1 Introduction: Emergence and World Englishes

How different are World Englishes (WEs) from their metropolitan donor varieties and from each other, and why have these differences evolved and become firmly established? Indirectly, this question touches upon even broader, more fundamental ones: how are language varieties related to each other, and how does Language "work"? In the early phase of the discipline of WEs (and still in some branches of linguistics), languages and language varieties were conceptualized as discrete, separate entities. More recent thinking has tended to downplay discrete nation-based distinctions and to highlight mutual interactions, contact-induced changes, and linguistic forms as floating resources (Blommaert 2010; see Section 4.1). Extending this line of thinking significantly, in this Element, I am proposing a novel approach and explanation, arguing that languages and language varieties are best understood as perpetually fluctuating, highly complex interactions between a large number of agents and attributes, best to be accounted for by the scientific theory of Complex Dynamic Systems.

English is the world's leading language today by far – this is simply a trivial statement. In addition to being used as the default choice in transnational and "lingua franca" communication of whatever kind (in international politics, business, academia, tourism, etc.), it has come to be rooted as national, official, semi-official or practically most widely used language in about 100 countries all around the globe (for lists, see McArthur 1998, 2002; for a map, see Schneider 2020a:64). In these countries, stable varieties with distinct properties of their own (on the levels of phonology, lexis, grammar, and pragmatics) have emerged (such as Canadian, Singaporean, Indian, Nigerian, Ghanaian, or Australian English). The label "World Englishes" has been widely accepted as a summary designation of both these varieties and the scholarly discipline that intensely investigates their structural properties and sociopolitical settings. Most of these varieties are "Postcolonial Englishes" (see Schneider 2007), products of British (and rarely American) colonial expansion since the seventeenth century, including countries where the majority of speakers are native or first-language (L1) speakers of English, such as the United States or New Zealand, perceived as (modified) continuations of erstwhile British English (BrE). In many other former colonies, national varieties, which have often been called "New Englishes" (NEs), have evolved – second-language (L2) forms of English in multilingual contexts created by the process of colonial linguistic re-rooting under language contact conditions. In recent decades, however, globalization and the "transnational attraction" of English (Schneider 2014) have produced distinct forms (and usage contexts) of English also in countries with no British colonial background, such as Korea

(Rüdiger 2019), the Netherlands (Edwards 2016), or Namibia (Schröder 2021), producing, consequently, "Korean English," "Dutch English," "Namibian English,"[1] and others, respectively. In general, the distinction between postcolonial and nonpostcolonial varieties has been downplayed increasingly, and many similarities have been identified (Buschfeld & Kautzsch 2017, 2020). Since roughly the 1980s, the discipline of WEs, investigating and describing these varieties, has grown tremendously; access to it and survey representations can be found in many useful textbooks (Mesthrie & Bhatt 2008; Schneider 2020a), handbooks (Filppula, Klemola & Sharma 2017; Nelson, Proshina & Davis 2020; Schreier, Hundt & Schneider 2020), and other resources.

World Englishes are conventionally seen as distinct linguistic systems, separate from though related to their respective "parent/donor varieties." They are commonly regarded as "daughter varieties" of (mostly) British or (occasionally) American English (AmE), similar in some respects but distinct in others, with specific words, pronunciation habits, and grammatical patterns of their own, which are unique to certain regional or national varieties and contribute to constituting their independent systemic status. The assessment of these distinctive, innovative features is often a topic of quite some controversy, especially in local contexts where for historical and social reasons the erstwhile donor variety (typically BrE) is still considered to be the only "pure" and "correct" one, to be striven for in language teaching (see Schneider 2023). Consequently, when judging innovative structural properties of these varieties, utterances are commonly classified as "grammatical"/"acceptable" or not, or, in prescriptive and often educational contexts, "correct" or "wrong." There are many discussions in WEs and applied linguistics as to whether or when structurally unusual phenomena observed in NEs are errors or emerging properties, to be recognized as such (see Mukherjee & Hundt 2011).

This categorial line of thinking on different language varieties, viewing them as discrete and distinct entities[2] and their features as locally established and acceptable or not, is deeply engrained in linguistic thought and theorizing, and results from the dominant early phases and schools of linguistic theory development. In leading grammatical theories, languages are conventionally conceived of as discrete systems of interrelated rules or units and their mutual relations. To illustrate this and put it in context (even if this may be familiar to

[1] Obviously, such labels need to be problematized somewhat, since they posit an essentialist, entity-like understanding of varieties. For a discussion of the distinction between "X English" and "English in X," see Schneider (2007:50).

[2] Larsen-Freeman (2018b:82) labels this perspective "stasis – viewing it as a product, not a dynamic process." In contrast, she argues (and I fully agree) that "language is an inherently malleable, non-teleological system," and "learning a language is not about conformity to uniformity" (84).

many readers), let me take a short look at the theories that have dominated our understanding of "how language works" up until and through the twentieth century (and into the recent past), a short detour with a concise survey of established linguistic thinking, as it were.

Looking into and analyzing modern living languages began in early modernity (roughly the sixteenth and seventeenth centuries), when educated scholars in Europe came to apply the categories of ancient Greek and Latin, the only type of language analysis familiar to them, to living languages like English, thus establishing a line of thinking that came to be called "traditional grammar." It has fundamentally shaped traditions and concepts of language teaching in England, the western world, and elsewhere, and continues to be strongly influential to the present day. Educated language students will be familiar with many of its core concepts: Words are classified into "parts of speech" (word classes such as nouns, verbs, adjectives, etc.), and semantically interpreted grammatical categories (such as number, person, gender, or tense) are attached to these words, expressed by specific inflectional endings (e.g. *-s* for noun plurals, *-ed* for verbal past, *-er* for adjective comparatives, etc.). The modern thought that each language has a structure of its own that needs to be understood in its own right was unfamiliar to the proponents of these descriptions, who regarded Latin as the "ideal" language, sanctified by tradition. However, this scheme does not work well for modern English, a language that has changed from an erstwhile strongly synthetic type (with many inflectional endings in the Old English period) to a largely analytic language, with very few endings left and word class assignments being notoriously fuzzy and variable. Younger, internally important distinctive properties of modern English, which have evolved only in the last few centuries (such as aspect, word order, or the special role of an operator in the verb phrase; see Quirk et al. 1985) and which Latin does not have, were given short shrift and disregarded in school teaching. The attitude associated with this scheme, inherited from days of old, is thoroughly prescriptive – so unusual patterns (including innovative ones in NEs) are typically branded simply as "wrong" and "errors" to be avoided.

As is well known, modern linguistics as a scientific discipline started with structuralism, originated by Saussure (1916), and was then implemented in several regional schools in the first half of the twentieth century. In this framework, language is regarded as a purely synchronic system consisting of abstracted units of various kinds and levels (phonemes, morphemes, lexemes, etc.) that enter category relations through functional equivalence ("paradigmatic relations" of choice: *This animal is a dog/cat/horse/. . .*), delimitate each other by contrast (*Willow is a cat.* [and hence not a dog/horse/. . .]), and build higher-order entities (like phrases or clauses) in chain-like syntagmatic

relationships (*I love [my fluffy old cat]*). Language analysis operates mainly through inductively abstracting specific sequences of units categorized into paradigmatic classes. Hence, a sequence of symbols such as "Det + Adj + N" describes a prototypical noun phrase pattern (*my good friend*, *these lovely flowers*, . . .). This type of structural approach is observation-based (as to how words behave, in which contexts and patterns they occur) and inductive (starting out from multiple empirical observations and then generalizing from them), and hence "objective," not based on preliminary dispositions. Thus, it is practically useful and has come to be influential, in particular methodologically (e.g. in corpus linguistics, where the behavior of words in context can be objectively observed and quantified). However, one obvious question is whether its purely surface-oriented, categorizing, and compositional approach is not somewhat reductive, whether it is sufficient for understanding how language works. Its rather mechanical procedure fails to recognize surface ambiguities, underlying relations and movements (such as corresponding active–passive structures), schematic structures, and the like.

Generative (transformational) grammar, originally the brainchild of Noam Chomsky, has become a strongly dominant discipline in linguistics, and since his first book in 1957 has gone through various developmental stages (such as Principles-and-Parameters, Move α, or Minimalism). Basically, it offers a highly abstract account of the human language faculty and is not primarily geared toward recognizing any language's distinctive traits. Its central goal is modeling "competence," the language knowledge of an "ideal speaker-listener, in a completely homogeneous speech community," a being that simply does not exist in reality. Actual linguistic behavior, "performance," is rejected as irrelevant (possibly containing errors and random deviations), and the approach is devoid of all interest in social and pragmatic contexts and variability (which has inspired the growth of sociolinguistics, pragmatics, and discourse analysis as more realistically grounded counter-tendencies since the 1960s). Essentially, generativism is a highly abstract rule-rewriting exercise – intellectually brilliant, perhaps, but with limited grounding in actual empirical observations of human behavior, constrained by axiomatic assumptions that deliberately exclude a huge amount of data, observations, and contextual conditions.

As to unexpected, "deviant" observations, structuralism would simply add further classification schemes to the existing set of possible patterns without evaluation, and in deductive generativism rules would have to be rewritten to account for ("generate") observed options. Both schools, however, and with them much of modern linguistics (with the exception of more recent theories as described in Section 3) build upon a notion of clear-cut entities, distinctions, and relations that deterministically describe what is possible or not in matters

linguistic. Similarly to traditional grammar, clear-cut boundaries as to what exists and is possible in a language or not are assumed to exist.

Such deterministic and categorial thinking meets with its limitations quite often in communicative and attitudinal reality, however. "All grammars leak," as Sapir (1921:39) observed a century ago: All rules have exceptions, with or without a specific discourse function; grammars allow "multiple analyses" of the same surface structures and often show gradience in the assignment of forms to analytical categories (Quirk et al. 1985), and so on. Similarly for the relationship between varieties: They show similarities and differences, they overlap in some respects and with some forms but not in others (see Section 4 for an application to WEs). Prescriptive thinking that posits clear-cut rules and distinctions between "right" and "wrong" may be pedagogically helpful to some extent and certainly helps to uphold conservative social divisions in societies (between those with and those without access to higher education, where these rules and myths are perpetuated); but it does not conform to the realities of linguistic behavior: It has often been argued and shown that "usage guides" prescribe rules that are not widely observed in natural conversations. Adhering to simple rules may conform to a deeply human desire for clarity, conformity, and guidance in a challenging world, and result from cognitive principles favoring clear-cut categorizations and binary divisions, but, regrettably, that is not how in most spheres of life reality works. Language is not mathematics.[3] Language, and life, for that matter, are inherently fuzzy.

Similar trends toward the dissolution of once clearly established realities have marked developments in many other disciplines, including the sciences. Take physics as a case in point: Newton's beautifully deterministic equations seemed to have offered a solution to all puzzles regarding the revolution of planets in orbit and movements of objects in time and space, until Einstein's relativity came and overthrew it all. And that is not to talk of Heisenberg's indeterminacy principle or the insights of quantum physics. Nature and life are simply immensely complex processes. And it is time to recognize that this applies to language as well, and to ask which insights may result from this. Furthermore, the natural and social sciences do provide us with insights and theoretical frames that may also serve to advance our understanding of how languages, including WEs, "work."

Consequently, in this Element I advocate an alternative approach to theorizing on the nature of linguistic evolution, the emergence of language varieties, and I apply it specifically to the domain of WEs. I suggest that language and

[3] See Chambers (2003:26–38) for a similar criticism of the "categorial" approach from the perspective of modern variationist sociolinguistics.

language varieties should be integrated into and highlighted from the scientific sphere of Complex Dynamic Systems (CDSs). This is a scientific framework, a meta-theory, that has been growing strongly over the last few decades and that has been immensely successful in explaining many processes, properties, and relations in many natural and social sciences. Complex Dynamic Systems are systemic relations that, as the name suggests, are complex (i.e. consist of many locally interacting agents and components) and dynamic (i.e. continuously evolving in time and space – in addition to many other, more specific properties; see Section 2.2). They produce networks of higher-order interrelationships, and they are self-organizing, bringing about novel functional entities without a steering authority. I presuppose the basic assumption (standing behind this Element, and argued for by other linguists as shown in Section 2.3) that precisely these qualities characterize language and the relationship between language varieties as well, and that "Language," the human language faculty, and existing languages, constitute another important domain of life shaped by and manifesting CDS principles. Consequently, I argue that English(es), the sum total of all manifestations of this particular language, with no need for clear-cut boundaries, constitute a broad, important, multilayered domain (or fraction) within this CDS of Language, and that WEs, in turn, are components within such an overarching complex-systemic context.

In the next section, I introduce the approach and its theoretical framework, and, most importantly, I will discuss the properties that have been suggested as characteristic of CDSs. I believe it is necessary to introduce these principles and properties first, and independently, but I will provide cross-references to language whenever possible and offer sample applications to properties of English to build the connection from the outset. There have been earlier suggestions to apply this line of thinking to language studies, and I will briefly survey and discuss these (especially in Section 2.3). Many of the CDS principles, and some of these earlier language-related suggestions, are rather abstract in nature, and one core question is how these principles manifest themselves in linguistic evolution in everyday practice, on the ground, as it were. I believe that the school of usage-based linguistics, in association with construction grammar as a mode of formally modeling some of these interrelationships, offers a superb explanation of how emergence and evolution happen (see Hoffmann 2022:2.1), and so I will discuss these in some detail in Section 3. In Section 4, I will zoom in on the domain of WEs more specifically and will provide many illustrative examples showcasing select properties of WEs, which highlight how specific principles of CDSs manifest themselves in structural reality (with these principles as organizing guideline, similar to Section 2.2). And finally, in Section 5, I briefly illustrate how agent-based modeling, a procedure that is routinely

employed to simulate and understand CDSs, can be applied to model the emergence of novel features of WEs, using the freeware programming environment NetLogo.

2 The Theory of Complex Dynamic Systems

2.1 Origins, Definitions, and Applications

Compared to other branches of the sciences, the theory of CDSs is relatively young: It can be traced back to precursor disciplines from the earlier twentieth century and to the work of some individual scholars, notably mathematicians, who laid foundations even earlier (see Larsen-Freeman 2017:12–13), but as a coherent framework it has evolved strongly and become influential in various subdisciplines of the sciences since roughly the 1960s (for a historical survey, see Mobus & Kalton 2015:32–40). It has been established and is growing vigorously in the natural and social sciences, with a wide range of successful applications in biology, medicine, social organization, natural phenomena, technology, businesses, and other disciplines. It captures basic properties of many domains in life marked by complex systemic relationships that are perpetually developing, with many agents and units that are interconnected, exchange information (in a very general sense) just locally, without any overarching steering authority, and thereby build something larger (so that the whole is more powerful than the sum total of its parts, a consequence of nonlinearity). The theory of CDSs argues against reductionism and categorial, deterministic thinking and highlights the productive power of interacting multi-agent systems without central control. It is a "holistic," "cross-disciplinary" approach, though at this stage not yet a fully-fledged theory with a "canonical" form but rather a "meta-science" (Mobus & Kalton 2015:3), a "transdisciplinary" "metatheory" (Larsen-Freeman 2017:14, 21).

Terminology is still somewhat unstable – the framework comes under varying labels such as systems science, complexity science, or theory of complex adaptive systems. The latter has been widely adopted when applied to language, but I prefer "CDS" because this term captures two of the most central properties of such systems (and "adapt" to me implies a given target [to what?], while here we witness a continuous process of different components evolving through time, influencing each other mutually).[4] By now the approach has been widely accepted and applied in the natural and social sciences, and established as a major all-encompassing research branch in itself, with powerful institutions devoted to developing it further (e.g. Max Planck Institutes for Dynamics and

[4] For a discussion of terminological issues with the same outcome, see de Bot (2017).

Self-Organization, for Intelligent Systems, or for the Dynamics of Complex Technical Systems). Its leading, longest-standing global research center is the Santa Fe Institute in New Mexico.[5] It offers a wide range of activities and resources (see complexityexplorer.org), including many courses on specific CDS-related topics (like Rand 2024 on agent-based modeling). The Institute promotes the CDS theory through several publications, for example Mitchell (2021; a recommendable, slightly technical introduction), or a lecture by its president on the epistemology of complexity science (Krakauer 2023). Other introductions and surveys that I have found useful include Bossomaier and Green (2000), Johnson (2009), Holland (2014), Mobus & Kalton (2015), and Jensen (2022). Researchers into CDSs have investigated the nature of developmental processes and outcomes in many domains of life, including ant colonies and many other life and ecology associations (e.g. tropical rainforests), the human brain (Bossomaier 2000), the economy (including markets), social networks (Buchanan 2003), or the growth of cities, to name but a few. The discipline's main tools are computational modeling and simulations (Mitchell 2021; Rand 2024; see Section 5).

Conceptual basics and processual explanations of CDS have their roots in nonlinear equations in mathematics described by chaos theory (e.g. Gleick 1987), one of its major precursor frameworks. Nonlinear equations have distinctive and apparently "strange" properties, such as extreme sensitivity to (minimal variation in) initial conditions, which may lead to qualitatively different leaps and outcomes, and the fact that after very many iterative, recursive runs of the equation the resulting phase model tends to approximate self-similar "attractor" states, something like order. Chaos theory is mostly seen as a core branch, one particular type, of CDSs (see, e.g., Guastello & Liebovitch 2009; Johnson 2009:ch. 3; Mitchell 2021:units 2.1, 2.6, 2.7). The core property of nonlinearity is usually taken to be shared (Mufwene et al. 2017:2; Larsen-Freeman 2018a:52) or even central to both disciplines: "Non-linear behaviour is one of the cornerstones of complexity" (Bossomaier & Green 2000:7; see also Mobus & Kalton 2015:202, 251–2, 592). Some central properties (e.g. the interrelationship of order and chaos, or bifurcation) derive from chaos theory.

Complex Dynamic Systems are characterized by a range of distinctive core properties, with the most important and relevant of them to be surveyed and contextualized in the next few sections. In each case CDS properties will be defined and described in their own right, but these descriptions will be followed by brief language-related observations or examples from English, to imply the claim (and hopefully make it compelling) that both are inherently related.

[5] www.santafe.edu/.

Section 4 will then be devoted to showcasing the application of these principles to WEs in greater breadth and detail.

2.2 Properties of Complex Dynamic Systems – and Effects in Language

While there is no generally accepted finite listing of the distinctive properties of CDSs in the discipline, and different authors define the notion slightly differently and highlight varying aspects, there is a basic consensus on some core components that constitute and characterize CDSs. I subdivide them into and present them as eight different central properties that encompass the most important qualities under discussion. In each case, I circumscribe the property in question as presented by CDS theorists before I consider some reflections and voices on its applicability to language in general and English in particular.

2.2.1 Systemness

In CDS theory, "a system is a whole of some sort made up of interacting or interdependent elements or components integrally related among themselves" (Mobus & Kalton 2015:73–74). This seems a rather trivial notion, but of course it constitutes a fundamental prerequisite for complexity to emerge. Systems are to be understood by knowing the units they are built of (and possibly relevant associated attributes) and the relations between them.

It is self-evident and axiomatic in linguistics that this holds for language as well, that languages are systems. The idea was introduced by Saussure (1916), the founding father of structuralism, who argued that language is a "system of signs" (and signs, in turn, are defined as pairings of *signifié* and *signifiant*, meaning and formal expression). Systemic relations mean that synchronically, at a given point in time and in a single language, units on several levels (phonemes, morphemes, lexemes) enter mutual relations of various kinds. Fundamentally, these relationships are either syntagmatic or paradigmatic. In syntagmatic ("chain"-type) relations items follow each other and jointly build a higher-order unit. For example, sounds build words (/m/+/æ/+/n/ > *man*), words build phrases (*the*+*old*+*man* is a prototypical noun phrase), and phrases build clauses ([*The old man*]$_{NP}$ [*is writing*]$_{VP}$ [*a letter*]$_{NP}$.). Paradigmatic ("choice"-type) relations hold virtually between units that can be exchanged for each other (can undergo "substitution" and stand in "contrast" to each other): they can fill the same slot, they function equivalently (and thus functionally defined constituent classes can be identified). For example, in the word *man*, the final consonant can be substituted by /t/, yielding another word, *mat*, and in the clause the object noun phrase slot after *writing* could also be filled by a single