact processes. Metastability – or the process-like nature of entities that are taken for inert objects – underlies individuals and signifies that we incorrectly conceptualize processes as fixed states. The dynamic understanding of components of a system and the system itself as processes implies that the system strongly depends on its history. Path dependence implies that the momentum of the system is regulated by the precise – already dynamic – conditions where the first “step” was made. The intensity of this dependence is evident in the fact that due to the non-linear amplification even the most insignificant feature in the past may have a global and drastic effect after a time (see next paragraph). Hence, in order to understand the system, its past is as relevant as is its present situation (Dobzhansky 1973:125 [71], Yates 1987:414 [72], Werndl 2009:197 [73], Prigogine 2009:155 [67], Schneider & Sagan 2009:151-152 [66], Hooker 2011b:20-21, 33 [64], 2011c:867 [74], Bickhard 2011:95 [69], Høffkirchner & Schafanek 2011:188-189 [75], Cilliers et al. 2013:2 [34]).

2.1.4. Complexity Means Non-Linearity, Sensitivity to Initial Conditions, Exponential Amplification and Chaos

The concept of non-linearity has already been mentioned on various occasions. Non-linearity\(^8\) is a property of complex systems that, even more than cardinality, renders them uncontrollable, both synchronically and historically. A non-linear system does not satisfy the superposition principle: its functioning cannot be described by equations of the first degree and its outputs are not directly proportional to the inputs so that a microscopic disturbance is typically amplified in an exponential manner. Synchronously, the linear increase in the quantity of components causes that the amount of configurations among them expands exponentially and becomes unmanageable. Historically, the insignificant behavior of a single piece of the system may trigger a dramatic macroscopic fluctuation after a time. The historical non-linearity makes complex systems highly sensitive to initial conditions, which, in turn, emphasizes their infiniteness and uncontrollability. The sensitivity is understood as an exponential divergence of processes issuing from neighboring initial states, i.e. states that are finitely identical or identical within a margin of error. Because of this sensitivity, the behavior of complex systems is chaotic – it is unpredictable although laws governing such organisms are, in principle, deterministic. The margin of error or rounding assumed in any approximation (due to the fact that realistic infinite series must be made finite) will, after a time, exponentially inflate the previously controlled inaccuracy, rendering any exact prediction invalid (Yates 1987:412-416 [72], Gleick 1987 [76], Smith 1988 [77], Strogatz 1994 [78], Alligood, Sauer & York 1997 [79], Auyang 1998:11-14 [41], Elaydi 1999:117 [80], Wagensberg 2007:56-57 [63], Prigogine 2009:222-223, 324 [67], Schneider & Sagan 2009:45, 115, 319, 350, 363-369, 377-379 [66], Werndl 2009:203-204 [73], Hooker 2011b:21, 25-26 [64], Bishop 2011:105-111 [65], Cilliers et al. 2013:2 [34]).

2.1.5. Complexity Means Emergence, Non-Additivity, Non-Modularity, Irreducibility and Organizational Intricacy

Another phenomenon that derives from the non-linearity of complex systems is emergence or the capacity of developing emergent properties. Emergent traits are characteristics that fail to be qualitatively comparable and analogous to the properties present in constituents or that are not directly derivable from lower-level entities. Inversely, systems that are emergent are non-resultant, non-additive and non-modularized: they cannot be explained by their microanalysis into independent parts because they are not mere superposed computations of their isolated components. Emergence emphasizes the existence of multiple echelons in a system (each one with their own properties, processes, terminology and behaviors) and the interplay between them. Accordingly, complex systems are irreducible – it is impossible to divide the system into subsystems without the important loss of information. As certain important features are recognizable only from the whole system’s perspective, any modularization will trigger a damage of information. In other words, since the behavior of the components depends on the emergent properties of

---

\(^8\)The term ‘non-linearity’ employed in complex-system theory refers to an entirely distinct phenomenon than the notion of non-linearity used by Duhl (2011 [7]).
the whole – and important characteristics of a lower level are dictated by a higher level – the system cannot be deconstructed into isolated individual portions, where the behaviour of a constituent appears as independent from the rest. Rather than being compounded of modules, complex systems are self-organizing organisms in which all the components are embedded and to which they all contribute. The organizational depth of complex systems is itself highly sophisticated: multi-dimensional, multi-level, multi-phasic with intra and inter-level relations and with top-down causation in addition to a down-top one. In this global non-modularizable coherence, a mechanic modular view breaks down and an organic one comes instead (Crutchfield 1994 [81], Casti 1995 [82], Mihata 1997:31[83], Auyang 1998:178-179, 342-343, 2000:170 [41], Schlindwein & Ison 2007:237 [40], Prigogine 2009:177 [67], Hooker 2011b:21-22, 28-29, 40, 50 [64], Bishop 2011:126, 128 [65], Cilliers et al. 2013:2-3 [34]).

2.1.6. Complexity Means Infiniteness

The above discussion shows that complex systems are extremely intricate, being persistently infinite: a) as boundaries are artificial inventions and the system is always a subsystem of a higher organization, the constituents and their types or varieties – should the system be complete – are infinite; b) as everything is connected to everything else, the amount of relations existing among the constituents is infinite; c) the phenomenon of non-linearity triggers an infinite amount of the system’s configurations even if the number of constituents is restricted to a finite one; d) due to the lack of boundaries and fluidity of individuals, the elaborateness of the system into organizational levels – from the most microscopic to the most macroscopic – is also infinite; e) to be complete, the series representing the system’s state or states of its components should be infinite; f) the system’s static or synchronic infiniteness is further complicated by its dependence upon history as all the previous infinite states somehow contribute to the present situation, sometimes in an exponential manner.

If two real-world systems are complex, offering the properties discussed previously in this section, the comparison of their respective complexities typically includes the comparison of their cardinalities. The cardinality, itself, may involve the number of elements or the size of a set. The elements or sets can refer to constituents, taxonomical types, relations, configurations, descriptive empirical series, historical states, organizational depth etc. However, as mentioned above, any realistic complex system is infinitely complex – the cardinality of their constituents, taxonomical types, relations, configurations, descriptive empirical series and historical states upon which they depend is infinite as it is so their organizational depth. Hence, the information included in real-world complex systems can be represented as an infinite set. In mathematics, under the Dedekind definition of infinity, the whole has the same size as its parts if the part is already infinite. For example, the union (i.e. summation) of an infinite set with a finite or infinite set is infinite; the power-set of an infinite set (i.e. the set of all subsets of the set S, the empty set and the set S itself) is also infinite; and any superset of an infinite set (i.e. a set that contains an infinite set) is likewise infinite. Accordingly, the determination which one of two real-world complex systems is more complex is futile. It is pointless to determine which of two infinitely complex systems is more complex as both are so in an infinite manner. What can be finitely quantified, measured and estimated as more – or, on the contrary, less – cardinal, intricate and complex are not realistic complex systems themselves but their models.

2.2. Models of Real-World Complexity

The properties of realistic complex-systems outlined above have some important bearings on the scientific treatment of such organizations in models, distinguishing them from other simpler structures and their representations. In other words, apart from being characterized in terms of their traits such as those described in the previous section, complex systems can also be defined in terms of their models. To be exact, “complexity is the property of real-world systems that is manifest in the inability of any one formalism being adequate to capture all its properties” (Mikulecky 2007 [84]). While simple systems can be fully described by their models, a complex system never can (Cilliers et al. 2013:3-4 [34]).

2.2.1. Complexity Means Incompleteness and Provisionality of Models

As complex systems are incompressible and irreducible, their models – irrespective of the grade of sophistication – are always incomplete. Since the information included in a complex system is infinite and the system cannot be sliced up into subsystems “without suffering an irretrievable loss of the very information that makes these systems a system” (Casti 1995 [82]), a complex system can never be entirely compressed by models, which are by definition finite, isolated and partial. Hence, models never contain all the information that exists in the realistic system. If a model of a complex system were complete and able to represent all the possible behaviours of that system, the model in question would have to be at least as complex as the system it represents. Inversely, since all models of the real world inevitably simplify, only a limited portion of information

---

9 It is important to note that although the cardinality is not a unique or even the most important property of a complex system, the comparison of complexities of two, or more, complex systems can only be made in a numerical manner. Hence, it will necessarily involve a comparison of cardinality of a certain type.

10 In general terms, all models of realistic phenomena are theoretical hypotheses that (because of approximations, idealizations and rounding) drastically simplify reality. However, in comprehending reality, no other solution is available. In fact, science is possible only because of approximations, idealizations and rounding (Rosen 1985 [85], 1991 [86], Futuyma 1998:128 [87], Auyang 1998:69-70 [41], Cilliers 2007:82-83, 88 [88], Diéguez Lucena 2010:66, 75 [89]).
corresponding to the real world will be present. A part of the data will always be left outside the arbitrary limits of description, imposed by the model itself. As explained, this portion will, in fact, be infinite and, due to the sensitivity to minimal fluctuations in initial conditions, will inevitably affect the system’s running after a time. Furthermore, a model of any complex system must partially isolate or frame the system it aspires to represent from the environment. However, given that there is only one complete complex system, viz. the entire universe, and that no absolute boundaries are present in reality, the very isolation of the system and specification of its boundaries – separating it from the remaining portions of the universe – will render the model incomplete. In order to model a complex system accurately, a scientist should model everything, life and reality included. Consequently, both incompressibility and irreducibility of real-world complex systems and the incompleteness of their models jointly imply that representations of complex systems are per se provisional. Any model is inherently tentative and fragmentary – it can always be expanded or comprised as far as its sophistication is concerned and developed within a different framework. Our representation and understanding of a complex system always change as the framework is revised (Richardson, Cilliers & Lissack 2007:26-28 [70], Schindwein & Ison 2007:237 [40], Allen 2001 [90], Allen et al. 2010 [91], Cilliers et al. 2013:3 [34]).

2.2.2. Complexity Means Plurality of Models

By recurring to approximations in delimiting its content and limits of representation, any model of a complex system inevitably falsifies the real picture of affairs. The relation type between the models and the states of a realistic target complex system is many-to-many: there are an infinite number of states of the target system which can be mapped into the same state in the model (for instance, the series of a model is finite and corresponds to an infinite number of similar series in the target system that, however, diverge after the end point of a chosen approximation) and an infinite number of models can map a state of the target system (for example, each model employs a different cardinality of the series representing a given state of the target system). Accordingly, scopes and boundaries of models can be multiple and diverse. There is no unique model – no perspective can represent all the properties of a complex system. Hence, the study of complex systems necessitates a number of perspectives – an epistemological principle in analysis of such compositions is the exploration of perspectives.

As the determination of the boundaries of a model and the limits of its precision (and, thus, designation of the rest as a non-relevant noise and/or inactive environment) rather than being dictated by the system itself, is a pragmatic question of cutting up the ‘system’ and ‘environment’ that depends on convenience and suitability for a given analysis (so-called framing), the observer’s position, scientific needs (description purposes) and the model’s constructor or human actor should all be incorporated into the representation (or at least acknowledged in it). Epistemologically, it is impossible to separate the observer from the world: to a degree, reality results from the decisions made by the observer, just like the explanations concerning the world depend on the explainer. In complexity thinking, a basic assumption is that subject and object cannot be radically separate: “complexity resides as much in the eye of the beholder as it does in the structure and behavior of a system itself […] and requires the reintegration of the observer in his observation” (Schindwein & Ison 2007:236 [40]). Since complexity results from the position and perception of the observer, a certain portion of complexity is, in fact, subjective and hinges on how the explainer looks at and analyzes the system. Complexity is something that exists and something we construct in models. Hence, the degree of the complexity offered by a model will depend not only on the system represented by that model but also by this system’s interaction with another system, observers or explainers. Complex-system thinking demonstrates that there is no one right answer when approaching a complex system – as a complete representation of a complex-system is impossible, various manners of representation are conceivable. Each model answers only the questions relevant to itself – it does not respond to all the questions (Senge 1990:281 [92], Casti 1995:269-270 [82], Cilliers 1998:4 [61], Smith 1998:127 [77], Richardson, Mathieson & Cilliers 2000 [68], Richardson, Cilliers & Lissack 2007:30-31 [70], Cilliers 2007:82-83, 88 [88], Prigogine 2009:222-223 [67], Schindwein & Ison 2007:233-238 [40], Bickhard 2011:101 [69], Hooker 2011b:43, 84 [64], Bishop 2011:112, 115, 117, 121-123[65], Cilliers et al. 2013:3 [34]).

2.2.3. Complexity Means a Continuum of Models of Increasing Complexity

Just like any model of realistic phenomena, in order to be scientific, models of complex systems necessarily approximate the system under analysis and represent it in partially ideal, scientific, terms. Only by making complex simpler, complexity can be controlled becoming knowledgeable for us. There is no a rule of thumb for developing models of complex systems and delimiting their (i.e. the models’) minimal complexity. Generally speaking, the less reductionist and simplistic a model is and/or the more accurately it preserves typical properties of complex systems outlined above – being still treatable or transparent enough to be comprehended – the “better” it is. There are three main ways of approaching complex systems in scientific representations: by designing many-body models (a large number of constituents of a few types are connected to each other by a few types of relations – it is non-linearity that renders these models complex); organic models (highly specialized constituents of a variety of types are strongly coupled and integrated in the whole) and cybernetic models (these representations combine many-body and organic models). Nowadays, many-body
theories are the most advanced, probably, because they generalize the most (for detail, see Auyang 1998 [41], Hooker 2011b [64] and 2011c:864 [74]).

As a main rule, a conceptual framework of complex systems – and, in particular, of many-body models – necessitates, at least, two somehow related scales: a composite macro-scale or macro-explanation and a variety of situated micro-scales or micro-explanations in which individuals can be accommodated and by means of which they can be connected. By representing macro- and micro-levels and the relation existing among them, the framework appears as both holistic and atomistic. Macro-explanations solve for the behavior and dynamics of the system as an all-inclusive rounded individual while micro-explanations couple the dynamics of the whole to the dynamics of underlying constituents and their connections. The relation between the levels is not only bottom-top but also top-bottom. Thus, the explanation of complex systems in their models “involves integrated holistic processes that also top-bottom. Thus, the explanation of complex systems – and, in particular, of many-body models – range from more resultant, isolated, coarse-grained (less precise), endogenous and dynamic to more emergent, open, relational, fine-grained (more precise), endogenous and dynamic. In all such cases, the exact shape of a representation is dictated by the aimed statistic treatment, generalizations to be discovered and the required precision in controlling causal factors in the system. By doing so, each model unveils different macro-truths and their relation to micro-states, and solves for distinct fragments of the target real-world system, distinguishing diverse patterns and dissimilar facets of its organizational consistency (Auyang 1998:11, 15, 67-70, 342-344 [41], Prigogine 2009:177 [67], Diéguez Lucena 2010:66, 75 [89], Hooker 2011a [42], 2011b [64] and Cilliers et al. 2013 [34]).

Apart from this, as already mentioned, the more properties prototypical to a realistic complex system a given model preserves, the “better” it is. In this manner, models range from more resultant, isolated, coarse-grained (less precise), settled to equilibrium, with fixed boundaries or external world relegated to exogenous parameters, and with endogenous variables externalized or regarded as given and fixed, to more emergent, open, relational, fine-grained (more precise), endogenous and dynamic. In all such cases, the exact shape of a representation is dictated by the aimed statistic treatment, generalizations to be discovered and the required precision in controlling causal factors in the system. By doing so, each model unveils different macro-truths and their relation to micro-states, and solves for distinct fragments of the target real-world system, distinguishing diverse patterns and dissimilar facets of its organizational consistency (Auyang 1998:11, 15, 67-70, 342-344 [41], Prigogine 2009:177 [67], Diéguez Lucena 2010:66, 75 [89], Hooker 2011a [42], 2011b [64] and Cilliers et al. 2013 [34]).

2.2.4. Complexity Means Multiple Manners of Complexity Quantification in models

The two main manners in complexity quantification mentioned above – i.e. Kolmogorov and Gell-Mann complexity – can be viewed as only two types among many other possible ways of measuring a system’s complexity. In scientific literature, two main modes in quantifying the complexity of complex systems are distinguished. One corresponds to the epistemological side of complexity, while the other to its ontology. The quantification of the epistemological mode – or formulaic complexity – consists of determining the length of the series that could give an adequate description of the system under analysis (descriptive complexity),\(^\text{11}\) the length of the series that could simulate the system (generative complexity) or amount of time, effort and energy in resolving a problem (computational complexity). The quantification of ontological modes includes, itself, three subtypes: compositional complexity (the number of components [constitutional complexity] and the number of types of components [taxonomical complexity]); structural complexity (the number of possible ways of arranging components or the number of relations and configurations [organizational complexity] and elaborateness of organizational levels or subordination/inclusion relationships [hierarchical complexity]); and functional complexity (variety of functions or operations the system can perform [operational complexity] and intricacy of laws governing the system [nomic complexity]).\(^\text{12}\) These three modes reflect three perspectives: formulaic complexity measures how we comprehend a complex system; ontological complexity measures the structure of the system and its organization; and functional complexity measures the behaviour of the system. Given the properties of real-world complex systems, the three facets and measurement are closely related and equally relevant. In concrete cases, one typically selects a certain perspective and does not approach the complexity of a system in its entirety. This, however, does not signify that the measurements that have been ignored reveal less truth about the system at issue (Cejnárová 2005:57-59 [45]; see also Casti 1995 [82], Auyang 1998:344 [41], Schindlwein & Ison 2007:232-233 [40], Wagensburg 2007:60 [63], Schneider & Sagan 2009:55 [66], Hooker 2011b:30, 37 [64], Sinnemäki 2011 [57]).

In all the types of measurement specified above, more complexity implies a higher cardinality of a chosen factor. One should note that the increase of the cardinality of one factor can correspond to the decrease of the cardinality of another. For example, randomness implies a high cardinality of possible configurations and a low cardinality of rules. On the other hand, order can mean a higher cardinality of rules and, thus, a lower cardinality of configurations as certain combinations are constrained. Additionally, although each mode can be represented numerically as a finite series that quantifies the number of elements, each cardinal chain is, in fact, different as it represents entirely distinct phenomena. The commensurability of these cardinalities is possible only within a model – at a meta-level – and it is entirely model dependant. Thus, the outcome of the comparison of these cardinal series or their total summation is strongly conditioned by the model and has little to do with the system under analysis itself. This, in turn, demonstrates that the measuring of total complexity is, to a great degree, an elusive task. As a result, rather than comparing systems or models in their totality, scientists evaluate them in respect

\(^{11}\) Under a narrow view, Kolmogorov complexity would correspond to this type.

\(^{12}\) Under a narrow definition, Gell-Mann complexity could be understood as nomic complexity.
to a given – more specific – mode.

To conclude, as all models of complex systems are incomplete and provisional, no model can claim to be exhaustive. A variety of models is admissible and, in fact, necessary to treat a given complex system: in each one of them, different properties typical to realistic complex systems will be acknowledged and different modes of complexity measured. Hence, the comparison of models and the measurement of their complexity make sense only if the fragmentary position of a proposed model in a more global analysis is fully recognized.

3. Language as a Complex System

As already explained, complexity underlies all real-world systems, not only physical, biological and chemical ones, but also – and, in fact, especially – those related to human activity, be they economic, social or cultural (Cilliers et al. 2013 [34]). Accordingly, language, a phenomenon where physical, biological and socio-cultural factors coexist and intervene is viewed as an exemplary real-world complex system (Pinker 1994 [47], Li & Vitanyi 1995 [48], Cilicover & Nowak 2003 [50], Ceninarova 2005 [45], Solé 2010 [46], Dahl 2011 [7], Andrason 2012 [55], Massip-Bonet 2013 [51], Munné 2013 [52], Mufwene 2013 [53]). By doing so, it is assumed that it will offer all the characteristics typical of realistic complex bodies, being open, situated with fluid boundaries and unstable individuals; “infinitively” cardinal in respect to its components, relations and configurations; dynamic, metastable and path dependent; non-linear, sensitive to initial conditions, exponentially amplifiable and in regions chaotic; emergent, non-additive, non-modularizable and organizationally intricate. As far as the modelling of language is concerned, this will typically be incompressible, model-specific and model-plural, thus leading to a profoundly model-dependent representation of complexity and its quantification. Again, as is the case with other real-world complex organizations, a high cardinality (especially, high cardinality of components and relations or rules) constitutes only one of the features that make language an exemplary complex system.

3.1. Language Complexity

Language is a prototypical open system that constantly exchanges material and energy with the environment. It influences our reality and perception (for instance, through categorization), being, at the same time, affected by the external world (for example, though the creation of new lexemes necessary to represent new objects). The openness of a linguistic organization – or anyone of its sub-parts – is also evident in a constant grammatical renewal of languages. Words like energy propel the formation of novel, usually periphrastic, grammatical constructions so that the grammatical inventory of forms (‘core grammar’) is constantly renovated by using lexical and syntactic material. As older categories become obsolete and disappear, new formations are continuously derived. By interacting with its milieu, language is always a situated phenomenon: it is invariably embedded in a culture, social organization and higher bodies on which it depends and to which it simultaneously contributes. Accordingly, any fragment of a language is embedded in a larger system so that no components can be viewed as isolated with clear boundaries and fully externalized exogenous settings. In general, boundaries at which language comes into contact with the physical world, biology of human mind or socio-cultural institutions, are vital parts of the linguistic system itself, so that, under certain approximations, physics, biology, sociology and culture can be considered important spheres of languages. As everything is open, situated and embedded with no non-arbitrary boundaries, the individuality of components of a language is fluid and unstable. An entity that, at a certain level, appears as an individual may, at more macroscopic levels, be a component of another individual. Inversely, if envisaged from a more microscopic perspective, an individual can constitute a system of closely collaborating elementary individuals. The meaning of a grammatical category such as a verbal tense – when approached from the perspective of cognitive linguistics – constitutes a typical example of such a behavior. At a macro-level, a gram – treated in its totality as an individual – offers a certain global meaning that interacts with other macroscopic grammatical objects in the system. At this moment, the meaning is the information which is attached to the form as such. However, when analyzed at a lower level of description, the formation equals a fluctuating mass of more atomic cases, senses, each one of them with its particular individuality and network of relation with other lower-level individuals (e.g. words appearing in the immediate vicinity that contribute to the context and the specific sense of this formation). For example, in this manner, a Biblical Hebrew macroscopic individual gram qatal is a composition of microscopic individual qatal forms appearing on concrete occasions (Massip-Bonet 2013 [51], Munné 2013 [52], Mufwene 2013 [53], Andrason 2012 [55]; on the openness of language, see also Jenner, van Peursen & Talstra 2006 [54], van Uden 2007:148-150 [93]).

The cardinality of the components of a language is extreme. Only the number of words is enormous and due to the derivation or composition de facto infinite. If other purely linguistic components are accounted for (sounds and their formants or morphemes such as inflectional endings, derivational suffixes, etc.), the amount of constituents of a language is absolutely untreatable. In fact, due to the instability and fluidity of individuals, it can always be increased to infinitum. The immensity of the components and factors that concern language becomes even more evident if one includes physical, biological or sociological parameters and variables. An extreme cardinality of components, even if approximated to purely linguistic ones, renders the number of relations among constituents absolutely overwhelming. As everything interacts with
everything else, language establishes a gigantic set-up of
connections and the amount of combinations between
components that belong to various levels in clauses, phrases
and sentences is absolutely unmeasurable. These interactions
are infinite due to the fact that all the components of a
written text, oral discourse or pragmatic situation constantly
interact and influence one another in a circular, feedback
loop manner: the environment and the entity which is
embedded by it are given simultaneously and cannot be
separated. To put it simply, the context influences the entity,
being at the same time already influenced by it. There is no
exact starting point of this mutual interrelation as neither the
individual nor the context comes first: their relation is
absolutely intricate. The infinite cardinality of language
may also be seen at another plane. As the number of features,
parameters and variables is infinite (language being
embedded in the realistic world), the series with which one
would wish to encapsulate the total information provided,
even, by a microscopic fragment of a language is
uncontrollable: it can always be expanded ad infinitum. For
instance, a series that describes a single sense offered by a
gram in a precise time and place can be made infinite as
more factors – from coarse-grained and purely linguistic to
fine-grained and pragmatic, up to the entire universe – could
be incorporated (Čejarova 2005 [45], Solé 2010:191-217
[46], Andrason 2012 [55], Munné 2013 [52]).

Language is an inherently dynamic phenomenon
and uncontrollable: it can always be expanded ad
infinitum. For instance, a series that describes a single sense offered by a
gram in a precise time and place can be made infinite as
more factors – from coarse-grained and purely linguistic to

\[ X(n) = f(Y(n-1)) \]
and
\[ Y(n-1) = g(X(n-2)) \]

Thus, variable \( X \) at a moment \( n \) depends – by the way given by the function \( f \) –
on variable \( Y \) but calculated in the moment \( n-1 \). On the other hand, at the same
moment \( n-1 \), variable \( Y \) depends on \( X \) but counted at the moment \( n-2 \). In
consequence, we obtain the following equation:

\[ X(n) = f(Y(n-1)) \]

This leads to the following differential equations where \( n \) is a real number:

\[ \frac{dX}{dn} = f(Y(n)) \]
and
\[ \frac{dY}{dn} = g(X(n)) \]

Thus, by differentiation:

\[ \frac{d^2X}{dn^2} = \left( \frac{df}{dY} \right) \frac{dY}{dn} \]

And next, by substitution

\[ \frac{d^2X}{dn^2} = \left( \frac{df}{dY} \right) g(X(n)) \]

where the expression

\[ \frac{df}{dX} = \frac{df}{dY} \]

is a new function of \( Y \), noted as a function from the derivative \( X \):

\[ \Delta Y = k \cdot f \cdot g(X(n)) \]

The final result is a differential second-order non-linear equation.
Language fails to be a simple aggregate of its atomic material. On the contrary, new emergent and non-resultant properties appear as constituents organize so that the system, as a whole, develops novel characteristics which did not exist at the constituents’ level. For instance, cognitive linguistics teaches us that the meaning of a form as such – i.e., the total meaning of an entity, be it grammatical (e.g. a tense) or lexical (e.g. a word) – is much more than a mere summation of concrete microscopic instances where this item appears. It has novel properties non-existing at a lower level. To be precise, at the macroscopic plane, where the form is analyzed as a holistic phenomenon, the vector of direction or change becomes an integral feature of the semantic representation: atomic empirical instances belonging to the micro-level generate a dynamic structure of higher rank, a gram viewed as a developing phenomenon where time and evolution are central parameters. In this manner, a set-theoretic union of microscopic senses available on concrete occasions (the polysemy of a gram) comes to make reference to the evolutive capacity of grammatical forms and of the language in general. This novel property can clearly be recognized in semantic maps organized along the grammaticalization paths which are commonly used to encapsulate and define the meaning of grams viewed macroscopically. In these representations, the time-dependency or vectored orientation is a new emergent characteristic of a formation, unperceivable at the microscopic level where the description of atomic cases is conducted. As macro-levels are not additive conglomerates of micro-properties, the whole is not directly reducible and merely modularizable into unrelated parts. The existence of emergent properties emphasizes the relevance of organizational depth and its intricacy. Various levels exist embedded in one another and influence one another: each one with its specific emergent properties that differentiate it from the lower and higher planes (Culicover & Nowak 2003 [50], Croft 2003:288 [97], Cejnarova 2005 [45], Massip-Bonet 2013 [51], Munné 2013 [52], Mufwene 2013 [53], and Andason 2012 [55]).

To sum up, language is a prototypical realistic complex system and offers all the properties typical of complex systems. Although cardinality is only one of these characteristics, the quantification of complexity can only be effectuated in a numerical manner. When quantifying the cardinality of language, it is evident that language is complex in an infinite manner, just like any real-world complex organization is. To be exact, as in real-world complex systems, the cardinality of language constituents, relations, configurations, empirical series that describe it, historical states upon it depends or organizational hierarchical structures are all infinite. Hence, the information included in a language – as in any real-world complex system – equals an infinite set. Since the whole has the same size as its parts if the part is already infinite, the issue of which one of two languages – viewed as real-world complex systems – is more complex seems to be vain, because both of them are complex in an infinite manner. Languages viewed as realistic systems resist any finite measurement and estimation in the way that one would be more complex than another – all languages regarded as real-world complex-systems are infinitely complex. What can successfully be finitely quantified and compared as more cardinal or, on the contrary, less cardinal, intricate and complex are not realistic languages but, again, their scientific models.

### 3.2. Models of Language Complexity

Although complexity of language can be quantified in models, all models of linguistic complexity – irrespectively of their sophistication – inevitably face exactly the same problems which accompany representations developed for other realistic complex organizations: they are incomplete, provisional, pluralistic, characterized by distinct modeling intricacy and by different manners of complexity measurement.

Language viewed as a realistic complex system is incompressible and irreducible, which means that its model will always be incomplete and fragmentary. As has been explained previously, if one wishes to model a complex system accurately, he or she would have to model everything, including life and reality. Given that, to be complete, a series describing a linguistic phenomenon (or the entire system) would have to be infinite, and that our descriptions – if they are to be manageable – must be finite, it is necessary to establish limits of precision and, hence, idealize by means of rounding, introducing boundaries and isolating the organization to be represented. Under such an approximation, the infinite amount of data is neglected: any series and thus the entire model are made uncertain. Since all the linguistic models drastically simplify and since this simplification typically depends on utilitarian factors and on the perspective adopted by the researcher, all the models are provisional. Given that the determination of the limits of the model and the extent of its precision, rather than being dictated by the system itself, is a pragmatic question of framing, the observer’s position and model constructor’s objectives must somehow be acknowledged, if not explicitly incorporated into a proposed representation. This also means that there is no unique model – no perspective can represent all the properties of a complex system. Hence, an infinite number of models can map a state of the target system (each model employs a different cardinality of the series representing a given state of the target system). Accordingly, scopes and boundaries of models can be multiple and diverse – the representation of a language is inherently pluralistic.

Paralleling the quantification of complexity in other non-linguistic models, the linguistic complexity can be equalled to the epistemological formulaic cardinality of the series (be they descriptive, generative or computational); to the ontological cardinality concerning the number of components (constitutional), kind of components (taxonomical), number of relations and/or combinations (organizational) and number of levels (hierarchical); or to the functional cardinality related to the number of operations performable by the system (operational) and the
number of laws regulating the system (nomic). The selection of the measured properties, their precision and assumed error, the determination of a numerical correspondence between the digits symbolizing different properties and series, and the position of the observer, all, affect the numerical outcome of the quantification of the overall complexity and, hence, its comparison between languages, be it typological (synchronic) or historical (diachronic). There is no objective procedure that could a priori establish, once and for all, how the global complexity should be represented and measured as already the above list of different types of complexity corresponds to an artificial – external to the world itself – categorization imposed by the model and observer. Other fragmentations of the overall complexity are likewise possible.

As a result, the types of complexity one selects to be measured, the number of categories within each type, the precision of values that specify each category, and the manner of relating numbers of one category and/or type to another category and/or type will all decide how the complexity “test” will be performed by a chosen language. How should the number of words (e.g. the number of the most basic non-derived and non-compounded lexemes) be related to the number of certain morphological categories, such as grammatical cases? Is this relation 1:1 or 1000:1? How should the number of exceptions be related to the number of rules; 1:1 or in any other manner? The answer to these and similar questions will always depend on the explainer and his or her model, rendering both the overall complexity and its local measuring entirely model and theory laden. This stems not only from the limitations of any complexity model – which are finite/isolated models of infinite/situated phenomena with a strong degree of the explainer’s intervention – but also from the fact that the quantifications refer to singularities and objects that are not originally numbers but features whose sets are counted. What is quantified is how many $x$ can be identified for a given feature. However, how many features can be identified for a language or should be included in an analysis (i.e. the selection of properties, its precision and variety of kinds) and how should the relation between a feature and the size of its set be linked and, additionally, correlated with other features and their sizes is far from being strictly numerical and straightforward. On the contrary, it is always dictated by the explainer’s position and the model’s needs or objectives. As a result, any measurement will inevitably be model-driven and, to an extent, be arbitrary.

Since the representation of the complexity of a language and its measurement always depend on the researcher’s necessities and model’s scope, the outcome of comparison between languages will likewise be conditioned by the properties of the model within which it has been elaborated. Thus, different perspectives and different approaches will give distinct outcomes of complexity measurement. This relativity should not be regarded as a weakness typical of and limited to linguistics – it is exemplary to any modelling of real-world complex systems by which such organizations are distinguished from non-complex and non-realistic systems (Cilliers et al.2013 [34]; cf. Munné 2013 [52] and Mufwene 2013 [53]).

4. Conclusion

Complex-system theory demonstrates that the worries expressed by Deutscher (2009:248-250) [39] and Dahl (2011) [7] who question whether it is ever possible to define the overall complexity of a language are more than appropriate. To be exact, Deutscher (ibid.) [39] proposes that, instead of being numerically quantified, global complexity should rather correspond to a vector of separate values referring to different and incomparable domains and features. As these values represent entirely different objects and phenomena, one cannot collapse the distinct complexity measures into a single, summed up, figure.

Complex-system theory demonstrates that any numerical comparison of the overall – or even local – complexity of languages (i.e. comparison of their cardinalities) is, if not futile, at least, deeply theory conditioned. First, complex-system theory teaches us that languages viewed as real-world systems are infinitively complex. The cardinality of their constituents, relations, configurations, empirical series, historical states, organizational hierarchical structures are all infinite. As any two linguistics systems are infinitively complex if approached as realistic phenomena, their comparison and measurement is, in a way, irrelevant: since they are infinite, they will always be equally complex. What can be compared in finite terms are models. However, since all the manners of quantification are invariably driven by the theory, its axioms, and utilitarian or pragmatic choices, the measurement and, hence, comparison between different languages, even though possible and sometimes practical, should be relativized and always viewed – just like the models themselves – as provisional, incomplete, fragmentary and pluralistic. Accordingly, no model or measurement may claim to establish the ultimate hierarchy of less or more complex languages. The complexity will constantly depend on the scientific frame of reference and theory within which it has been quantified. Thus, any synchronic and diachronic pyramid representing complexity of languages is nothing but relative.

The present paper has also demonstrated that complexity cannot be equated to cardinality. Although a comparison of complexities can probably be effectuated only in numerical or set-theoretical terms, the infinite or uncontrollable cardinality of a series or set is not a unique or even the most important property of complex system. Equally – or even more – relevant is openness, situatedness, lack of boundaries, individual instability, uncertainty, non-linearity, exponential sensitivity to initial conditions, dynamcility, metastability, path dependency, emergence, regional chaos, non-additivity, non-modularization,
irreducibility and organizational intricacy. All of this means that while the idea of complexity should definitely be incorporated into linguistics, this should be done with caution. Talking about complexity certainly makes sense but it must be performed carefully (cf. Dahl 2011 [7]). First, scholars should make a clear distinction between complexity of language and its model. The former is unquantifiable and infinite, while the latter, although quantifiable and possibly finite, is a fictionalized approximation, invariably tied to theoretical frameworks and observers. Second, models of linguistic complexity should not be traditional but should benefit from advances of the theory of complex systems and complexity science. If possible, they should go beyond an additive resultant representation of a language and modularization of an equilibrium state as has typically been done thus far. Instead, they should incorporate the concepts of dynamics, non-linearity, emergence and situatedness (see Muñoz 2013:215 [53]). And third, more advanced manners of the computation of complexity should be used, where measures of complexity developed within different frameworks (Kolmogorov-Chaitin or algorithmic information theory, classical information theory of Shannon and Weaver, Fisher information, logical depth, thermodynamical depth, computational mechanics, etc.) and distinct information magnitudes (Shannon entropy, Fisher information, disequilibrium, variance, etc.) are compared, giving a more accurate estimation of the complexity of a particular system (Angulo & Antolin 2008 [98], López-Ruiz, Mancini & Calbet [99] and Abe et al. 2004 [100]). This could partially free calculations of complexity from relativism. I am convinced that by doing so various phenomena – and the language itself – can be explained in a manner that would more closely approach our representation to reality, where language functions as an exemplary complex system.

References


[44] Cilliers (Ed.), Thinking Complexity (pp. 229-238). Mansfield: ISCE.


