C O P *i* L CAMBRIDGE OCCASIONAL PAPERS *in* LINGUISTICS
Volume **14(1)**, Article 1: pp. 1–33, 2022 | ISSN: 2050-5949

Emergence, Complexity and Developing Grammars: A reinterpretation from a Dynamical Systems perspective[*]

Núria Bosch
University of Cambridge

Abstract Any theory of language acquisition logically calls for a theory of the development and the epistemological foundations of individual grammars, yet the exact manner with which grammars emerge has been perennially debated (see Bavin 2009 for a review). Against this background, this work advocates for the potential of a Dynamical Systems take on grammar construction and generative grammar. I assume here Chomsky's (2005) Three Factors approach, as well as neo-emergentist approaches to language variation, which put forward a radically impoverished Universal Grammar (Biberauer's 2011, *et seq.*, Maximise Minimal Means model; cf. also Ramchand & Svenonius 2014, Wiltschko 2014, Wiltschko 2021). Taking as a point of departure a maximally poor set of starting conditions (Universal Grammar) and the assumption that there exists a third-factor principle that aims to maximise minimal means, I then show how Dynamical Systems Theory (DST) naturally complements these perspectives on learnability and offer one possible theoretical implementation of DST in this context. The suggested architecture attempts to relate acquisition, cognition and representation explicitly: symbolic dynamics and contextual emergence analyses of DST allow us to interrelate, both metaphorically and topologically, (i) acquisitional dynamics, (ii) conceptual spaces (*à la* Gärdenfors 2000, 2014) and (iii) the representational system being derived from these interactions. The acquisitional and theoretical consequences of the proposal are also discussed.

1 Introduction and Theoretical Background

One of the basic questions to be asked about language and language variation is the extent to which apparent principles of linguistic systems are unique to this cognitive system or whether similar arrangements can be observed in different cognitive areas in humans or other species. An even more fundamental task is to determine how much of language can be given a principled explanation based on well-known properties of natural and biological systems.

An important theoretical backdrop upon which this work rests is the idea that language, biology and mathematics share crucial common traits. Language is a

---

[*] The work presented here is a slightly modified version of my undergraduate dissertation (submitted in March 2022). I am immensely indebted to Theresa Biberauer for introducing me to complex systems and for her invaluable support and advice during this dissertation and my undergraduate studies. Her footprint is visible throughout this work. I also thank the audience at SyntaxLab (8 February) for stimulating feedback and discussion. All remaining errors are, of course, my own.

biological property of human beings. As far as we know, it is a species property, meaning it is identical across humans, pathologies aside (Hauser, Chomsky & Fitch 2002). It is a defining feature of human beings and, as a biological system, it deserves scientific treatment on par with the way we study the visual system and any other biological system. Taking an evolutionarily- and biologically-oriented perspective on language means posing questions about what language is, which (modular or non-modular) components it comprises, how it develops in children and adults and why it evolved in some way and not another. This has become known as the *biolinguistic enterprise* (Berwick & Chomsky 2016), whose essence could be captured with the following exhortation: let's approach language in the same way any other scientist would look at a biological system (Chomsky 2010). This paper seeks to make progress in this domain by reinterpreting grammar construction and language development through the lens of Dynamical Systems Theory (henceforth, DST), a branch of mathematics that describes the behaviour, development and change in complex systems.

In connecting biolinguistics, systems theory and acquisition together, my point of departure will be, on the one hand, Chomsky's (2005) Three Factors approach and, on the other, Biberauer's (2019) Maximise Minimal Means (MMM). Moving forward from the Principles & Parameters era of the 1980s and 1990s, the Minimalist Program (Chomsky 1993, 1995) reshapes the theoretical framework and its methodological and ontological apparatus to a considerable degree. Not only is the language faculty endowed with those operations and components that hold a substantial degree of 'virtual conceptual necessity' (the Strong Minimalist Thesis), but there is a strong commitment towards gradually shrinking the content of Universal Grammar (UG) in order to comply with theoretical parsimony and general evolvability requirements. Chomsky (2005) thus proposes three factors that enter into the growth of language systems. This is schematised as follows (adapted from Biberauer 2019: 47):

(1) (Maximally poor) UG + Primary Linguistic Data (PLD) + 3rd factors → Adult (steady-state) grammar

These third factors are argued by Chomsky to fall into various subtypes: (i) 'principles of data analysis that might be used in language acquisition and other domains' and (ii) 'principles of structural architecture and developmental constraints that enter into canalization, organic form, and action over a wide range, including principles of efficient computation, which would be expected to be of particular significance for computational systems such as language' (Chomsky 2005: 6).

In this context, Biberauer (2019) proposes an explicit third-factor, general-cognitive principle, Maximise Minimal Means (MMM), which unifies two independently proposed principles in the linguistic domain: Roberts's (2021) Input Generalisation and Roberts & Roussou's (2003) Feature Economy. In her approach, Factor 1 comprises a radically impoverished UG, with the formal operations Merge and Agree, and some notion of formal feature. Taking a step towards the general minimalist desideratum of diminishing the role of UG, formal features (FFs) and their substantive content are assumed to be emergent, with only a feature template of some kind or the notion of feature being encoded in UG.

Language is viewed, following Abler (1989), as a Humboldt-type system, a class of emergent complex systems that 'makes infinite use of finite means' (von Humboldt 1836: 70), and whose 'synthesis creates something that is not present per se in any of the associated constituents' (von Humboldt 1836: 67). In other words, Humboldt systems place emphasis on the emergent (rather than so-called 'resultant') products of complex systems, where 'the whole is more than the sum of its parts'. MMM thereby encapsulates a general cognitive principle, evident in language and acquisition, that makes maximal use of minimal means. The emphasis on grammars universally structured via generativist notions such as Merge, Agree and formal features (cf. Baker's 2008, Borer-Chomsky Conjecture) clearly sets the approach apart from traditional emergentist approaches such as those in the usage-based or Construction Grammar tradition, hence the name *neo-emergentism* (see, i.a., Ramchand & Svenonius 2014, Wiltschko 2014, Wiltschko 2021, Samuels 2009, for comparable approaches).

Throughout this work, I will be assuming a neo-emergentist approach to language acquisition and linguistic variation, with a maximally poor Universal Grammar. I will also be agreeing with Mielke (2008), Dresher (2009) and Biberauer that formal features (both phonological and syntacticosemantic) are themselves emergent, with only the *notion* of feature being encoded in Universal Grammar.[1] Given the above, I work, too, under the assumption that there exists some cognitive bias that attempts to maximise available resources. The cognitive reality and plausibility of MMM (summarised in Biberauer & Bosch 2021), along with emergent systems, will become important when discussing DST in the context of dynamical systems, cognition and language (section 2.2, section 3).

The overall aim here is to draw a general outline of a systems-theoretic acquisition model, which is consistent with and complements neo-emergentism and generativism, and to generate research avenues which might not otherwise have been instigated from a distinct methodological standpoint. DST provides a new perspective on the development and change in linguistic systems and it also operationalises the interaction of their component parts. This concept encourages a *Gestaltwechsel* – a change in perception – which fosters the search for new answers to burning questions about the interaction of the cognition-acquisition-language interface (Ratter 2012: 85).

The paper is organised as follows: in section 2, I review the main characteristics of Dynamical Systems Theory and the role of DST in cognitive science. This second section presents the basis of and foresees a potential application of DST to grammar construction, which is sketched in more detail in section 3. Section 4 outlines novel predictions of the model and presents some additional hypotheses in the domain of learning paths, formal features and linguistic representations. Section 5 concludes.

---

[1] As I will be discussing both syntactic and phonological features, I do not assume that a specific formal template is indispensable (e.g., [uF, iF], [Att:(val)]). Broadly, UG is simply assumed to restrict linguistic systems to those compressible in hierarchical and featural terms. Even if a feature template ends up being required in UG, though, it does not fundamentally alter the suggestions made in this work.

2 Dynamical Systems Theory and Cognitive Science

2.1 General overview

Dynamical Systems Theory (DST) is a sub-branch of mathematics and physics that describes the long-term behaviour of complex dynamical systems, usually via the use of differential equations or difference equations. In general, DST relies on a geometric approach to systems, by understanding the states and the evolution of a system geometrically, in terms of its position with respect to other states and features of the system's landscape.[2]

Loosely speaking, a dynamical system is a mathematical object that describes how the states of a system develop over time. Any system that changes through time can be understood as a dynamical system. More precisely, a dynamical system is formally defined as a triple $< T, X, \Phi >$, consisting of a set of times T, a state space X, and a transition rule Φ that specifies the evolution of a state with time. A system's phase or state space (X) is the set of all its possible states; each state corresponding to a unique point in the state space. For example, a graphical plot of a phase space for the pendulum equation is shown in Figure 1. The evolution rule Φ is a function whose domain is $X \times T$ and whose co-domain is X, namely $\Phi : X \times T \to X$. The function Φ takes two inputs ($\Phi = \Phi(x, t)$), where $x \in X$ is the initial state (at time $t = 0$, for instance) and $t \in T$ is a future time. Put simply, $\Phi(x, t)$ returns the state at time t when the initial state was x.

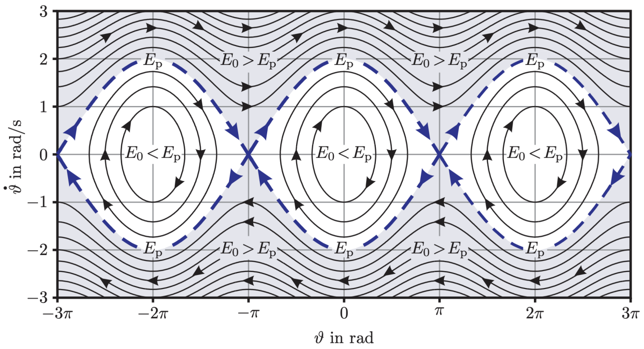


Figure 1 A phase portrait of the pendulum equation (Ochs 2011: 482).

An important goal in DST is to describe the fixed points, or steady states, of a dynamical system, which are the values of a variable that are invariant over time. Some of these points are attractive. These have come to be known as *attractors*: a set of states towards which the system tends to drift, regardless of the system's initial conditions. An attractor's *basin of attraction*, then, is the region of the phase space such that any point (any initial condition) in that region will be asymptotically

---

[2] The following paragraphs are not meant as a comprehensive overview of dynamical systems. For greater mathematical and conceptual detail, we refer the reader to Gottlieb (1992), Thelen & Smith (1996) or Glendinning (1994).

iterated into the attractor. As the treatment given here of attractors and dynamical systems will be largely non-mathematical, see Hawkins (2021) or Milnor (1985) for the mathematical definition of an attractor.

Within this geometric and topological phase space, new behaviour or structural reorganisations can emerge, not only as a result of the system's adaptation to the input and environment (cf. section 2.2.3 later), but also due to changes in one of its *control parameters*. Control parameters (akin to independent variables in conventional research) are a formalisation of the parameters to which the collective behaviour of a system is sensitive. They lead to phase shifts (system reorganisations) by threatening the stability of the current attractor. The notion of control parameter will become relevant in section 3.1.3.

Considering that DST will be applied over acquisition in section 2, we will only be concerned with a subset of possible dynamical systems. Our primary attention will be on *open*, *non-linear* and *complex* dynamical systems, i.e., dynamical systems with multiple interacting components that are in constant flux as a result of environmental fluctuations and whose behaviour may not be reducible to the behaviour of its individual components. To be more specific, given the nature of the explanandum here, we will be concerned with *complex adaptive systems*, or CASs, (a subset of non-linear dynamical systems), namely systems whose apparently complex behaviours emerge as a result of non-linear spatio-temporal interactions among component parts at distinct levels of organisation (Chan 2001).

2.2 Properties of complex adaptive systems

There are a number of properties central to CASs, which this work borrows, and therefore it is worth commenting briefly on them. These include:

2.2.1 Emergent, non-linear and self-organising

CASs share many of the properties with which Abler (1989) was concerned, in particular their emergence and self-organisation properties. In CASs, new forms or properties come into existence as a result of interactions among smaller or simpler forms that do not have these properties themselves. The crucial result is that systems are not reducible; process transformations cannot be reduced back to their original state. Therefore, emergence is effectively a violation of the superposition principle, where the net result of the sum of influences is not equal to the sum of their individual results. It implies that the systems are *non-linear*: the change in their output is not proportional to the change observed in the input.

Another important feature of a complex system is the idea of *self-organisation*. Systems organise and structure themselves without need for external direction, manipulation or control, leading to an increase in the structure or order in the system. Order and behaviour is thus not predetermined: pre-formationism is no longer a necessary requirement for developing systems (Tucker & Hirsh-Pasek 1993). The system creates its own form and order (i.e., behaviour and order is 'softly-assembled', rather than hard-wired), which will contain emergent features that are qualitatively distinct compared with earlier organisations (cf. the distinction in Gottlieb 1991 between probabilistic and predetermined epigenesis).

2.2.2 Sensitive to initial conditions

Non-linear systems are highly *sensitive to initial conditions* or *path dependent*. Systems can start in a nearly identical state and can develop in opposite directions as the system amplifies initially minor differences (Warren, Franklin & Streeter 1998). As a result of this sensitivity to initial conditions, a dynamic model is 'recursive', or 'iterative'; that is, it describes a procedure or function that transforms x_t into x_{t+1}, x_{t+1} into x_{t+2}, x_{t+2} into x_{t+3} and so on. The output of a preceding application of the transition rule Φ is the input to the next application.

A logical by-product of path dependence is that CASs are also *feedback sensitive*. The output of a process within the system is 'recycled' and becomes a new input for the system. The nature of this feedback can be positive or negative: negative feedback reverses the direction of change of a given variable, while positive feedback increases the rate of change of the variable in a given direction (Rickles, Hawe & Shiell 2007).

In complex systems, feedback occurs between levels of organisation. Open and self-organising systems are composed of three basic and inseparable dimensions of space-time (Fenzl 2018: 34):

- The **microscopic dimension** on the level of the individual **elements**.
- The **mesoscopic dimension** of the whole structure, limited by the **structural boundary**.
- The **macroscopic dimension** of the **field of interaction or relevant environment**, limited by the **system boundary**.

This leads to a bidirectional and highly interactionist view of systemic development: higher levels can 'back-react' onto subunits, causing them to generate new patterns, which back-reacts again, and so forth. This type of 'global to local' positive feedback is sometimes termed *coevolution*, a term from evolutionary biology that describes the ways in which organisms create their environment and are in turn shaped by that environment. An important point that emerges from these inter-dimension interactions is the concept of *structural homology*: later forms are built up of earlier forms. The reorganisations evident at each successive level are composed of structures that were present at earlier stages (Tucker & Hirsh-Pasek 1993).

2.2.3 Adaptive, critical and self-similar

If feedback loops involve a macroscopic dimension, as set out above, it stands to reason that CASs are *adaptive*: behaviour mutates and self-organises in response to a changing situation in the environment. An adaptive tension emerges from the energy differential between the system and its environment (Turner & Baker 2020). CASs are characterised by a high degree of adaptive capacity, allowing them to be resilient even in the event of perturbations. Systems are therefore *open* to the environment, meaning that external forces can shift the system into a new state or way of interacting (Perone & Simmering 2017). There is thus a certain dependence on continued interaction with the environment, as they require a constant flow of energy to sustain the organisation of the system (i.e., they are *out of thermodynamic equilibrium*).

However, adaptation to the environment does not, in fact, proceed in just any imaginable manner. These systems have an ability to avoid chaos by self-organising to a state roughly midway between globally static (unchanging, ordered) and chaotic (random, disordered) states. Systems that are simultaneously ordered and disordered are more adaptable and resilient, by adapting to the *edge of chaos*.

We will capitalise further on the 'edge-of-chaos' properties of CASs. The edge-of-chaos zone is analogous to the scientific phenomenon of the *phase transition* or *phase shift* – the point at which a system transitions from a state to another. Some theories of neuroscience have suggested that the human mind and cortical networks could operate at this edge-of-chaos zone, or what is sometimes termed the critical state (the Critical Brain Hypothesis or self-organised criticality). Although not totally free of controversies (Touboul & Destexhe 2010), the idea is that optimal information processing in non-linear systems is achieved close to phase transitions, a kind of 'Goldilocks zone' that optimises the transfer and processing of information while still maintaining stability.[3]

Genuinely complex structures and systems arise in this region on the edge of order and chaos, where they can take advantage of the possibility of the sudden fluctuations inherent in non-linear dynamics, while maintaining the order necessary for continuity (Waldrop 1993). In other words, systems operating at the edge of chaos are endowed with some sort of adaptive advantage, in being able to make the most of intermediate levels of disorder to optimise information processing capabilities. A particular type of attractor that arises in edge-of-chaos dynamics are so-called *strange attractors*, which provide an explicit link between edge-of-chaos and self-similarity. Fractal (self-similar) structures appear naturally in the phase space of non-linear systems, in such a way that the two are deeply related. Complex patterns recur across different scales, by repeating a simple process over and over in an ongoing feedback loop. While attractors can be seen as a way of representing the behaviour of a system in geometrical form, all strange attractors have in common the fact that their shape is geometrically a fractal. We will come back to the potential role of edge of chaos, fractals and strange attractors in section 3.1.2.

DST has been applied to a wide range of developing complex systems (from geology to cells or ant populations), which brings to the forefront the possibility that order and complexity in nature arise in an analogous way and can receive a unified metatheoretical treatment (i.e., that there is some 'ontological interconnection

---

[3] The observation that biological and neurological systems operate at this criticality level is widespread. We do not have space to review it here and we refer interested readers to literature on criticality and 'neuronal avalanches' (see Beggs & Timme 2012, Muñoz 2017, Beggs 2008, Chialvo 2010).

between mathematics and biology', Watumull, Hauser, Roberts & Hornstein 2014: 6). Before applying these notions to grammar construction, I turn briefly to literature on cognitive science that suggests cognitive systems can likewise be modelled with a systems-theoretic mindset.

2.3 The Dynamical Hypothesis in cognition

In human cognition and cognitive development, DST has been employed both as a conceptual framework and as a literal description of a dynamical system. Thelen & Smith's (1996) seminal work on motor and cognitive development supposed a revolutionary shift in thinking about child development. Instead of characterising what changes over development, there is new emphasis on the *how* of systemic change. Similarly, in the domain of cognitive science, the Humean possibility that cognitive systems can be characterised in dynamic terms began to gain momentum (Smolensky 1988, Schöner 2012). Although DST is often presented as an anti-representationalist alternative to computational (Hobbian) cognitive science (Haselager, Bongers & van Rooij 2003, van Gelder 1995), we will follow some DS theorists in assuming dynamicism is not inherently contradictory with representationalism (Shapiro 2013). In fact, it will become clearer in section 3.2 in which respects dynamics and representations are dissociable, but interestingly interconnected.

We will subscribe to the Dynamical Hypothesis by van Gelder (1998: 615), at least insofar as the behaviour and development of cognitive systems can be modelled as dynamical systems:

(2) **The Dynamical Hypothesis (DH)**:
Cognitive agents are dynamical systems.

The DH has two major components: the *nature hypothesis* (cognitive agents are dynamical systems) and the *knowledge hypothesis* (cognitive agents can be understood dynamically). The division of the DH into two parts is important: it is one thing for cognitive systems to be dynamical systems, but it is another for scientists to understand and model them as such. The claim behind the DH is that a DST toolkit should be part of any cognitive scientist's workshop. The fundamental move is to conceptualise systems geometrically, that is, in terms of distances, regions and paths in a space of possible phases and states, and in doing so, attempt to understand the structural properties of this *flow* and the range of (im)possible paths (van Gelder 1998: 621).

It is apparent that many of the processes cognitive scientists study (development, learning, the spread of activation in a network, patterns of cortical activity, motor behaviour, among many others) are intrinsically dynamic (Riley & Holden 2012). My proposal appends grammar construction to this list of dynamical processes and harnesses the tools and concepts supplied by DST to understand this cognitive phenomenon. As such, it takes a cognitively- and dynamically-oriented approach to grammar construction and generativism and, in the following sections, it argues that DST may help bridge the gap between cognition, acquisition and generative syntax.

3 A DST APPROACH TO GRAMMAR CONSTRUCTION AND NEO-EMERGENTISM

3.1 The model

In this section, I present a novel analysis of grammar construction drawing on the toolkit of notions and concepts introduced in section 2.

The main proposal I will be putting forward is largely a rehashing of van Gelder's Dynamical Hypothesis formulated over language acquisition. This is summarised in (3):

(3) **The Dynamical Hypothesis of Acquisition (DHA)**:
The cognitive process of grammar construction can be modelled as a complex dynamical system.

I will suggest that applying systems-theoretic notions as sketched in section 2 can provide (i) a natural, scientific and biolinguistic characterisation of well-known observations on language acquisition and (ii) fruitful avenues of research for generativists, with new consequences for how learning paths and learnability are approached. Although DST is not alien to functionalists or in the domain of L2 acquisition (Larsen-Freeman 1997, de Bot, Lowie & Verspoor 2007, Lund, Basso Fossali, Mazur-Palandre & Beldame forthcoming), to the best of my knowledge, a learnability model that is both generativist (in particular, neo-emergentist) and systems-theoretic is non-existent in the literature. The purpose, then, is to call attention to the similarities among complex non-linear systems in nature and language acquisition. While it remains to be determined whether the value of the analogy is metaphorical, sometimes 'you don't see something until you have the right metaphor to perceive it' (Bowers 1990: 132; cited in Larsen-Freeman 1997). Although the discussion is restricted to First Language (L1) acquisition, the approach could be extended to diachrony (as in Niyogi & Berwick 1997), L2 acquisition and other branches of acquisition such as heritage speakers and attrition (including complex multilingual situations). For reasons of space, we do not discuss these options here.

Let us, then, turn to applying and translating the properties of a complex adaptive system to grammar construction.

3.1.1 Emergence, self-organisation and the role of Universal Grammar

Dynamical systems start off their life in a relatively general and undifferentiated state (their first set of initial conditions) and then progressively diversify, self-organise and specialise. All that is really needed for complexity to arise in these systems is their sensitive dependence on initial conditions and a context against which the system adapts and develops (Tucker & Hirsh-Pasek 1993). Self-organisation and emergence, then, allows for a self-sufficient construction of complex systems that requires very little pre-ordained instruction. During this initial period of progressive specialisation, the system's 'learning potential' (its degrees of freedom) is much greater than it will be in later developmental periods. This notion is consistent with the traditional account of a developmental critical period, a window within which learning takes place with comparative ease. DS terminology provides a way to see how considerable complexity and structure can emerge from a relatively 'simple' system with initially substantial degrees of freedom. The system will undergo subsequent complexification and specialisation to the task domain and environment (the target language), with concomitant reduction in degrees of freedom. Applied to grammar construction, these self-organising and emergent properties of linguistic systems seem to allow us to begin the acquisition path with relatively little pre-specified structure (which is *not* to say that innate knowledge is superfluous; addressed now).

We might call UG the very first set of initial conditions of the system – it contains certain formal universal principles that constrain the shape of attested and attestable human languages. In this approach, UG has a 'navigation' role, in restricting the possible phase spaces of dynamical systems towards linguistic organisations that are driven in the by-now classical minimalist way (namely, syntactic systems characterised by application of the binary, set-theoretic operation Merge, and structured in featural terms; recall section 1). As UG remains one of the core system-structuring principles here, it very much retains part of the 'steering' role it had in the Principles & Parameters era, but crucially in a much more underspecified way. UG specifies the *shape* of grammars, but it does not provide *substance*. It does not itself dictate or hard-wire development (parameters and representations are emergent, so is the substantive content of formal features); instead, it determines the very initial state of the system, with sensitivity to initial conditions, the input and the attractors (see section 3.1.2) enacting the characteristic 'softly-assembled' development found in dynamical systems. To the extent that the system is also iterative (meaning that the output of a system at time t becomes the basis for $t+1$), one ensures the construction of a linguistic system that is (neo-)emergent and structurally homologous in this very sense: later forms are built up based on earlier ones. A maximally poor UG therefore remains well-founded upon dynamical systems methodology.

3.1.2 Attractors and life at the edge of chaos: Goldilocks, MMM and fractals

Much like cognitive systems, dynamical systems are not free of biases. Language acquisition systems are likewise biased in which facets of the input they pay particular attention to and which decisions are prioritised. Attractors in dynamical systems may thus be seen as corresponding to third-factor learning biases, such as Culbertson & Kirby's (2016) 'simplicity bias' or Biberauer's MMM. It is also perfectly possible that some attractors may be of a functional and processing kind. To name one relevant example: the well-known fact that rightward movement is difficult to parse compared to leftward movement could have important consequences when steering linguistic systems towards certain types of grammars and not others (e.g., see Ackema & Neeleman 2002, Abels & Neeleman 2012, Biberauer, Roberts & Sheehan 2014b, Kayne 2013, Uriagereka 2012, for distinct proposals for how this functional pressure relates to linearisation and syntactic derivations, both Kaynian and non-Kaynian).

These attractors may be of possibly various kinds: point attractors, limit-cycle attractors or strange attractors. The potential significance of distinct types of attractors was noted by Lass (1997), albeit in the domain of diachrony. It remains to be seen whether this typology of attractors is useful in diachronic syntax or in language acquisition, but it is arguably worth contemplating, as Lass does (see also Roberts 2021: 490-492). The aim here, though, is not to provide an exhaustive list of the attractors possibly involved in this dynamical system (in part owing to the fact that there is little agreement as to precisely which third-factor optimisation principles shape language acquisition). We also do not make any distinctions between point attractors or limit-cycle attractors. Instead, we propose to focus on a specific kind of attractor that would seemingly have rather direct parallels in language acquisition and general cognition: these are the strange attractors outlined in section 2.

By drawing an analogy between acquisition and complex non-linear systems found in nature, cognition and biology, it follows that grammar construction systems go through the same edge-of-chaos stages, as do many other living systems, and that they are frequently biased by strange attractors. Recapitulating, the creative growth of these systems takes place at an *intermediate* zone between chaos and order (Waldrop 1993). In other words, self-organisation needs to find a balance between no order and too much order, which fosters systemic development (De Wolf & Holvoet 2004). Assuming that chaos and order can in some respects be equated with complexity and simplicity, respectively, Edge of Chaos (Langton 1990) is a classic case of the Goldilocks concept of complexity. Borrowing a well-known example of this phenomenon, Langton formulates a value λ (with range 0 to 1), which can be used to help predict whether a given cellular automaton (CA) will fall in the ordered realm, in the chaotic realm, or near the boundary, on the edge of chaos. $\lambda = 0$ leads to a quiescent state, and the entire CA quickly becomes completely ordered, or 'frozen'. When $\lambda = 1$, no neighbourhood state moves a cell to the quiescent state, and the CA continues to fluctuate wildly and chaotically. Langton tunes the lambda value for the cellular automaton to the critical value ($\lambda = 0.5$) so that the CA is not too ordered nor too chaotic, but just right for exhibiting complex behaviour. Self-organised criticality on the edge of chaos is, thus, the zone where complexity and computation are possible, and where self-similar fractal structure is evident.

Self-organised criticality is found in physical, biological and other systems, prompting us to wonder what its materialisation might be in linguistic systems. Consider, in this connection, two proposed principles and biases in acquisition: the Goldilocks Principle (which suggests that infants direct attention towards input sequences that are neither too simple or nor too complex, but 'just right'; Kidd, Piantadosi & Aslin 2012, 2014) and Maximise Minimal Means. To the extent that brains and minds display self-organised criticality (section 2) and complex adaptive systems often present strange attractors, the expectation in a DST approach to grammar construction systems is that they will also operate in this manner. That is, that the system will work in such a way that it tackles the input and areas of the hypothesis space that are intermediately complex, from which most learning advantage may be extracted, from which it avoids complete systemic disorder, and in which one can maximise the available computational resources (as predicted by

Goldilocks and MMM). Systems meandering around the edge-of-chaos zone enjoy an adaptive advantage over other complex systems.

This edge-of-chaos and Goldilocks zone is thus the ‘progress niche’ (in the sense of Oudeyer, Kaplan & Hafner 2007, Oudeyer & Smith 2016) of our dynamical system, where complexity-driven activities progress to yield an improvement over mastery or control over an activity or knowledge and thus a reduction in uncertainty. In these systems, then, the pathway *complexity-to-simplicity-to-complexity* emerges without prespecification (Thelen & Smith 2007). The interaction of their parts and the environment will require the system in a currently ‘disordered’ state (complexity) to self-organise, leading to ‘order’ (simplicity), which, upon greater interaction with the input, will drive the system towards disorder. And the cycle continues. This is illustrated in the diagram in Figure 2.

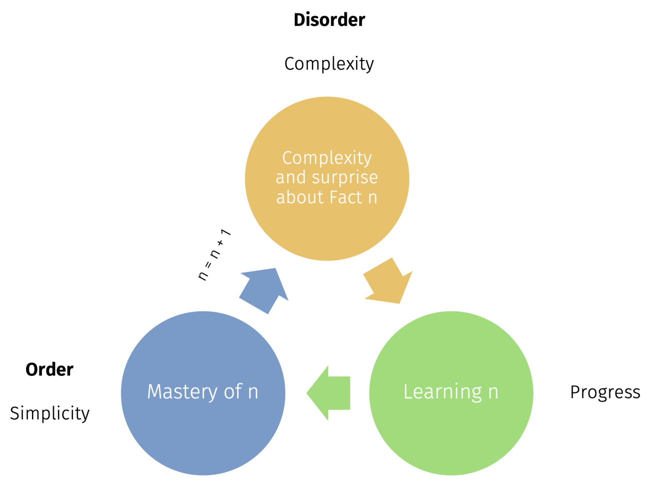


Figure 2 The Goldilocks Cycle.

Importantly, then, the picture that emerges, both in DST and independently in the language acquisition literature, is of acquirers exploiting the cues available to them and tackling them in such a way that the areas of the input with intermediate complexity get priority. Cumulative learning means that one begins with the problems that are ‘just right’ before one gets a clearer picture of more intricate ones. Language acquisition proceeds via a radical reduction of the input (in the words of Evers & van Kampen 2008: 498 ‘leave out what you cannot fit in’). If grammar construction is modellable in the systems theoretic way advocated here, what is precisely ‘left out’ in the reduction of input-to-intake (see also Lidz & Gagliardi 2015 for this distinction) are levels of either very low or very high complexity, from which less systemic development can be achieved. That is, children will preferentially focus on stimuli that they can *partially* integrate into their knowledge (cf. also Gagliardi 2012 on ‘partial encoding’ as a catalyst driving the language acquisition system forward). This edge-of-chaos perspective on MMM and complex linguistic systems may, then, be the dynamical essence behind Abler’s (1989) Humboldt systems.

Since the edge of chaos and strange attractors are intimately related to fractals, I will additionally be extending this Goldilocks-style approach to acquisition to our view of natural language fractals, and therefore suggest that there may exist an important link between MMM, Goldilocks, fractals and edge-of-chaos dynamics. While at the edge of chaos, the driver behind fractals is the strange attractor which attracts an organisation to a recognisable structure or process. It is worth contemplating the possibility, then, that the fractal structures observed in natural language syntax might be related to acquisitional dynamics that operate at an edge-of-chaos and Goldilocks zone.

Fractals pervade in natural language. Ever since Abney (1987), Chomsky (1970), Ritter (1991), Giusti (1997) and others, for example, syntacticians have been working with a fractalian view of the structural organisation of the nominal and clausal domain. A potential reason for the pervasiveness of fractals is that they have minimal Kolmogorov complexity (in any pragmatic computational model, in fact), and hence may be plausibly harnessed in zones that make maximal use of minimal means. The learner will, all other things equal, pick the analyses with the shortest description length of the linguistic input (Biberauer, Holmberg, Roberts & Sheehan 2014a; see also Chater & Vitányi 2003, on the role of simplicity in cognitive science and its relation to Kolmogorov complexity). Note also how the link is rather natural: inasmuch as strange attractors attract systems to a recognisable structure as set out above, it would not be surprising if fractals were tightly connected with acquisitional dynamics that focus on intermediate levels of complexity. If children paid attention to parts of stimuli they can *entirely* integrate into their knowledge (i.e., with very low complexity), the expectation is that this would lead to an identity condition, not a fractal. If fractals are a kind of structure that is partially encodable in this sense, one might predict their locus to be with Goldilocks and edge-of-chaos zones.

A word on the relationship between Goldilocks and MMM. For expository convenience, they have been discussed separately thus far. This is imprecise: the Goldilocks Principle may feasibly be a reflex of cognitive systems that maximise available resources, and so may be subsumed under MMM (see Biberauer 2018a, and Biberauer & Bosch 2021, for some general-cognitive parallels); the discussion here, where Goldilocks and MMM are embedded within an edge-of-chaos framework, points in exactly the same direction. If the links presented here are correct, it might be worthwhile considering whether fractals, Goldilocks and MMM dynamics are in fact linguistic and behavioural manifestations of a single strange attractor observed in complex non-linear systems. Therefore, they would have a unified source, presumably found in the transparent tendency of dynamical systems to operate at zones of behaviour where performance is optimised.

Readers may have noticed that nothing in the above provides any means with which to feasibly relate the structural fractals we observe in natural language syntax and the proposed fractalian and edge-of-chaos dynamics in language acquisition. To the extent that we have proposed that developmental fractals are plausibly linked to structural ones (namely, that structural fractals arise when the dynamical system follows fractalian paths), there is a noticeable leap in our argument. More generally, if we are to make any claims about the representations being derived from dynamics, one needs an explicit theory about the relationship between the two. We return to this in section 3.2.

3.1.3 Control parameters: features and emerging representations

As mentioned earlier, control parameters are cues or sets of cues that drive system reorganisations. Considering that control parameters have been often argued to be established and related to the edge-of-chaos zone (see Baym & Hübler 2006), I would like to entertain the possibility that control parameters play a direct role in MMM-driven feature postulation by the child.[4] More specifically, the possibility is that the cues to which feature postulation is sensitive are (at least) one of the linguistic analogues of control parameters. Put another way, the suggestion is that examining the types of correlations to which control parameters are sensitive may prove informative in deciding how children systematise the PLD they are confronted with (cf. also section 3.2 and section 4.2 later).

That control parameters are relevant for the study of language acquisition, and in particular bootstrapping, is not a new claim. In this respect, we highlight two notable sets of cues that have been linked to control parameters in the DST literature:

- **Contrast and discrepancies**: discrepancies between what the system expects (i.e., what the system has learned) and what the context provides drive the system forward through successive reorganisations. Imperfect relationships between expectations-reality provide these discrepancies and discontinuities, which ultimately motivate systemic development (Tucker & Hirsh-Pasek 1993)

- **Acoustic cues for prosodic bootstrapping**: ‘acoustic control parameters drive the perceptual system towards more detailed, language-specific relevant analysis’ (Hirsh-Pasek, Tucker & Golinkoff 2014: 444). These cues (e.g., those that coincide with boundaries of linguistic units) may include, but are not limited to: syllable lengthening, fundamental frequency shifts, and pausing.

The former point aligns directly with theories of (phonological and syntactic) featural organisations that piggyback on contrast, such as Dresher’s (2009) and Dresher’s (2014) Successive Division Algorithm, Cowper & Hall’s (2014) *Reductiō ad discrīmen*, Hall’s (2007) Contrasitivity Hypothesis and Biberauer’s (2019) ‘systematic departures from Saussurean arbitrariness’, among others. In accordance with these hypotheses, each formal feature (FF) is postulated and becomes part of the system if it fulfils a contrastive role in some context in the system being acquired. The role of contrast as a control parameter and as driver of feature postulation is probably well-motivated. The impact of prosodic bootstrapping for devising featural systems is less immediately clear, though. But considering that prosodic bootstrapping can provide some cues for the establishment of head-directionality (see the Iambic

---

[4] At the very minimum - this view could easily be extended to other computational units such as morphosyntactic categories and heads.

Trochaic Law and work by Christophe, Nespor, Guasti & van Ooyen 2003, Nazzi, Bertoncini & Mehler 1998, Wagner, Zurita & Zhang 2021), the link may be seen as less indirect (we return to some speculative suggestions on prosody and formal features in section 4.3).

It is conceivable that control parameters for prosodic domains and boundaries may be related to contrast and that they are another consequence of detecting systematic correlations between a cue and a relevant linguistic property, but we do not discuss this further (see Biberauer 2019: 52-53, on 'edges' and their relation to 'departures from Saussurean arbitrariness'). Regardless, it is likely not a mere coincidence that both dynamical systems and linguistic systems are swayed by contrast.

3.2 Dynamics vs representations: Who wins?

Dynamical systems have sometimes been framed as a replacement for computational representation. There exists a general commitment in some DST approaches to anti-representationalism (Haselager et al. 2003, van Gelder 1995, Smolensky 1988), namely to the idea that cognition is not concerned with the domain of symbol-level 'effective computation', as identified by the Church-Turing thesis and computability theory (*contra* Fodor & Pylyshyn 1988). For rather self-evident reasons, however, getting rid of linguistic representations would create a vast array of non-trivial issues and we disagree that DST is inconsistent with computational and information-processing cognitive science.

Indeed, not all DS theorists argue that dynamicism is diametrically opposed to representationalism, with some presenting them as distinct but complementary (Dale, Dietrich & Chemero 2009), or even consistent with each other given the implicit computations performed by dynamical systems (Giunti 2006, Siegelmann & Fishman 1998, Crutchfield & Mitchell 1995). We take the view here, with Dale et al. (2009), that explanatory pluralism is desirable in cognitive science. The idea that one theory or framework can singly account for the vast complexity and variety of linguistic and cognitive processes seems at best unlikely. There need not be anything in the mathematical description of a system's flow that precludes it from being representational. In fact, through a by-now prominent mathematical framework of analysis, symbolic dynamics, it can be shown that both systems of measurement (continuous/dynamic and discrete/symbolic) have important qualities of equivalence in the limit (Dale & Spivey 2005, Dale 2008, Edelman 2008).

Following this latter strand of work, the intuition to be sketched here is that the putatively incompatible duo of dynamics and representations can be bridged in a principled and illuminating way. Dynamics and representations are dissociable, but tightly interconnected. In accord with Dale & Spivey (2005), I suggest that a 'mathematisation' of the problem space in terms of complex non-linear systems can reconcile representation in higher-order cognition, revealing a possible epistemological synthesis of the dynamics-symbols debate in the context of acquisition.

Let us take as a point of departure the fact that dynamical systems can be conceptualised as geometric and topological spaces. Let us assume, too, that there exists a level of conceptual representation that is also geometric (as famously suggested in Gärdenfors's 2000, 2014, theory of Conceptual Spaces, where conceptual spaces are made up of spatially contiguous convex regions, split up via Voronoi tessellations). I will be following already-existing DST literature that assumes compatibility between representationalism and dynamicism (e.g., Shapiro 2013, Tabor 2009). In particular, largely following some proposals made in Fekete (2010), I set out to probe a possible topological mapping between the dynamics of a system and the conceptual spaces being derived from those dynamics. Activity spaces (i.e., spaces of spatio-temporal events generated by a dynamical system) act as representational media as a result of their geometry and topology. Instances of activity serve as the *vessels* of content within a representational system. The representations of a system, then, can be analysed, at least in part, in terms of the geometrical and topological properties of the systems' activity space.

The reasoning behind this mapping is the following:

> 'If the dynamics of a system are to give rise to representational content, i.e., make perceptual and conceptual information both explicit and accessible, this information should be "entirely present" in some sense, or intrinsic, to the system's activity. Anything less would mean that some other entity or process is taking care of business, or simply that representation hasn't been achieved by the system.' (Fekete 2010: 74)[5]

Now, whether or not representational content is *entirely* present in the dynamics of a system is probably subject to debate (see below, too), but that some relationship can be established between the two spaces seems promising. Instead of restricting this approach to dynamics and conceptual spaces only (as Fekete does), I suggest an extension of his model to strictly symbolic computation, in the spirit of Aisbett & Gibbon (2001). Conceptual spaces constitute a mesolevel between the macrolevel of symbolic computation and the microlevel of dynamics. The resulting architecture as conceived of here is schematised in Figure 3[6] (nothing precise should be read from this diagram, it merely serves a visualisation purpose).

Importantly, this view is compatible with independent DST models such as aforementioned symbolic dynamics (Dale & Spivey 2005), which attempts to analyse dynamical systems by *discretising* space and by demonstrating that a space of representational symbols Σ and its shift map σ (the progression in time of symbols emitted by a system) have a certain geometrical equivalence to a dynamical system's continuous mapping and the set of states it visits (e.g., a system's dynamics can be rendered symbolic by carving partitions in its phase space and assigning a unique

---

[5] On the importance of making the semantic interpretation of a formal symbolic system intrinsic to the actual system, see Harnad's (1990) symbol grounding problem.

[6] This is not to be confused with the three macro/meso/micro dimensions of dynamical systems introduced in section 2.1, which, unhelpfully, have received very similar names in the DST literature. The notions are distinct. System dimensions need to be seen as *intra*-systemic relations (namely, between the structures and environment of a single dynamical system), whereas the levels presented above are fundamentally *inter*-systemic relations (among three distinct systems: a dynamical system, conceptual spaces and a representational system).

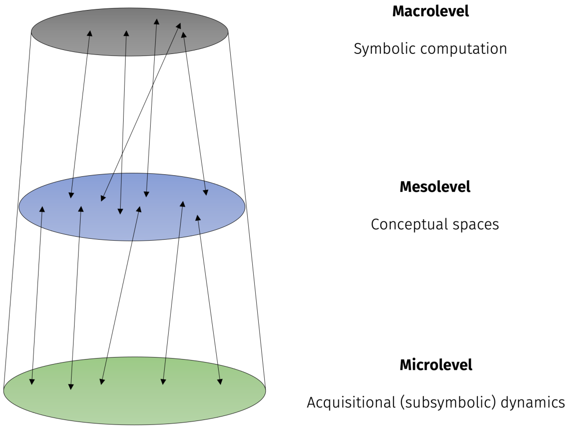


Figure 3 A topological mapping between acquisitional dynamics, conceptual spaces and representations.

numeric state or label to a partition). The systems (X, Φ) and (Σ, σ) are thus related to each other by:

(4) $\pi \circ \Phi = \sigma \circ \pi$

which can be represented as a commutative diagram:

$$\begin{tikzcd}
x \arrow[r, "\Phi"] \arrow[d, "\pi"'] & \Phi(x) \arrow[d, "\pi"] \\
s \arrow[r, "\sigma"] & \sigma(s)
\end{tikzcd}$$

where $\pi : X \rightarrow \Sigma$ acts as an intertwiner and where s consists of the sequence of symbols emitted in the space Σ (Atmanspacher & beim Graben 2007).[7]

The three-dimensional architecture in Figure 3 also abides by the principles of contextual emergence (i.a. Atmanspacher & beim Graben 2007, 2009, Bishop & Atmanspacher 2006), which seeks to characterise relationships between different domains and levels of description of particular phenomena, with increasingly more complex phenomena residing at higher levels of description. We can see, then, in which respects this contextual emergentist architecture shows how lower-level features are necessary (but not sufficient) conditions for the description of higher-level features. The whole is, yet again, more than sum of its parts.

Although this mapping remains unexplored in linguistics, the intuition that dynamics and representations are inherently symbiotic is widespread in the DST literature. It is worth stepping back and contemplating whether it is in fact so new in linguistics. The idea that representations in some way reflect the learning path (in the

---

[7] This approach of course only maps *two* spaces, rather than three as in Figure 3. I treat this formalisation as an illustration. My objective is not to provide an exhaustive mathematical breakdown of this tripartite mapping.

sense of Dresher 1999), is already implicit, to varying degrees, in some phonological and syntactic theories, such as Biberauer & Roberts's (2015) NO>ALL>SOME, or Dresher's (2009) Successive Division Algorithm. Additionally, from an MMM perspective it may not be highly surprising if dynamics and representations are tightly interconnected in the way presented here. If cognitive systems somehow maximise the use of available means and spaces, this case of 'recycling' an already-existent topological space as the basis for later representations would potentially be another instance (even if just metaphorically) in which neurological, biological and cognitive agents take advantage of available resources for future goals.

Overall, then, we observe a general architecture in which the acquisitional dynamics are topologically mapped onto conceptual spaces, which, in turn, are mapped to computational symbols, such that it enables the establishment of equivalence and ontological relations between these (*a priori* quite distinct) levels of abstraction. Although, as presented here, there is indeed a substantial degree of unidirectionality in this dynamics-to-symbols pathway, it is also sensible to expect a fair amount of interactionalism and bidirectionality from a systems-theoretic perspective (see the discussion on system dimensions in section 2.2). For neo-emergentism and linguistics, this is everything but implausible: we expect the current state of the representational/conceptual level to have some top-down effects on the acquisitional dynamics. This is precisely what seems to occur: take as illustratory examples the observation that participants do not seem to consider non-convex categories in artificial learning experiments (Heffner, Idsardi & Newman 2019), or the general truism that L1 acquisition is structure dependent. Such a bidirectional flow of information, aside from being psychologically plausible (cf. top-down and bottom-up effects in processing), also restricts the search and phase space in a welcome manner, in that certain domains of the state space need not be explored if the representational or conceptual dimensions rule them out in a principled way.

We assume this mapping to be, at the very least, a useful metaphorical tool, but the possibility of it being mathematically real is tantalising. Of course, the outstanding question is what sort of rigidness we afford this mapping, namely how high a degree of isomorphism we assume to hold between a dynamical system and its higher levels of organisation. Determining which mathematical mapping is the correct one is well beyond the scope of this work. However, certain topological mappings are arguably more suited to language acquisition than others.

The correspondence suggested in contextual emergentist architectures gives rise to two ways of describing the system, which beim Graben & Atmanspacher (2006) refer to as the 'ontic' (vector space) and 'epistemic' (symbolic) levels. beim Graben (2004) goes on to suggest that the dynamical notion of *topological equivalence* (or *topological conjugacy*) can formalise compatibility between levels of description. Two systems $f : X \to X$ and $g : Y \to Y$ are topologically conjugate if there exists a homeomorphism among them, i.e., $f : X \to X$ and $g : Y \to Y$ are effectively the same topological space and thus isomorphic to each other.

However, a note of caution is in order. The most evident weakness in this equivalence is that it assumes an isomorphism between dynamics and representations, which may be too restrictive for language acquisition. Whether such an intimate connection is appropriate for our linguistic purposes here certainly requires additional attention. A less unbending avenue would be to take the two topological domains as being homotopically equivalent. In the case of homotopy, the continuous deformation from one map to the other is the *essence* of the spaces (a map between *functions* of two topological spaces), and it is also less rigid, since none of the maps involved need to be bijective (i.e., one-to-one and onto). For simplicity, however, I will not go into the mathematical properties of homeomorphisms and homotopies, nor discuss them further.[8]

The aim here is neither to prove that such a specific mathematical mapping exists, nor to adjudicate between possible topological equivalences. We can take, for now, this mapping to be metaphorical, but the establishment of some precise mathematical equivalence between these topological spaces, albeit speculative, does not seem impossible to formulate. I leave this open, but I outline the (potentially wide-ranging) consequences this theory of representations and acquisitional dynamics would have for linguistic theory in the coming section.

4 Novel Predictions of the Model

The following section aims to show that the proposed model is *not* a mere restatement of already-existing acquisitionist literature. A systems-theoretic model of grammar construction makes a series of new predictions and results, which appear to be both theoretically and empirically sound.

In the interest of space, we will consider the following predictions and consequences, but the list is by no means exhaustive: those relating to the general architectural links between dynamics and representations, and their results for representations and learning paths in phonological and syntactic theory, and, on the other hand, those relating to theories of formal features in language acquisition.

4.1 Emergence in a dynamic and representationalist system

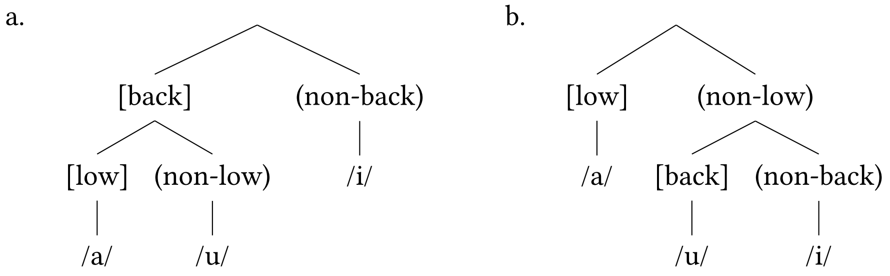
The general picture of representations outlined in section 3.2 puts forward a view of symbolic systems that hinges on learnability and conceptual spaces to a considerable degree.

If acquisition and representation are as intimately interweaved as suggested, this has important consequences: most (if not all) representations in steady-states should reflect the learning path, itself derivable from the dynamics. Although a strong prediction, it has not, again, gone untouched in linguistics. Dresher's (2009) Successive Division Algorithm outputs distinct orderings for features subject to which contrast is teased apart first by the acquirer. This is shown below, where the distinct orderings lead to different patterns of contrastive activity. In (5a) both /a/ and /u/ can trigger backing, since both are specified as [back]; in (5b) only /u/ is [back], so neither /a/ nor /i/ should cause backing (Dresher 2021):

---

[8] See `https://en.wikipedia.org/wiki/Homeomorphism` and `https://en.wikipedia.org/wiki/Homotopy` for their respective mathematical definitions.

(5)



A side-effect of adopting Gärdenfors's (2000) Conceptual Spaces[9] as a mesolevel in the architecture is that symbolic representations should obey the same structural principles of conceptual representations. A notable one is their convexity and contiguity requirement, which postulate visual and linguistic categories that are self-contained domains of homogenous properties. This indeed readily visible in lexical acquisition (Dautriche & Chemla 2016, Xu & Tenenbaum 2007, Plunkett, Hu & Cohen 2008), as well as in both perceptual and linguistic categories (for the former, see the discussion in Biberauer & Bosch 2021, on Structural Information Theory and Recognition-by-components Theory; as regards the latter, see Gärdenfors 2000, Gärdenfors 2014, for convexity in semantic categories and Heffner et al. 2019, for evidence that phonological systems and categories are encoded in a convexity-like manner).

A related, and far-reaching, consequence of the above (particularly if an isomorphism between the topological spaces is assumed) is that it becomes imperative to consider lower-level dimensions (learnability) when theorising about higher-level ones (symbolic computation). Contextual emergence dictates that lower-level descriptions provide the necessary foundations for more fine-grained and higher levels of description. As an overall goal of cognitive science remains to relate the genetic, neural, cognitive and behavioural accounts of human development, devising theories that specifically relate the different levels of observation is crucial (Johnson 2011).

More fundamentally, the proposed mapping brings us to consider, too, what is universal and variable in learning paths. A mapping-based view of representations which hinges on the flow, the hallmark properties of dynamical systems (self-organisation, emergence, edge of chaos, control parameters, among others) and a radically impoverished UG suggests that what may be crosslinguistically universal in learning paths is precisely the geometric properties of the flow, rather than the substantive content attached to it. That is, when dealing with universalism in learning paths, we need to draw a distinction between shape and substance.

4.2 Shape and substance: DST-resultant learning paths

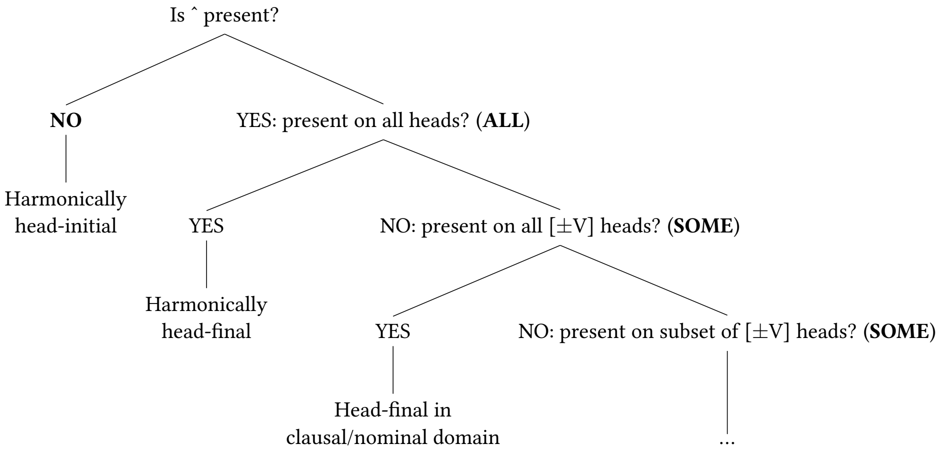
Both FFs and general development are emergent and language-specific. The general *shape* of variation and acquisitional dynamics, on the other hand, is crosslinguisti-

---

[9] Thanks are due to Theresa Biberauer for drawing my attention to this theory and its potential implications for syntax.

cally consistent. In a parametric approach to language variation (following the Borer-Chomsky Conjecture), this takes the form of a NO>ALL>SOME learning path, as derived via MMM (Biberauer & Roberts 2015, Biberauer 2019). NO>ALL>SOME is schematised in (6) for head-directionality parameters:

(6)



Crucially, however, it is not predicted that Roberts's (2019) parameter hierarchies will be universally ordered in the same way across all possible L1s, just like it is not predicted that contrastive phonological hierarchies will have identical orderings crosslinguistically (cf. our discussion on Dresher's SDA above and see also Biberauer 2018b, for this same proposal in the context of pro-drop). Distinct orderings of geometries and hierarchies follow directly from sensitivity to initial conditions and the mapping proposal in section 4.1, in which representations will derive (more or less isomorphically, depending on the equivalence assumed) from acquisitional dynamics. If a given feature or generalisation is acquired earlier in some languages, but not others, this will necessarily result in divergent representations.

We need to understand, then, how it is that the present proposal does not predict rampant and unconstrained crosslinguistic variation, in acquisition, linguistic representations and in language typology. Notably, the role of attractors in acquisitional dynamics (e.g., MMM, Goldilocks and the NO>ALL>SOME learning path) is to ensure some crosslinguistic uniformity in learning, but they still leave important room for substantial, though crucially constrained, variation among L1s (see Rice & Avery 1995 on uniformity and variation in learning paths). The important point that ought to emerge from the discussion so far is that the 'landscape' of a dynamical system makes any form of teleological drive superfluous. Considering the emphasis on self-organising 'sofly-assembled' development and since UG provides no relevant substantive content to acquirers, there is possibly no or very little room for (biologically-instigated) maturation in this model (*pace*, i.a., Radford 1988, Friedman, Belletti & Rizzi 2020, Rizzi 2021). It is the combination of initial conditions, attractors, the system's three dimensions (recall section 2.2) and their interaction with abstract control parameters that determines the shape of a system's flow throughout its phase space (see also Lass 1997 for a similar discussion applied to language change). This is also in line with our assumptions about UG: Universal Grammar provides shape specifications, not substance. In this approach, then, just like in MMM more generally, the likelihood of observing crosslinguistically universal (that is, formally identical) syntactic categories is considerably low (cf. Ritter & Wiltschko 2014, Haspelmath 2010).

DST as applied here leads to a view where there is crosslinguistic uniformity, due to common shape, topology and attractors, but what is exactly picked up on at each stage of acquisition in each L1 will be subject to variation. The content and the order in which elements are tackled is input-dependent, language-dependent and environment-dependent. This leads to an apparent, but sound contradiction: crosslinguistically, learning paths are *dynamically similar*, yet *substantively dissimilar*.

Together with the consequences of a topological mapping (section 4.1), this may prompt a reanalysis of how certain feature geometries of hierarchies in the literature (e.g., those in Roberts 2019, Harley & Ritter 2002) are interpreted. Taking as a main hypothesis stemming from the work here the fact that the shape of learning paths is crosslinguistically stable, it is likely that the content and orderings of a parameter hierarchy will vary somewhat from language to language (e.g., as a function of which formal features and contrasts are postulated earlier or later in a given L1).[10] This would catalyse various consequences, notably the fact that the same parameter might be more or less marked depending on the L1 being acquired (particularly if one understands markedness to be associated with acquisition difficulty and complexity, cf. Biberauer et al. 2014a).

4.3 The flexible life of features

The point about how features develop in an emergentist and dynamical approach deserves further attention. The standard proposal in the nativist literature has been that children make a one-time selection from an inventory of formal features supplied by UG (Chomsky 2001). In emergent feature theories, likewise, there is often the tacit assumption that once a relevant feature is postulated, it ‘stays as is’, and, even if embedded in feature geometries, the emergence of distinct FFs (e.g., phonological, syntactic, semantic) are in some sense independent of each other. Yet there has been little consideration in both sides for the possibility that features may neither be so ontologically independent of each other nor as immobile. Precisely, if one adopts a truly emergentist approach to FFs, it requires a stipulation to render them ‘fixed’ once the child has filled in (some of) their substantive content.

One of the hallmark properties of dynamical systems, in this context, is the idea of feedback, in which the output of some processes within the system is ‘recycled’ and becomes a new input for the system. Additionally, we know that, in a system’s control parameters, the cues associated with a particular control parameter may vary and that the weights assigned to control parameters can change in order to direct the system towards more detailed, language-relevant analyses.

---

[10] See Cowper & Hall (2019), who reanalyse morphosyntactic feature geometries as acquisitionally-informed contrastive hierarchies.

An extension of this observation that control parameters can change profile and specifications during time, and that grammars are built and refined incrementally as a function of the feedback loops with their environment and present organisation, would be to hypothesise that linguistic dynamical systems may likewise 'recycle' the use made of a given feature and that its specification can get progressively refined during acquisition, following the general-to-specific pathway that characterises these systems. The former point is not novel in generative syntax (see Biberauer 2019), but the latter has been considerably less explored.[11] If true, it predicts the existence of acquisitional stages where a feature has been successfully postulated but its substantive content has not yet reached adult-like status.

In other words, the neo-emergentist view on FFs, coupled with this dynamical framework, brings us to reconsider whether FFs are so ontologically rigid. To assume that children latch on directly to the correct substantive specification of a FF is not impossible, but it may be acquisitionally more credible to probe this DS intuition that the content and function of FFs is not perennially fixed upon postulation (a 'no-going-back' scenario), and that it is instead subject to some refinement. FFs are malleable in acquisition. The resulting idea follows Hale's (1986) suggestion that features are 'semantically broad, ontologically flexible, and category independent' (cf. also Song 2019 on the formal flexibility of syntactic categories). Such an approach may therefore present a clearer picture as to how FFs precisely *emerge* throughout the course of development and eventually reach a steady-state system.

For concreteness, we may call this the *Formal Feature Adaptability Hypothesis* (or FFAH), which is formulated as a bipartite statement:

(7) **The Formal Feature Adaptability Hypothesis (FFAH)**:
  a. The formal features of complex dynamical (linguistic) systems may get 'recycled' throughout the course of acquisition and may take on new, but developmentally related, functions (following Biberauer 2019).
  b. The substantive content of a feature may start out being comparatively underspecified and may undergo later refinement.

The question about the formal and substantive flexibility of features does not really arise if features are encoded in UG and linked to Logical Form in the genetic encoding. Yet, it is not straightforward how to solve Pinker's (1984) Linking Problem for both innate morphosyntactic and phonological features. Even if they are somehow linked to some perceptual, articulatory or semantic correlate, these correlates may differ in important respects from language to language, as has already been noted in the literature (see, i.a., Ritter & Wiltschko 2014 on [±COIN(CIDENCE)] or Liter, Huelskamp, Heffner & Schmitt 2018 on crosslinguistic interpretations of number features; for some issues with innate phonological features, see Mielke 2008). It is thus unclear what learnability advantages an innate inventory of FFs would bring to the child if the Linking Problem is still non-trivial. What is more, if features

---

[11] Though the possibility has been suggested for linguistic categories, e.g., see Fourtassi & Dupoux (2014) for a proof of principle where the learner first acquires approximate representations before progressively refining and constraining them.

become incrementally more fine-grained and specific in the manner outlined above, then it provides a motivation for why an inventory of FFs in UG will be largely unhelpful to the child: UG-given features provide steady-state specifications, yet what FFAH is suggesting instead is that there is more to the life of features than what is readily visible in steady-state grammars. If FFAH is on the right track (or at least, proves a useful heuristic), this is another example of how taking a poor UG as a point of departure raises productive questions that would otherwise not arise in a rich-UG world.

There is no room to discuss or develop FFAH further here (I leave this to future work), but I comment on some observations independently made in syntax, phonology and acquisition that suggest 'feature recycling and refining' may be a productive logic.

Take, for expository purposes, the research demonstrating *in utero* and very early sensitivity to various aspects of prosody (Gervain & Werker 2008), which appears to allow 6-month-old infants' to distinguish the basic head-directionality of the system they are acquiring (i.a., Christophe et al. 2003, Nazzi et al. 1998; cf. also Wexler 1998, on basic word-order as 'very early parameter'). Very roughly, phonological phrases in OV languages have a strong(=complement)-weak(=head) prosodic outline ($[\mathbf{s}\mathrm{w}]_{\phi}$), while VO shows weak(=head)-strong(=complement) prosody ($[\mathrm{w}\mathbf{s}]_{\phi}$). One perspective on this state of affairs would be to interpret it as following from the fact that the detection of prosodic regularities (call them 'cues' or 'control parameters') may feed directly into their subsequent recycling into formal, syntactic features encoding head-directionality. More specifically, what we could suggest is that we are not taking seriously enough the possibility of an inherent interactional and developmental connection between linguistic modules that have thus-far been treated as largely modular and encapsulated, and that seemingly independent prosodic cues may be the coarse-grained formal basis for later syntactic ones.[12]

Consider also, as a second example, the possibility of recycling features in syntactic domains beyond the one they were initially postulated for. Harbour (2020) discusses so-called 'Frankenduals', a crosslinguistically widespread phenomenon where a non-primitive dual is stitched together out of singular and plural components. This is shown in (8a), (8b) and (8c) for Hopi (Uto-Aztecan):

(8) a. *Pam* *wari*
that.SG run.NPL

'(S)he ran.'

b. *Puma* *wari*
that.NSG run.NPL

'They$_2$ ran.'

---

[12] Note that this may imply these strong-weak regularities would be encoded as respective *prosodic* features before 'developing' into syntactic head-directionality features. This is not the standard analysis, but I see nothing in the MMM model, and in particular its emphasis on contrast and systematic regularities, that would prevent these 'cues' from being encoded as features (in fact, see Biberauer 2017, especially fn. 7, who independently arrives at a similar proposal).

c. *Puma* *yùutu*
that.nsg run.pl

'They$_{3+}$ ran.'

Frankendurals show two distinct categorial behaviours of [±minimal] at once: the verbal distribution characteristic of aspectual uses and the nominal interpretation diagnostic of its numeral use. Supporting Hale's (1986) case for ontologically flexible, category-independent features, Harbour (2014) likewise argues that [±minimal] is not a number feature as such, but is logically equivalent to concepts of aspect/telicity. As shown in the patterns in (8), the [±minimal] feature on the verb contributes to the nominal number only when the noun is unspecified for that feature. If the noun is specified for [±minimal], the verb takes its value from the noun instead. This 'feature trading', as Harbour calls it, has other precedents in morphosyntax, such as Adger & Harbour's (2007) treatment of the Person Case Constraint (PCC). It is suggested the applicative head requires an argument with [±participant], but where the argument is 3rd person (which lacks specification for [±participant]), the applicative head itself can endow third person arguments with a [-participant] specification (for other recycling phenomena, see Biberauer 2019, Douglas 2018, Wiltschko 2014, Ritter & Wiltschko 2014).

That features will be 'flexible' and 'polyvalent' is expected in a structurally homologous system that will 'recycle' available FFs and employ options already available in its current organisation. The fractal-like structures mentioned in section 2 and section 3 would appear to be another case where structural organisations or FFs in a system can be reused in distinct transcategorial domains. Recall also the edge-of-chaos dynamics in these systems: if they encode linguistic organisations that are 'just right' in learnability terms and maximise applicability of a feature while minimising its description length, it is computationally advantageous to select a feature that can be distributionally adaptable. In fact, if edge-of-chaos dynamics are a driving force of complex adaptive systems, we expect to find a significant repertoire of fundamentally flexible features, either at some point during development or at steady states.

5 Conclusion

My goal here has been to attempt to show why it is productive for generativists to incorporate systems-theoretic notions in their theoretical models. I introduced the basis of a model of grammar construction that draws together current neo-emergentist approaches to learnability with Dynamical Systems Theory and exploits their strengths. As such, the nature of this work remains programmatic, but, if adopted, it provides a natural framework with which to characterise well-known observations in language acquisition, syntax, phonology and language variation and from which properties of linguistic systems can fall out. Additionally, such an approach provides innovative tools and means with which to shed new light on the content of Universal Grammar, the three factors of language design and perspectives on featural and representational systems. Most importantly, it allows us to see whether we can start making progress in integrating together mathematical systems, learnability and language in such a way that it embellishes our understanding of the nature of all three.

REFERENCES

Abels, K. & A. Neeleman. 2012. Linear asymmetries and the LCA. *Syntax* 15(1). 25–74.

Abler, W. L. 1989. On the particulate principle of self-diversifying systems. *Journal of Social and Biological Structures* 12(1). 1–13.

Abney, S. P. 1987. *The english noun phrase in its sentential aspect*: Massachusetts Institute of Technology dissertation.

Ackema, P. & A. Neeleman. 2002. Effects of short-term storage in processing rightward movement. In S. Nootebook, F. Weerman & F. Wijnen (eds.), *Storage and Computation in the Language Faculty*, 219–256. New York: Springer.

Adger, D. & D. Harbour. 2007. Syntax and syncretisms of the Person Case Constraint. *Syntax* 10(1). 2–37.

Aisbett, J. & G. Gibbon. 2001. A general formulation of conceptual spaces as a meso level representation. *Artificial Intelligence* 133(1-2). 189–232.

Atmanspacher, H. & P. beim Graben. 2007. Contextual emergence of mental states from neurodynamics. *Chaos and Complexity Letters* 2(2-3). 151–168.

Atmanspacher, H. & P. beim Graben. 2009. Contextual emergence. *Scholarpedia* 4(3). 7997.

Baker, M. 2008. The macroparameter in a microparametric world. In T. Biberauer (ed.), *The Limits of Syntactic Variation*, 351–373. Amsterdam: John Benjamins.

Bavin, E. (ed.). 2009. *The Cambridge Handbook of Child Language*. Cambridge: Cambridge University Press.

Baym, M. & A. W. Hübler. 2006. Conserved quantities and adaptation to the edge of chaos. *Physical Review E* 73(5). 056210.

Beggs, J. M. 2008. The criticality hypothesis: how local cortical networks might optimize information processing. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 366(1864). 329–343.

Beggs, J. M. & N. Timme. 2012. Being critical of criticality in the brain. *Frontiers in Physiology* 3. 163.

Berwick, R. C. & N. Chomsky. 2016. *Why Only Us: Language and Evolution*. Cambridge, MA: MIT Press.

Biberauer, T. 2011. In defence of lexico-centric parametric variation: two 3rd factor-constrained case studies. Paper presented at the *Workshop on Formal Grammar and Syntactic Variation: Rethinking Parameters* (Madrid).

Biberauer, T. 2017. Word-order variation and change in systems that Maximize Minimal Means. Talk presented at *Variation and Change in the Verb Phrase* (Oslo), 16 May.

Biberauer, T. 2018a. Less IS More: On the Tolerance Principle as a manifestation of Maximize Minimal Means. *Linguistic Approaches to Bilingualism* 8(6). 707–711.

Biberauer, T. 2018b. Pro-drop and emergent parameter hierarchies. In F. Cognola & J. Casalicchio (eds.), *Null Subjects in Generative Grammar: A Synchronic and Diachronic Perspective*, 94. Oxford: Oxford University Press.

Biberauer, T. 2019. Factors 2 and 3: Towards a principled approach. *Catalan Journal of Linguistics* 45–88.

Biberauer, T. & N. Bosch. 2021. The case for Maximise Minimal Means: linguistic and general-cognitive perspectives. Talk presented at *SyntaxLab* (University of Cambridge), 16 November.

Biberauer, T., A. Holmberg, I. Roberts & M. Sheehan. 2014a. Complexity in comparative syntax: The view from modern parametric theory. In F. Newmeyer & L. Preston (eds.), *Measuring grammatical complexity*, 103–127. Oxford: Oxford University Press.

Biberauer, T. & I. Roberts. 2015. Rethinking formal hierarchies: a proposed unification. *Cambridge Occasional Papers in Linguistics* 7(1). 1–31.

Biberauer, T., I. Roberts & M. Sheehan. 2014b. No-choice parameters and the limits of syntactic variation. In *Proceedings of the 31st West Coast Conference on Formal Linguistics*, 46–55. Somerville, MA: Cascadilla Proceedings Project.

Bishop, R. C. & H. Atmanspacher. 2006. Contextual Emergence in the Description of Properties. *Foundations of Physics* 36(12). 1753–1777.

de Bot, K., W. Lowie & M. Verspoor. 2007. A Dynamic Systems Theory approach to second language acquisition. *Bilingualism: Language and cognition* 10(1). 7–21.

Bowers, R. 1990. Mountains are not cones. In J. Atlantis (ed.), *Georgetown University Round Table on Languages and Linguistics (GURT) 1990: Linguistics, Language Teaching and Language Acquisition: The Interdependece of Theory, Practice and Research*, 123. Washington, DC: Georgetown University Press.

Chan, S. 2001. Complex Adaptive Systems. ESD.83 Research Seminar in Engineering Systems.

Chater, N. & P. Vitányi. 2003. Simplicity: a unifying principle in cognitive science? *Trends in Cognitive Sciences* 7(1). 19–22.

Chialvo, D. R. 2010. Emergent complex neural dynamics. *Nature Physics* 6(10). 744–750.

Chomsky, N. 1970. Remarks on nominalization. In R. Jacobs & P. Rosenbaum (eds.), *Reading in English Transformational Grammar*, 184–221. Oxford: Ginn.

Chomsky, N. 1993. A minimalist program for linguistic theory. In K. Hale & S. J. Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, 1–52. Cambridge, MA: MIT Press.

Chomsky, N. 1995. *The Minimalist Program.* Cambridge, MA: MIT Press.

Chomsky, N. 2001. Derivation by Phase. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language*, 1–50. Cambridge, MA: MIT Press.

Chomsky, N. 2005. Three factors in language design. *Linguistic Inquiry* 36(1). 1–22.

Chomsky, N. 2010. Poverty of stimulus: Unfinished business. Transcript of a presentation given at Johannes-Gutenberg University, Mainz.

Christophe, A., M. Nespor, M. T. Guasti & B. van Ooyen. 2003. Prosodic structure and syntactic acquisition: the case of the head-direction parameter. *Developmental Science* 6(2). 211–220.

Cowper, E. & D. C. Hall. 2014. Reductio ad discrimen: Where features come from. *Nordlyd* 41(2). 145–164.

Cowper, E. & D. C. Hall. 2019. Scope Variation in Contrastive Hierarchies of Morphosyntactic Features. In D. Lightfoot & J. Havenhill (eds.), *Variable Properties in Language: Their Nature and Acquisition*, 27–41. Washington, DC: Georgetown University Press.

Crutchfield, J. P. & M. Mitchell. 1995. The evolution of emergent computation. *Proceedings of the National Academy of Sciences* 92(23). 10742–10746.

Culbertson, J. & S. Kirby. 2016. Simplicity and Specificity in Language: Domain-General Biases Have Domain-Specific Effects. *Frontiers in Psychology* 6. 1964.

Dale, R. 2008. The possibility of a pluralist cognitive science. *Journal of Experimental and Theoretical Artificial Intelligence* 20(3). 155–179.

Dale, R., E. Dietrich & A. Chemero. 2009. Explanatory Pluralism in Cognitive Science. *Cognitive Science* 33(5). 739–742.

Dale, R. & M. J. Spivey. 2005. From apples and oranges to symbolic dynamics: a framework for conciliating notions of cognitive representation. *Journal of Experimental & Theoretical Artificial Intelligence* 17(4). 317–342.

Dautriche, I. & E. Chemla. 2016. What Homophones Say about Words. *PloS ONE* 11(9). e0162176.

De Wolf, T. & T. Holvoet. 2004. Emergence Versus Self-Organisation: Different Concepts but Promising When Combined. In S. Brueckner, G. Di Marzo Serugendo, A. Karageorgos & R. Nagpal (eds.), *Engineering Self-Organising Systems*, 1–15. New York: Springer.

Douglas, J. 2018. Māori subject extraction. *Glossa: a journal of general linguistics* 3(1). 110.

Dresher, B. E. 1999. Charting the Learning Path: Cues to Parameter Setting. *Linguistic Inquiry* 30(1). 27–67.

Dresher, B. E. 2009. *The Contrastive Hierarchy in Phonology*. Cambridge: Cambridge University Press.

Dresher, B. E. 2014. The arch not the stones: Universal feature theory without universal features. *Nordlyd* 41(2). 165–181.

Dresher, B. E. 2021. Contrastive hierarchies and phonological primes. *Perspectives on Element Theory* 143. 33.

Edelman, S. 2008. On the nature of minds, or: truth and consequences. *Journal of Experimental & Theoretical Artificial Intelligence* 20(3). 181–196.

Evers, A. & J. van Kampen. 2008. Parameter setting and input reduction. In T. Biberauer (ed.), *The Limits of Syntactic Variation*, 483–515. Amsterdam: John Benjamins.

Fekete, T. 2010. Representational systems. *Minds and Machines* 20(1). 69–101.

Fenzl, N. 2018. Emergence and Self-Organization of Complex Systems: the role of energy flows and information a philosophical approach. *Complexitas–Revista de Filosofia Temática* 2(1). 31–45.

Fodor, J. A. & Z. W. Pylyshyn. 1988. Connectionism and cognitive architecture: A critical analysis. *Cognition* 28(1-2). 3–71.

Fourtassi, A. & E. Dupoux. 2014. A Rudimentary Lexicon and Semantics Help Bootstrap Phoneme Acquisition. In *Proceedings of the Eighteenth Conference on Computational Natural Language Learning*, 191–200. Baltimore, Maryland.

Friedman, N., A. Belletti & L. Rizzi. 2020. Growing Trees. The Acquisition of the Left Periphery. *lingbuzz/005369* .

Gagliardi, A. C. 2012. *Input and Intake in Language Acquisition*: University of Maryland, College Park dissertation.

Gärdenfors, P. 2000. *Conceptual Spaces: The Geometry of Thought.* Cambridge, MA: MIT Press.

Gärdenfors, P. 2014. *The Geometry of Meaning: Semantics Based on Conceptual Spaces.* Cambridge, MA: MIT Press.

van Gelder, T. 1995. What might cognition be, if not computation? *The Journal of Philosophy* 92(7). 345–381.

van Gelder, T. 1998. The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences* 21(5). 615–628.

Gervain, J. & J. F. Werker. 2008. How Infant Speech Perception Contributes to Language Acquisition. *Language and Linguistics Compass* 2(6). 1149–1170.

Giunti, M. 2006. Is being computational an intrinsic property of a dynamical system? In G. Minati, E. Pessa & M. Abram (eds.), *Systemics of Emergence: Research and Development*, 683–694. New York: Springer.

Giusti, G. 1997. The categorial status of determiners. In *The New comparative Syntax*, 95–123. New York: Wiley Blackwell.

Glendinning, P. 1994. *Stability, Instability and Chaos: An Introduction to the Theory of Nonlinear Differential Equations.* Cambridge: Cambridge University Press.

Gottlieb, G. 1991. Epigenetic systems view of human development. *Developmental Psychology* 27(1). 33–34.

Gottlieb, G. 1992. *Individual development and evolution: The genesis of novel behavior.* Oxford: Oxford University Press.

beim Graben, P. 2004. Incompatible Implementations of Physical Symbol Systems. *Mind and Matter* 2(2). 29–51.

beim Graben, P. & H. Atmanspacher. 2006. Complementarity in Classical Dynamical Systems. *Foundations of Physics* 36(2). 291–306.

Hale, K. 1986. Notes on World View and Semantic Categories: some Warlpiri Examples. In P. Muysken & H. van Riemsdijk (eds.), *Features and projections*, 233–256. Berlin: De Gruyter Mouton.

Hall, D. C. 2007. The Role and Representation of Contrast in Phonological Theory. *Toronto Working Papers in Linguistics* .

Harbour, D. 2014. Paucity, abundance, and the theory of number. *Language* 90. 185–229.

Harbour, D. 2020. Frankenduals: Their typology, structure, and significance. *Language* 96(1). 60–93.

Harley, H. & E. Ritter. 2002. Person and number in pronouns: A feature-geometric analysis. *Language* 78(3). 482–526.

Harnad, S. 1990. The symbol grounding problem. *Physica D: Nonlinear Phenomena* 42(1-3). 335–346.

Haselager, W. F., R. M. Bongers & I. van Rooij. 2003. Cognitive science, representations and Dynamical Systems Theory. In W. Tschacher & J.-P. Dauwalder (eds.), *The Dynamical Systems Approach to Cognition: Concepts and Empirical Paradigms based on Self-organization, Embodiment, and Coordination Dynamics*, 229–241. Singapore: World Scientific.

Haspelmath, M. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86(3). 663–687.

Hauser, M. D., N. Chomsky & W. T. Fitch. 2002. The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science* 298(5598). 1569–1579.

Hawkins, J. 2021. Attractors in Dynamical Systems. In J. Hawkins (ed.), *Ergodic dynamics: From basic theory to applications*, 27–39. New York: Springer.

Heffner, C. C., W. J. Idsardi & R. S. Newman. 2019. Constraints on learning disjunctive, unidimensional auditory and phonetic categories. *Attention, Perception, & Psychophysics* 81(4). 958–980.

Hirsh-Pasek, K., M. Tucker & R. M. Golinkoff. 2014. Dynamic Systems Theory: Reinterpreting ”Prosodic Bootstrapping” and Its Role in Language Acquisition. In J. Morgan & K. Demuth (eds.), *Signal to Syntax*, 461–478. Hove, East Sussex, UK: Psychology Press.

von Humboldt, F. W. H. A. 1836. *Über die Verschiedenheit des menschlichen Sprachbaues und ihrer Einfluss auf die Geistige Entwickelung des Menschengeschlechts.* Berlin: Dümmler.

Johnson, M. H. 2011. Interactive Specialization: A domain-general framework for human functional brain development? *Developmental Cognitive Neuroscience* 1(1). 7–21.

Kayne, R. 2013. Why Are There No Directionality Parameters? In T. Biberauer & M. Sheehan (eds.), *Theoretical Approaches to Disharmonic Word Order*, 219–245. Oxford: Oxford University Press.

Kidd, C., S. T. Piantadosi & R. N. Aslin. 2012. The Goldilocks Effect: Human Infants Allocate Attention to Visual Sequences That Are Neither Too Simple Nor Too Complex. *PloS ONE* 7(5).

Kidd, C., S. T. Piantadosi & R. N. Aslin. 2014. The Goldilocks Effect in Infant Auditory Attention. *Child Development* 85(5). 1795–1804.

Langton, C. G. 1990. Computation at the edge of chaos: Phase transitions and emergent computation. *Physica D: Nonlinear Phenomena* 42(1-3). 12–37.

Larsen-Freeman, D. 1997. Chaos/Complexity Science and Second Language Acquisition. *Applied Linguistics* 18(2). 141–165.

Lass, R. 1997. *Historical linguistics and language change*, vol. 81. Cambridge: Cambridge University Press.

Lidz, J. & A. Gagliardi. 2015. How Nature Meets Nurture: Universal Grammar and Statistical Learning. *Annu. Rev. Linguist.* 1(1). 333–353.

Liter, A., T. Huelskamp, C. C. Heffner & C. Schmitt. 2018. Grammaticalized number, implicated presuppositions, and the plural. *Glossa: a journal of general linguistics* 3(1). 39.

Lund, K., P. Basso Fossali, A. Mazur-Palandre & M. Beldame. forthcoming. *Language is a complex adaptive system.* Berlin: Language Science Press.

Mielke, J. 2008. *The Emergence of Distinctive Features.* Oxford: Oxford University Press.

Milnor, J. 1985. On the concept of attractor. *Communications in Mathematical Physics* 99. 177–195.

Muñoz, M. A. 2017. Criticality and dynamical scaling in living systems. *arXiv preprint arXiv:1712.04499* .

Nazzi, T., J. Bertoncini & J. Mehler. 1998. Language discrimination by newborns: toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance* 24(3). 756.

Niyogi, P. & R. C. Berwick. 1997. A Dynamical Systems Model for Language Change. *Complex Systems* 11(3). 161–204.

Ochs, K. 2011. A comprehensive analytical solution of the nonlinear pendulum. *European Journal of Physics* 32(2). 479.

Oudeyer, P.-Y., F. Kaplan & V. V. Hafner. 2007. Intrinsic Motivation Systems for Autonomous Mental Development. *IEEE Transactions on Evolutionary Computation* 11(2). 265–286.

Oudeyer, P.-Y. & L. B. Smith. 2016. How Evolution May Work Through Curiosity-Driven Developmental Process. *Topics in Cognitive Science* 8(2). 492–502.

Perone, S. & V. Simmering. 2017. Applications of Dynamic Systems Theory to Cognition and Development: New Frontiers. *Advances in Child Development and Behavior* 52. 43–80.

Pinker, S. 1984. *Language Learnability and Language Development.* Cambridge: Cambridge University Press.

Plunkett, K., J.-F. Hu & L. B. Cohen. 2008. Labels can override perceptual categories in early infancy. *Cognition* 106(2). 665–681.

Radford, A. 1988. Small children's small clauses. *Transactions of the Philological Society* 86(1). 1–43.

Ramchand, G. & P. Svenonius. 2014. Deriving the functional hierarchy. *Language Sciences* 46. 152–174.

Ratter, B. M. 2012. Complexity and Emergence: key concepts in non-linear dynamic systems. In M. Glaser, G. Krause, B. Ratter & M. Welp (eds.), *Human-Nature Interactions in the Anthropocene*, 108–122. New York: Routledge.

Rice, K. & P. Avery. 1995. Variability in a Deterministic Model of Language Acquisition: A Theory of Segmental Elaboration. *Phonological acquisition and phonological theory* 23–42.

Rickles, D., P. Hawe & A. Shiell. 2007. A simple guide to chaos and complexity. *Journal of Epidemiology & Community Health* 61(11). 933–937.

Riley, M. A. & J. G. Holden. 2012. Dynamics of Cognition. *Wiley Interdisciplinary Reviews: Cognitive Science* 3(6). 593–606.

Ritter, E. 1991. Two Functional Categories in Noun Phrases: Evidence from Modern Hebrew. In S. Rothstein (ed.), *Perspectives on phrase structure: Heads and licensing*, 37–62. Leiden: Brill.

Ritter, E. & M. Wiltschko. 2014. The composition of INFL. *Natural Language & Linguistic Theory* 32(4). 1331–1386.

Rizzi, L. 2021. Cartographic structures and the growth of trees in language development. Talk presented at *Zoom na Cartografia* (LaCaSa/UNICAMP and NEG/UFSC), 25 November.

Roberts, I. 2019. *Parameter Hierarchies and Universal Grammar*. Oxford: Oxford University Press.

Roberts, I. 2021. *Diachronic Syntax*. Oxford: Oxford University Press 2nd edn.

Roberts, I. & A. Roussou. 2003. *Syntactic Change: A Minimalist Approach to Grammaticalization*. Cambridge: Cambridge University Press.

Samuels, B. D. 2009. *The Structure of Phonological Theory*: Harvard University dissertation.

Schöner, G. 2012. Dynamical Systems Approaches to Cognition. In R. Sun (ed.), *The Cambridge Handbook of Computational Psychology*, 101–126. Cambridge: Cambridge University Press.

Shapiro, L. A. 2013. Dynamics and Cognition. *Minds and Machines* 23(3). 353–375.

Siegelmann, H. T. & S. Fishman. 1998. Analog computation with dynamical systems. *Physica D: Nonlinear Phenomena* 120(1-2). 214–235.

Smolensky, P. 1988. On the proper treatment of connectionism. *Behavioral and Brain Sciences* 11(1). 1–23.

Song, C. 2019. *On the Formal Flexibility of Syntactic Categories*: University of Cambridge dissertation.

Tabor, W. 2009. A dynamical systems perspective on the relationship between symbolic and non-symbolic computation. *Cognitive Neurodynamics* 3(4). 415–427.

Thelen, E. & L. B. Smith. 1996. *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge, MA: MIT Press.

Thelen, E. & L. B. Smith. 2007. Dynamic Systems Theories. In R. M. Lerner (ed.), *Handbook of Child Psychology*, vol. I. Theoretical Models of Human Development, Hoboken, New Jersey: John Wiley & Sons Inc.

Touboul, J. & A. Destexhe. 2010. Can Power-Law Scaling and Neuronal Avalanches Arise from Stochastic Dynamics? *PloS ONE* 5(2). e8982.

Tucker, M. & K. Hirsh-Pasek. 1993. Systems and Language: Implications for Acquisition. In L. B. Smith & E. Thelen (eds.), *A Dynamic Systems Approach to Development: Applications*, 359–384. Cambridge, MA: MIT Press.

Turner, J. R. & R. Baker. 2020. Just doing the do: A case study testing creativity and innovative processes as complex adaptive systems. *New Horizons in Adult Education and Human Resource Development* 32(2). 40–61.

Uriagereka, J. 2012. *Spell-Out and the Minimalist Program*. Oxford: Oxford University Press.

Wagner, M., A. I. Zurita & S. Zhang. 2021. Parsing speech for grouping and prominence, and the typology of rhythm. *Proc. Interspeech 2021* 2656–2660.

Waldrop, M. M. 1993. *Complexity: The Emerging Science at the Edge of Order and Chaos*. New York: Simon and Schuster.

Warren, K., C. Franklin & C. L. Streeter. 1998. New Directions in Systems Theory: Chaos and Complexity. *Social Work* 43(4). 357–372.

Watumull, J., M. D. Hauser, I. G. Roberts & N. Hornstein. 2014. On recursion. *Frontiers in Psychology* 4. 1017.

Wexler, K. 1998. Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua* 106(1-4). 23–79.

Wiltschko, M. 2014. *The Universal Structure of Categories: Towards a Formal Typology*. Cambridge: Cambridge University Press.

Wiltschko, M. 2021. *The Grammar of Interactional Language*. Cambridge: Cambridge University Press.

Xu, F. & J. B. Tenenbaum. 2007. Word learning as Bayesian inference. *Psychological Review* 114(2). 245.

Núria Bosch
University of Cambridge
`nb611@cam.ac.uk`