A farewell to UG and a welcome to CEG

Hubert Haider
Dept. of Linguistics & Centre for Cognitive Neuroscience, Univ. Salzburg

Abstract

It is essential for any scientific theory to demonstrate its predictive power, that is, to generate testable, subsequently empirically confirmed predictions and explanations. In this paper, predictions and explanations of the UG theory are contrasted with the theory of the cognitive evolution of grammars (CEG). For ten areas of grammar theory, it is shown how the two theories differ and why CEG performs better.

The hypothesis of an innate universal grammar (UG) turns out as underpowered and overburdened. The competing hypothesis presented in this paper – grammars as target and products of evolution on the level of cognitive systems – is well in line with the empirical situations of diachronic grammar change, cross-linguistic (in)variants, language acquisition, and system evolution in general. The theory of the cognitive evolution of grammars (CEG) proves its worth and avoids the notorious quandaries of UG. Moreover, it connects grammar theory to the most successful theory of system development and change, namely the theory of evolution, in a field of application outside of genetics, namely neuro-cognition.

1. Background

UG is an elusive subject because its content has always been adapted to the particular generation of model of grammar theory of which it is an axiomatic part. An authorized description reads as follows: "*The theory of the genetically based language faculty is called Universal Grammar; the theory of each individual language is called its Generative Grammar*." (Chomsky 2017: 3). Neither 'genetically based' nor 'language faculty' contribute much to a satisfactory explication. Geneticists have not detected "the genetically based language faculty" in the genome and Generativist linguists have more and more refrained from presenting testable versions of UG.

That humans are equipped with cognitive capacities for language processing in perception, production, and acquisition is undisputed and trivial. We experience it daily. The disputed issue is the claim that UG is the theory of these capacities and that it rests on a genetically coded blueprint for something whose implementation nevertheless "varies radically", given that "*languages appear to be extremely complex, varying radically from one another*." (Chomsky 2017: 2). UG is a much stronger claim than the traditional idea that our 'talent' for language is part of our human nature and thereby species-specific and ultimately somehow genetically conditioned, and it is empirically inadequate, as will be made evident in this paper.

The hypothesised existence of *innate* and homo-sapiens-wide uniform principles of the organization of human grammars is not so much an empirically confirmed discovery than a theoretical postulation. On the one hand, it is the essential 'antidote' for an induction gap. The very Generative grammar of her/his language that an L1 learner is supposed to acquire easily is underdetermined by the input. On the other hand, UG is meant to account for cross-linguistic variants of grammar systems. These two objectives are obviously antagonistic. Innate constraints that enhance acquisition should narrow the system space whereas cross-linguistic variance requires widening it. The more cross-linguistic variability is admitted by UG, the higher will be the effort of identifying the respective L1 grammar in the system space circumscribed by UG during grammar acquisition.

"Innate" is a multiply ambiguous notion, as Scholz & Pullum (2002: 189-191, 2006: 66) explicate. For capturing the linguistic reality, an appropriate reading of innateness seems to be this. It is agreed that there is a non-negligible set of grammatical properties that cannot be acquired bottom-up, that is data-driven. This set is not inferred ('learnt') from the input. Children would not be able to unerringly identify it with their general problem-solving abilities since linguists have not been spontaneously able to do so either.

At this point, the Generativist 'short circuit' took place. "*Not acquired bottom-up*" has been equated with "*innate domain-specific capacity"* and this has been equated with "*genetically coded domain-specific program"*. This is a precarious way of tackling the problem since it misjudges the real thing. The real thing is the fact that (mental) grammars are put to use as part of our human cognitive system. They run on a neuro-cognitive 'computational platform', which is a network of computational routines of our brain recruited for language processing. This is the locus of innate properties. The innate routines are the processing routines recruited for language processing. Their evolutionary history goes back further than the emergence of the human language abilities. We get an idea of their ways of working by studying how they constrain the structures of our languages. This is an essential part of the apparently a-priori qualities.

Second, in complex dynamic systems, the various components of a system interact in complex ways. What we observe as properties of the output are on the one hand *inherent* properties of the system and on the other hand, *emergent* properties that are necessary consequences of the ways the parts of the system operate and interact. The emergent properties are results of the processual characteristics of the system. The inherently determined ones may be regarded as pre-programmed. However, the 'programmer' is – just as in biological evolution – the ongoing processes of evolution operating on cognitive programmes. Grammars are products of evolutionary processes, namely the multi-millennial processes of the cognitive evolution of grammars, as explicated in Haider (2021a,b).

The cognitive capacities for language, with grammar as its core part, are species-specific because only our species has diverse and powerful enough, general cognitive abilities that make it possible to "subcontract" parts of these abilities and link them into a network that we may call – in modern parlance – a 'language app'. This app reveals properties of genetically-based capacities (due to the genetically determined general cognitive routines involved), but the app itself is not genetically coded although it is domain-specific. It is compiled (in the technical sense) and in its present structure and function a result of CEG, the *cognitive evolution of grammar*, that is, a result of variation & selection operating on self-replicating systems (viz. the respective grammars).

It is an important but too often overlooked aspect in the UG debate that grammars are learnable because variants of grammars that are difficult to acquire are immediately sieved out in the course of transmission from one generation to the other. This sounds trivial but it is the solution to the learnability puzzle that UG is meant to solve. Since the onset more than 300.000 years ago,[1] each grammar is in continuity with a preceding and a following grammar variant, each of

[1] As excavations in Jebel Irhoud have revealed (Richter et al. 2017), homo sapiens already settled in North Africa 300.000 years ago. Florisbad (Berger & Hawks 2023) and Omo Kibish fossils (Vidal 2022) are dated 260.000 and 230.000 years back, respectively.

which must have been able to be acquired and transmitted. This process guarantees that grammars are ‘purified’ of unlearnable or less-easily learnable variants. Unlearnable properties are trivially filtered out and preference is given to variants that are ‘easier’ to handle than others by the various processing routines. ‘Easier’ is the cover term for better adapted features of grammars. Economy conditions[2] are part of it.

The selection mechanism that is constantly at work is the learners’ brain itself and in particular the routines recruited for the handling of language (Haider 1999: 218). The inevitable outcome is evolution by natural selection (plus the analogues of genetic drift and gene flow; see Haider 2021a: 21-22) with a resultant adaptation to the selecting environment, that is, to the demands of the ensemble of processing routines of our brain (rather than alleged communicative desires of language users, as functionalists would have it).

Schoenemann (2012: 443) concludes that "*overall, language appears to have adapted to the human brain more so than the reverse*" and Kirby (2001: 110) underscores that “*we may need to concentrate less on the way in which we as a species have adapted to the task of using language and more at the ways in which languages adapt to being better passed on by us*.” As known from biological evolution, the process of variation and selection is able to produce highly complex self-replicative systems and one of them is the family of (neuro-)cognitively represented grammar systems involved in the handling of language (Haider 2015a, 2021a,b).

A negative trait of the UG hypothesis has been its low productivity when it comes to generate testable “stunning predictions”[3] that are empirically confirmed. On the one hand, UG provides so many degrees of freedom that there is (too much) room for auxiliary hypotheses to be adduced for blocking counterevidence. For historical reasons, UG is closely tailored to the characteristics of [S[VO]] languages like English. Consequently, other systems such as SOV or so-called free-word-order languages (= T3 languages) have been perceived as deviations requiring reintegration by means of auxiliary assumptions, but with the essential question unanswered: Why is there so much cross-linguistic variation with additional grammatical ado if a much simpler, UG-streamlined SVO grammar would suffice? The appropriate answer is that this is exactly the picture one expects if evolutionary processes are at work, but not if there existed UG as a master-grammar format. Evolution is dissipative since it is fed by random variation. UG-compatible grammar variation, however, is expected to be highly restricted and conservative.

In each case – the *innateness* claim and the cross-linguistic *universality* claim – the sustaining moment has been the absence of a superior theory. There are opponent positions but they are equally underpowered in their predictive capacity. This is a type of situation that philosophers of science call a crisis. There is an acknowledged problem, there are many alternative and incompatible attempts of overcoming it, without any broadly recognized success though. That the UG conjecture can be upheld until today is not so much a fact about human insufficiency but

---

[2] ‘Economy’ comprises the cost-benefit relations of storage, retrieval, reception, production, and acquisition. As Collins (2022) emphasizes that economy conditions do not play a role in UG, but they play a role in evolution in general and in the evolution of grammars in particular.

[3] „*The hallmark of empirical progress is not trivial verifications.* [...] *What really counts are* […] *unexpected, stunning predictions: a few of them are enough to tilt the balance.*“ Lakatos (1978: 6).

an excellent example of scientific conservativism. Scientists do not dump an unproductive theory if the alternative is a situation without any theory at all.[4] The longer a theory has been retained and bolstered, the stouter it tends to be defended because loss aversion is a strong cognitive bias.

A delaying feature in the whole debate has been the all-or-nothing attitude of the confrontation. For Generativists, potentially, *all computational principles and processes* of grammars are directly or indirectly dependent on an innate UG; for opponents (e.g. proponents of construction-grammar and other functionalist grammar theories) no *grammatical principle* is innate.[5] Altogether, the controversy is unrewarding since the crucial point is not *innate* vs. *not-innate* but the exploration and explication of the conditions of efficient grammar acquisition and use.

Opponents tend to overlook an evident fact. If language acquisition were merely a facet of the general problem-solving capacities of children (and adults), then children would be in about the same position as linguists, except for their less effective cognitive abilities. Both parties try to arrive at the grammar of a language by data analysis (rather than genetic guidance). Linguists still struggle while children succeed. Half a century of dedicated research in grammar theory has not produced an accepted compendium of UG or a comprehensive grammar based on it.[6] How would children individually manage a task without instruction with much less elaborate problem-solving capacities than teams of professional linguists within ten years? Patently, general problem solving does not seem to be the whole story either.

The present situation is one of reciprocal denial without alternatives. Opponents of the UG conjecture deny not only the existence of an innate "language acquisition device", aka innate UG, as guide through the grammatical mazes; they typically deny any domain-specific advantage of language learners. But on the other hand, they are unable to present a compelling demonstration based on empirical evidence that the grammar of a ten-year-old is the result of domain-general problem-solving capacities. After all, problem solving is known to be much more dependent on general intelligence than language acquisition. Language acquisition does not co-vary with general intelligence. Facts show that double dissociations between the two capacities are possible and attested.[7]

The fact that this debate has remained inconclusive over a period of several decades shows that neither position has been able to prevail in the scientific debate and that there must be a reason for this stalemate situation. The reason is that arguably neither position is right. During the same period, an already well-established line of research has flourished and continued to develop one of the most successful scientific theories addressing the same type of problems that linguists deal with, namely the problem of understanding how complex self-replicating systems emerge, how they are represented, how they change over time, why they change, and why some thrive and others disappear. This field is the modern synthesis of Darwin's theory of evolution and in particular the subfield of population genetics. This field is a valuable source of insights for

---

[4] "*Contrary to naive falsificationism, no experiment, experimental report, observation statement or well-corroborated low-level falsifying hypothesis alone can lead to falsification. There is no falsification before the emergence of a better theory.*" (Lakatos 1970: 119).

[5] "[...] c*onstructionist approaches do not rely on innate universal principles*" Goldberg (2013: 23).

[6] The first and last endeavour has been Stockwell & Schachter & Hall-Partee (1973) exactly fifty years ago.

[7] SLI is a dissociation between a typically developed general intelligence and impaired language acquisition. In Williams-Beuren Syndrome, grammar development by far exceeds impaired non-verbal cognitive abilities.

linguists, too. First of all, it is a law of nature that complex systems do not materialize in a flash.[8] Such systems develop over long periods of time, and they typically develop in a variation & selection process, aka evolution. "*The theory of evolution by cumulative natural selection is the only theory we know of that is, in principle, capable of explaining the existence of organized complexity*." Dawkins (1991: 317).[9]

It is important to understand that *the kind of evolution* this paper will focus on is neither a variant of the genetically based *biological* evolution nor of a so-called 'mimetically' based *cultural* evolution[10] (Aunger 2001), but plain Darwinian evolution[11] operating on a neuro-cognitively constituted, self-replicating system, see Haider (2021a,b, 2015a). The target of *cognitive evolution* is a neuro-cognitively represented 'programme package', or – in modern parlance – the "language app". This is a *domain specific* programme package in the ensemble of our cognitive computational capacities, assembled of *domain-general*, sub-contracted capacities. The structure and much of the content of this package is a product of cognitive evolution, that is, the present outcome of processes of variation & selection (plus drift and the analogue of gene transfer). The selection environment is the subset of general brain routines[12] recruited for language processing, that is, for the actual usage of the language app. In a nutshell, the evolutionary perspective on grammars is just the general Darwinian evolutionary perspective on dynamic systems, but applied to a neuro-cognitively based system.

Worded in computer terminology, the acquisition of language and grammar is neither a program upload nor an auto-installation process. It is the acquisition of something that has been custom-fitted to our human cognitive capacities in the course of a multi-millennial cognitive evolution of grammars. Custom-fitted language apps are the outcome of the evolutionary moulding exerted by the selection effects of the processing routines of the domain-general neuro-cognitive platform in the human brain on which the domain-specific language app runs. In other words, the specialized software has evolutionarily adapted to the general operating system for language acquisition and use. This guarantees effective acquisition and efficient use, as will be argued below.

2. How much UG?

Over the years, fittingly adjusted conceptions of UG have shadowed the various model generations of Generative Grammar. In the time of the Principles & Parameters model, UG was regarded as a full-fledged master grammar. Individual grammars were seen as particular, close manifestations of this master grammar, with language-specific values for parameter variables provided by UG. Thus, cross-linguistic grammatical variation is captured in terms of parametric

---

[8] "*Essentially, it is that no biologist imagines that complex structures arise in a single step*." (Maynard Smith 1986: 49)

[9] This is not only true of all forms of life but it seems to be true also for the cosmological dynamics of the universe, see Smolin (1992), or applications in computer science, see Holland (1992).Variation plus sieving out (i.e. selection) is the essence of evolution.

[10] Literature on language evolution tends to focus on aspects of *biological* evolution or even miss the point entirely. "*A look at the literature on evolution of language reveals that most of it scarcely even addresses the topic. Instead, it largely offers speculations about the evolution of communication* [...]." (Chomsky 2011: 265.)

[11] Substance-neutral theory of the behavior of self-replicative systems with variation exposed to processes of selection combined with the conserving of traits as main components.

[12] This is crucial. The selection environment is 'hard-ware' based. Functionalist ideas insinuating *communicative pressures* or *demands* are empirically and theoretically unsubstantiated concepts; see Haider (2021b: 110-111).

differences. With more and more cross-linguistic data available and in the absence of a theory of parameters, the continuously growing number of potential parameters eventually amounted to the very same kind of induction-gap dilemma that UG theory was meant to solve, see Section 3.5.

The Minimalist Program (MP), as a complete relaunch, goes hand in hand with efforts of devising a minimalist UG. Chomsky (2007) suggests that UG might be extremely simple and abstract, consisting only of a mechanism for combining symbols, which is called "merge", with recursion as unique feature of the language faculty. Hauser et. al. (2002) make a distinction between properties of the faculty of language (=FL) that are exclusively part of human languages (FL narrow) and general capacities (FL broad). The unique property of narrow FL is supposed to be recursion as a uniquely human capacity. In the meantime, evidence is available showing that other species master recursion as well: "*We reveal that crows have recursive capacities; they perform on par with children and even outperform macaques*. [...] *These results demonstrate that recursive capabilities are not limited to the primate genealogy and may have occurred separately from or before human symbolic competence in different animal taxa*." Liao et al. (2022: 1).

The UG of the MP is truly minimalist and a total about-turn. It sacrifices the original motivation for its postulation, namely the explanation of the induction gap and the explanation of cross-linguistic (in)variants. The minimalist UG is deemed to account for the mere fact that our brains are highly efficient in dimension management. We are able to map back and forth between linear arrays of sounds (one dimensional) and phrase structures (two-dimensional box-in-box configurations).[13] Nevertheless, most Generativists still cling to some sort of old-style UG in order to bridge the embarrassing induction gap.

Unfortunately, this claim has never been empirically validated. Therefore, this theory is in default. Kinsella (2009), at book-length, as well as Kinsella & Marcus (2009) show in great detail on empirical and conceptual grounds why the MP cannot and does not match its programmatic tenets. The monograph by Kinsella (2009) is one big argument that "*shows that the Minimalist Program is not a viable theory of the emergence of language under the criterion of evolvability*" (Karlsson 2010: 49).

Evidently, the minimalist UG theory does not answer the questions that motivated the postulation of the broad version of UG in the P&P version. The Minimalist concept needs to be accompanied by a theory that captures cross-linguistic (in)variants and guarantees the learnability of grammars. This is the theory of CEG. It fully replaces UG and thereby avoids its shortcomings.

3. UG versus CEG

In this section, ten assumptions or deducible predictions of the UG theory will be confronted with corresponding (counter-)positions in the theory of CEG, the cognitive evolution of grammars, see Table 1, which serves as menu for the following discussion. It lists and contrasts ten

[13] Dimension management is an essential part of vision, too. The retinal patterns are two dimensional, the constructed representation of images reconstructs a three-dimensional reality.

issues and discusses why the hypothesis of CEG clearly stands out positively against the hypothesis of an innate UG.

	Universal Grammar	Evolution of grammar [Table 1]
1.	genetically grounded	cognitively grounded
2.	universals by genetics	universals by cognitive evolution
3.	closed system space	open system space
4.	parametric	parameter-free
5.	parameter valuation required	no parameter valuation
6.	mono-genetic grammars	(potentially) poly-genetic grammars
7.	perfection	imperfection
8.	SVO by default	adaptive landscape (T3, SOV, SVO, VSO, ...)
9.	changes = parameter re-valuation	changes = ongoing evolution
10.	complete & consistent grammars	partially incomplete & inconsistent grammars

Table 1: A (non-exhaustive) synopsis of contrasting qualities

3.1 Genetic versus cognitive grounding

UG is supposed to be genetically coded, as Chomsky (2017: 3) has claimed from the beginning although compelling evidence is still missing. The language network involves large areas of the cortex of our brain. So, it is reasonable to expect anyone who claims that something which is innate and brain-power consuming to a high extent to produce evidence based on genetics or at least compelling experimental data in support of innateness. These data are wanting.

FoxP2, which is often invoked as *the* genetic evidence, is definitely not the *grammar* gene, as for instance, Xu et al. (2018: 8799) – "*FOXP2 might be important for anatomical features contributing to derived human traits, including speech and bipedalism*" – or Atkinson et al. (2018: 1424) note: "*An in-depth examination of diverse sets of human genomes argues against a recent selective evolutionary sweep of FOXP2, a gene that was believed to be critical for speech evolution in early hominins.*" Anyway, a single-gene mutation could not account for the ensemble of grammar capacities of humans, see De Boer et als. (2020).

It is amazing in this context to read Haworth et al. (2010: 1112): "*The heritability of general cognitive ability increases significantly and linearly from 41% in childhood (9 years) to 55% in adolescence (12 years) and to 66% in young adulthood (17 years) in a sample of 11 000 pairs of twins from four countries.*" More amazing than the content of the message is the fact that not a single linguist has been involved in the team of specialists from behavioural genetics, genomics, functional genomics, human development, psychology, and psychiatry. Language abilities evidently do not play a role in behavioural genetics.[14] In the past 50 years, genetics has made impressive headway, but linguists have not been part of it, although they ought to have, if an innate capacity such as UG had qualified as a prime target of behavioural genetics. The claims of linguists are rightly ignored by geneticists because they lack empirical substance. Even if it is a fact that many innate capacities[15] are involved in language acquisition and use, the claim of a rich UG is much stronger. It claims that the essentials of the neuro-cognitive grammar system with details of operation are innate and govern the child's build-up of grammar by innately-primed learning.

This is an overly bold claim without the necessary underpinning from a demonstration of the way complex systems come into being. Whenever a system reaches the level of innate capacities, a whole lineage of species shows at least precursors of it, since *genetically* determined complex systems take a *very* long sequence of generations for their evolution. During this time, sub-species develop and bud into new species. For homo sapiens, it is a biological fact that a complex UG cannot be the product of a singular fulguration event[16] in a single[17] species (see De Boer et al. 2020). Whoever boldly claims the opposite deserves no credibility unless (s)he is able to produce substantive and compelling data gained with standard methods of the discipline that studies innateness, namely genetics. Such evidence is missing, and missing evidence is the weak side of UG-based language acquisition geared by innately-primed learning, as Pullum & Scholz (2002) argue. Presently, UG as an explanation of cross-linguistic invariants and effective acquisition of grammar is nourished too much by the *confirmation bias*.[18]

In an evolutionary perspective, the grammar systems we see within our very narrow time horizon of roughly 5k years are the products of ongoing processes of cognitive evolution. This means that they are products of processes of variation & selection (plus drift and the analogue of gene transfer, that is, partial grammar transfer in massive bilingual situations). Variation is fed by the imperfect way of grammar transmission. Each generation acquires the grammar from being exposed to the outputs of their linguistic environment. Selection is the effect of sieving out. The variants that pass selection will have a chance to spread. They will enter more brains and thereby produce more grammar offspring in the following generation. The selector is no mysterious force. It is the ensemble of brain functions recruited for language learning and use. They have their own evolutionary history and their own restrictions and they exert them on the processes of language acquisition and processing. This is what is innate.

In the end, the trivial outcome is that elements of grammar that are too difficult to acquire

[14] "*Behavioural genetics* is the interdisciplinary effort to establish causal links between genes and animal (including human) behavioural traits and neural mechanisms. Methods used include twin studies, quantitative trait mapping by linkage to allelic variants, transgenic animals and targeted gene disruption or silencing." https://www.nature.com/subjects/behavioural-genetics

[15] Categorical perception, for instance, is an innate capacity involved in sound-to-symbol mapping. It is not primate-specific, though. Chinchillas, monkeys, chicken or rats dispose of it, too (Kriengwatana et al. 2015).

[16] Here, Sir Fred Hoyle's famous *tornado-over-a-scrapyard-leaves-behind-a-Boeing-747* argument would hold. This is impossible indeed, and nobody can reasonably deny it.

[17] Unfortunately, we shall never find out how language-talented Neanderthals, Denisovans, or homines floresienses have been.

[18] Adopters of a preferred system believe what they want to believe by favouring information that confirms pre-existing assumptions. As a consequence, they are looking for creative solutions that confirm their beliefs rather than challenge them. This makes them closed to new and more adequate possibilities. (Jermias 2001).

cannot come into being. The evolutionary approach to grammars as products of cognitive evolution, that is, evolution on the level of a cognitively represented system, is grounded in the most successful theory of the dynamics of complex systems, namely the theory of evolution we originally owe to Charles Darwin.

3.2 Universals by genetics vs. universals by cognitive evolution

If UG were the "*genetically based language faculty*", any human grammar would be moulded by it. This is what we should be able to identify as clear-cut cross-linguistically demonstrable invariants aka language universals. Is this a satisfactory characterization of language universals? The answer is negative. Invariants are not clear-cut but fuzzy, and a genetic basis has not been identified,

Proponents are unable to trace a causal chain between an alleged, genetically determined neuro-cognitive property and its phenotypical "expression" as a linguistic property. The neuro-cognitive processes underlying the computing of language are unknown and we do not know their genetic genesis. UG-based 'explanations' are postulates without empirical substance. They appeal to an explanatory background whose existence is merely stipulated. No causal chain is known, no experiment has proven any causal relationship, and the "ghost in the machine" required to monitor and carry out all the checks, probes and covert movements required by UG is a phantom.

Fortunately, there is a more straightforward alternative. What we linguists perceive as cross-linguistic invariants are reflexes of the *selection environment* of grammar systems, namely our neuro-cognitive equipment, and indirectly its limitations in linguistic applications. Crucially, universals are emergent, they are not programmed in. The processing brain evolutionarily shapes grammars by sieving out variants which are less adapted to the demands of the processing system. The major traits of variants that pass selection are the traits we observe cross-linguistically and construe as language universals. In the theory of evolution, this is known as *convergent evolution*.

Here is a biological showcase. Mammals have re-entered the sea as a habitat at several occasions millions of years ago. If animals change their habitat, major factors of selection change and we note the effects of adaptation to the new habitat, ranging from fish-like forms (see whales and dolphins), to transformations of limbs into paddles (see seals), or webbed feet, as in the case of water hounds or otters. So, for example, 50 million years ago, in the course of roughly 15 million years of evolution, the processes of evolution-by-selection have produced mammals that look like fish. This is known as convergent evolution; see Foote et al. (2015). Different life forms – fish and aquatic mammals – have arrived at the same morphological shape because of the permanent sieving out of less efficient forms. The message of this excursion is that nobody would call for a "universal grammar" of aquatic life forms in order to explain universals of marine vertebrates. The theory of evolution is sufficient. The cross-species invariants are emergent traits due to "*convergent evolution*", and convergent evolution is not restricted to biological evolution.

One of the favorite showcases in the discussion of allegedly innate properties of grammars is the *structure dependency* of grammatical rules. Chomsky (2017: 5) repeatedly emphasizes that *"the only plausible conclusion, then is that structure-dependence is an innate property of the* *language faculty,* [...] *ignoring linear distance*" and "*ignoring properties of the externalized signal, even such simple properties as linear order".*

The rhetoric of this claim outweighs its factual significance by far. Why would linearity-based properties be "*simple properties"* and mentally easier to handle than structure-based ones? Our brain is excellent when it comes to chunking but very poor when it has to apply operations on unstructured, only linearly organized representations. The simple inversion of a list of items, for instance, is a computation that is very easy to program[19] but very hard to carry out mentally. No grammar uses list inversion of terminals as a grammatical rule although it is *conceptually* much simpler than structure-based conditions.

Structure dependence is just one side of the medal, with chunking on the other side. Structure dependence is the linguistic reflex of chunking which is an efficient way of recoding information. Our information processing brain is excellent in effectively chunking linear arrays of the input into hierarchically organized 'constituents' on every level of representation, from phonology to semantics, via morphology and syntax. Moreover, chunking is a domain-general capacity that is operative in various modalities of human information processing, such as vision, action planning, event perception (Lashley 1951, Martins et. al. 2016), and also in language processing.[20] So, we should not be amazed at all that grammatical rules operate on categorized chunks (aka phrases) rather than on serial properties of linear sequences of terminals. Structure-dependency is just another way of describing the fact that rules of grammar operate on the level of chunks and not on serial representations of elements. This is neither surprising nor does it justify the invoking of an innate UG. Recursion, as a 'hot' topic, is just another way of describing chunks that contain chunks of the same category. Since chunks may consists of chunks, recursion of chunks is an expected consequence for sufficiently powerful information processing systems.

The supposedly innate "merge" operation is an upside-down view of the conditions of the possibility of the effective processing of linearly presented signals. Chunking works in such a way that larger units are perceived as complex units made up of atomic units, piece by piece. Structure processing is not reverse engineering based on a structure-generating merge operation. Structure processing is *pattern detection* rather than *structure generation*. "Merge" is the operational mimicking of pattern detection. Structures are not generated. They are projected over strings in language intake. In language out-put, strings are organized in such a way that well-formed structures can be projected on them.[21] "Merge" is a highly misleading operational metaphor for structure projection.[22]

3.3. Closed vs. open system space

The history and future of a given grammar of a human language is dynamic, open ended, and vectored. This is what the theory of evolution predicts and this is what we see. On the other side, UG theory, as a closed system, predicts small, oscillating changes, with swift returns to

[19] In Prolog, a simple command like "reverse ([a,b,c,d,e,f,g], Results)" yields the inverted list [g,f,e,d,c,b,a]. Our brains evidently do not provide a list-reverse function.

[20] According to Dirlam's (1972) mathematical analysis of efficiency, the "*most efficient chunk size*" is three or four items per chunk. This, by the way, is an answer to the vexing question why verbs typically have only three or four arguments at most.

[21] In this perspective, in MP diction, syntax would completely reduce to spell-out.

[22] Here is an obviously wrong consequence: The derivation of structures by 'merge' has to start at the deepest point, the foot position, which is the end of the clause. However, the human parser does not postpone analysis to wait until a clause or a complex sentence is finished but it would have to, otherwise merge could not start.

the UG defaults back into a terrain that is fenced in by UG. The limits of UG are the limits of change.

In UG theory, the properties of each grammar would be the union set of invariant properties plus subsets of properties stemming from parametrized conditions. UG theory necessarily and explicitly encompasses any parametric variant of each and any human language that has ever been spoken and will ever be spoken by human beings. It is not just a *potential* of the theory that would be activated in case a particular property of the given language requires it. UG is active in any native language user, just like any cell of our body contains the full-fledged DNA of an individual.

Let us be concrete once again. The different linking systems of languages that regulate the linking of argument expressions to the argument grid of their lexical (verbal) head partitions the system space. There is the purely *topological* way, as in English or Chinese, there are *morphological* systems with parametrized case assignment, (nom-acc, abs-erg, split), there are particle systems like in Japanese and there are mixed topological & morphological systems (as in Icelandic), to name just a few of the attested means of linking.

For a UG-theory this means that *for a given language*, all the responsible parameters behind this cornucopia of alternative ways of argument identification must be set and fixed. So, for example, English is set negatively for the bunch of parameters that define a split-ergative language. Parameter setting calls for decisions on all these UG options and also on the possibility that in present perfect tense the alignment mode could switch from nom-acc to abs-ergative case marking, as in Grusinian or Hindi, to name just a few randomly chosen characteristics.

Active parameter setting cannot be assumed for *all* instances of parameters which are *not* defined in the given language if the learner has no chance to meet decisive data, as for instance: Is there a genitive of negation in Mandarin? Is there a Dative subject in Afrikaans? Is Dative a lexical case in the intransitive passive in English? So, there must be defaults valuations of parameters. But if there are defaults, languages would necessarily gravitate towards the defaults (because language acquisition would start in the default mode as a constant attractor) and end up with uniform grammars. What we observe is the exact opposite. Diachronically, languages are fanning out rather than converging. This we see clearly in the Indo-European family.

The evolutionary view of CEG, on the other hand, is confirmed by what we observe. There is no pre-specified closed system of variability.[23] Grammars change and the changes are not always streamlined but may develop in unexpected ways, sometimes and in rare cases. Not only nature provides room for unexpected creatures, evolutionary grammar theory has room for 'exotic' grammars, too. Luxurious diversity is a sign of evolution but an embarrassment for an innate *universal* grammar.

[23] This is the place for an aside on Cartographic Syntax, which maps UG instantiations onto trees. There is a universal tree structure for sentences whose 'style of decoration' changes from language to language. By the same token, we should ask ourselves why no zoologist would ever be interested in proposing a universal format of organisms, with parametrized features. Looking back at grammar cartography in this perspective should be an eye opener.

3.4 (Non-)Parametric

In grammar theory, cross-linguistic invariants are rare and cross-linguistic *variation* is dominant*,* just like in the example of sea-dwelling mammals. Here is an arbitrarily chosen example, namely interrogative clauses, and in particular content questions. In Slavic languages, *multiple* wh-fronting is the rule. Germanic languages provide room for only a *single* wh-expression in the clause initial position. There are languages with optional fronting and n yet other languages, wh-items remain always in situ. There are even languages that discriminate between two sets of wh-expressions, namely those that are fronted and those that remain in situ (Dryer 2013). This is not the fingerprint of a "Universal Grammar" as taskmaster but an expected outcome of grammar evolution. For UG, all these variations (macro- and micro-parameters) are unpredicted. If it were true that human grammars closely follow an innate blueprint, a single invariant structure would suffice and be appropriate.[24]

Initially, parametrization seemed to be an elegant way of capturing the embarrassing amount of variation that has to be covered by a theory of universal grammar. In the best of all UG worlds, the grammars of human languages would differ only in the lexicon and in the lexical form of affixes and particles, but not with respect to structuring and to syntactic operations. Our linguistic world, however, belongs to a universe that does not qualify as an optimal UG world.

UG theory predicts a clear-cut partioning of languages by boundaries marked by parameters. This is not the linguistic reality, however. The reality is fuzzy. In the past decades, it has turned out that the number of parameters grows with every language adduced. Macro-parameters had to be amended with micro-parameters and eventually, parameter 'theory' more and more looks like a re-statement of the descriptive facts in abstract terminology. There is no predictive gain:

> "*Quite generally, language-wide parameters that have been proposed over the last three decades have, upon closer examination, turned out not to neatly partition the world's languages into two sets [...]. Instead, each parameter, under closer examination, turns out to fragment into smaller parts*." (Son & Svenonius 2008: 395).

How many (macro- and micro)-parameters would UG have to provide?[25] In generative grammar, there has never been a strong ambition of presenting a theory of possible and impossible parameters. Gradually, 'parameter' has become a byword for variation. Rizzi (2014: 22) updates the parameter issue for the Minimalist Program. "*I would like to propose the following informal characterization. A parameter is an instruction to perform a certain syntactic action expressed as a feature on an item of the functional lexicon and made operative when the item enters syntax as a head.*"

---

[24] The original hope for unified account based on "covert movement" was not fulfilled. Covert A'-movement, if it existed, would not meet essential restrictions on A'-movement, see Haider (2018: 7-13). Adverbial clauses, for instance, which are inaccessible for overt wh-movement would have to be transparent for alleged covert movement in multiple questions as in the German example (i):

i. *Wieviel* muss man bezahlen [wenn man es *wie lange* mietet]?
How-much must one pay [if one it *how long* rents]
'How much does one have to pay if one rents it, depending on the respective rental period?'

[25] Let us assume, just for the sake of a concrete estimate, UG would involve exactly 100 binary parameters (which is a modest estimate). This defines a system space of 2^{100} different grammars with potentially interacting valuations. So, grammar acquisition would have to check out a way through a maze of grammar alternatives in order to end up with precisely one out of 1 267 650 600 228 229 401 496 703 205 376 possible variants.

What Rizzi writes is far afield from the parameter idea of the Principles and Parameter program. There, a parameter is a variable in an otherwise invariant, universal principle that needs to be valued with the appropriate value. The UG principles are deemed to be universal and the cross-linguistic differences are characterized as differences in terms of the valuation of parameters. In the Minimalist perspective, there is just feature variation. Some functional items may be associated with some feature in one language and with another feature in another language. And, let me add, a theory of (possible or impossible) features has never been an ambition in the unfolding of the Minimalist Program either. Here is an example, the so-called null-subject parameter. In a volume devoted to this topic, Roberts & Holmberg (2010: 14) characterize it as follows:

(1) The Null Subject Parameter: *Does T bear a D-feature*?

The valuation of this parameter is binary, namely 0 (no) or 1 (yes). In French and German, for instance, it happens to be null, in Italian and Spanish it is 1. The "D-feature" is a special pronominal feature for the licensing of the null pronoun "*pro*" in Spec T. However, the actual situation is far from being so simple.

It is true that standard German is not a null-subject language (2a,b), but like many other languages, it provides a null-subject construction in a particular type of finite clauses, namely in imperatives (2c). In Minimalist terms, there must be a D-feature for T in German but it must not be used in declarative and interrogative clauses. It identifies a second-person null pronoun, in singular or plural. What limits the feature to imperatives?

(2) a. Gestern prophezeite *(sie), dass *(es) heute regnet.
yesterday prophesied (she) that (it) today rains
b. Darüber müsst *(ihr) noch einmal nachdenken.
about-it must$_{2\text{-pl.}}$ (you$_{\text{pl.}}$) once more think
c. Jetzt denk (du) noch einmal nach!
now think$_{2\text{-pl.-imp.}}$ (you$_{\text{sg.}}$) once more about!

Even if we grant that problems with details are business as usual in every theory, the decisive moment in our case is not the details; it is the question of the relation of parametrized to invariant elements in UG. The set of cross-linguistically attested *invariants* is little compared to cross-linguistic *variation*. This is exactly the picture predicted by CEG. After well more than three hundred millennia, what we observe is the 'shock front' of the linguistic 'big bang' in the history of homo sapiens. Grammars develop and they develop in dissipative ways. However, the direction of change is not the direction towards a UG-directed uniformity or invariance.

UG theory predicts assimilation of grammars, not dissimilation, for the following reason. The UG system would be the conservative moment for variation and changes since it is the attractor that repairs deviations by transmission errors from generation to generation. Minor changes would continuously disappear because of 'swamping.' In fact, this was a major problem for Darwin's theory, known as Jenkin's (1867) swamping argument (Haider 2021:13): "*Jenkin objected to Darwin's theory by pointing out that an accidentally appearing profitable variety could not be preserved by selection. It would be 'swamped' by the ordinary traits in the course of backcrossing in the population*."

UG is a mechanism of conservation, hence changes should not spread easily. The original parameter values would prevail in the population since any nascent change would be swamped in the speech community. Language changes would have a chance only in cases of bilingualism in the course of migration or foreign domination. Thus, for UG theory, "*languages varying radically from one another*" (Chomsky 2017: 2) are an embarrassment. UG should lead to uniformity but not to enhanced variation. Where does variation originally come from?

What appears to be parametric is, in reality, the expected cross-linguistic variation of grammars shaped by cognitive evolution. What is a serious challenge for parameter theory is a corroborative fact for CEG. Fuzzyness is intrinsic. There are two main sources. One is *dissipative* variation that leads to dialect-splits and eventually to different languages, as in the case of language families that are continuations of a common 'mother' language. This is the momentum behind Darwin's descent of species. The other source is *convergent evolution*. It makes grammars of different languages similar in certain respects, with clusters of similarities that we perceive as types. The closer the similarities, the easier it is to analyse them as parametric variants.

3.5 Parameter valuation during language acquisition

In Generative Grammar, parametric differences are not inherited.[26] The learner has to identify the appropriate value for each parameter and set it, whatever this may mean for a learning brain. In fact, it is entirely obscure. The syntactic system is cognitively encapsulated. Neither children nor linguists have introspective access to it. Nevertheless, a learner is supposed to interact in a highly precise way. Here is how a proponent of parameter setting perceives the situation:

> "*Under the assumption that acquisition proceeds by parameter setting, the child does not pick its language whole out of a set consisting of all possible languages. Rather, it sets individual (syntactic) parameters, the end result of which is (the syntactic component of) a grammar. If the number of possible languages were so large that the number of parameters the child had to set was unmanageable (i.e. not learnable in the amount of time available), there would indeed be a problem*." (Kayne 2000: 8)

How long would it take to set a parameter? Is it a matter of setting a switch or of getting accustomed to particular patterns? This is not the prime concern here, however, since before a child can set a parameter, (s)he must be able to trace its effect in the data in order to determine its valuation. The child cannot stroll along an alley of parameters with the values lined up for the child who walks along to pick them up, one after the other, in the appropriate order. Quite to the contrary, they are part & parcel of the utterances the child is exposed to. But crucially, no utterance is labelled for the values of the parameters to which it owes its form.

Let us assume, a child has fixed a few parameters and is about to fix the next ones. How does the child find out the appropriate valuation? First it has to become aware that its default setting is not appropriate, and then it can only proceed by trial and error. The child tries out a particular value and checks the outcome. And this is the point where children will get stuck in cluelessness. A sentence is not labelled for its parameter values. A well-formed input sentence is the

[26] Not only is it fully consistent with the innateness hypothesis, but it has to be expected that parametric differences are innate. They would be a necessary consequence of mutations of an innate UG. The grammar of the waggle dance of honeybees is innate and parametric differences are innate, too; see Rinderer & Beaman (1995) and Kohl et al. (2020) on bee 'dialects'.

result of the interaction of *all relevant parameters* set in the adult language. So, the child maybe happens to set a parameter correctly, but due to other still unset parameters, the utterance is deviant and maybe more deviant than it would be with a wrong setting. In short, complex parameter-setting is intractable for a child (and for most adults, too).

It is this cognitive intractability that would make parameter identification and setting a fundamental obstacle for a child (and for theoretical linguists, too). So, linguists and children are expected to fail. For linguists, this is a fact, but not for children. Why are ten years enough for children to master a task that professional linguists do not accomplish in a lifetime? Because linguists do not have access to UG? Children do not have access to it either. It is cognitively encapsulated. It could *prime* pattern detection but it cannot communicate with the conscious mind and there is no demon who evaluates the input and compiles the given parameter in the learner's grammar with its appropriate value.

The setability problem is one of five problems studied by Boeckx & Leivada (2013) in a program-based analysis, with a clearly worded outcome: "*The empirical issues uncovered cast doubt on classical parametric models of language acquisition as well as on the conceptualization of an overspecified Universal Grammar that has parameters among its primitives.*" (Boeckx & Leivada 2013:1) "*In their totality, these problems suggest that the notion of parametric dependencies runs into empirical problems that should cast doubt on the feasibility of parametric approaches to UG.*" (Boeckx & Leivada 2013: 8)

For the sake of concreteness, let us analyse some easy cases of entanglement. Wexler (1998: 25) claims that children at the age of two already set the OV/VO, the V2, and the null-subject parameter. This claim is an inference from what children utter at that time, but not a proof of parameter setting. Today we know it better.[27] The acquisition of V2 in Germanic languages is not instantaneous but a lengthy process.[28] Waldmann (2012) reports a long phase of V3 for an otherwise typically developing case in Swedish. Fritzenschaft et al. (1990) report and analyze data from five German longitudinal studies and show that there is no evidence for an early setting of the V2 parameter in combination with the OV parameter. Even at the age of nearly four, children have troubles with the interaction of V2 and the presence of complementizers. All this does not come as a surprise if we acknowledge the problem structure of the task.

At the beginning, children primarily master short utterances in early child-directed speech. Such utterances frequently do not contain more than one verb, which is the finite main verb. In a V2 language, this verb is in second position, frequently preceded by the subject and followed by an object. If the child takes this as input for parameter setting in Dutch or German, it will wrongly set the parameter as if it were confronted with an SVO language. The *base* position of the verb is difficult to identify in minimal V2 clauses. The child will have to compare utterances with an auxiliary in the presence of a main verb, must check particle verbs, and it must first of all be able to distinguish finite from nonfinite forms. But even after this phase, there is no abrupt change in the patterns the child produces, that would indicate a parameter reset. The 'wrong'

[27] Akhtar (1999) presents an intriguing type of (counter-)evidence. She reports that two-year- and three-year-olds repeat sentences in non-standard SOV and VSO orders, along with standard English SVO when they contain novel verbs, while by age 4, children changed non-standard orders to standard SVO order.

[28] And a lengthy process was the diachronic emergence of V2, too, which took more than two centuries, see Prell (2003)

patterns just become fewer over time and gradually disappear. Parameter-setting would be an inappropriate description for this kind of development. It is not disruptive but a continuous and lengthy elimination process. The acquisition function is not a step-like curve but a flat sigmoid. This is not parameter-setting but pattern identification and gradual consolidation in use.

3.6 Mono- or poly-genesis of grammars

When Ken Hale shared the results of his field-work, a new parameter was born, namely the (non)-configurationality parameter (Hale 1983). In fact, this was a capitulating move. Warlpiri was too exotic from the perspective of an English-biased UG, and so it had to be quarantined away. The new starting point for the learner in the decision tree of parametrization was [± configurational], where [-configurational] names the branch on which the 'exotic' languages are to be found. Do all these languages have a common origin within the limits of a common UG?

Indo-European studies have confirmed the evolutionary concept of a common origin followed by dissipative evolution. The research time depth is necessarily shallow. It comprises about 4 millennia and this is roughly 1% of the time that has elapsed since the attested appearance of homo sapiens in the Mediterranean region (Richter et al. 2017). So, 99% of this time and presumably more is inaccessible. It is pure speculation if one insists on a mono-genetic origin of human languages. An innate UG would be such a claim. The evolutionary approach[29] invites a more promising perspective on this issue that is consistent with the state of the art of the theory of evolution. It is plausible that our antecedents started just like children start with language, namely with one- or two-word utterances. In view of the then sparsely populated African[30] continent, polygenesis of grammars during and after such a "*Me Tarzan – You Jane*" epoch is more probable than monogenesis of UG. It is not unreasonable to assume that modern grammars are the result of cognitive evolution that started with very simple utterances. In Nichols (2011: 572) words "*language originated gradually over a diverse population of pre-languages and pre-language families*".

The small set of cross-linguistic invariants of modern languages are the expected reflex of *convergent* cognitive evolution by constant cognitive selection of grammar variants by the invariant neuro-cognitive processing resources that constitute the human language-processing facilities. In the evolutionary perspective, this is a well-known phenomenon (Foote et als 2015). From the UG vantage point, the dissipative nature of language change is unexpected and hard to explain. What circumstances would give rise to a UG with lots of parameters if a UG without parameters would do the same job, but better? UG theory predicts narrow channels of diachronic change, with UG as the strong attractor and gate keeper. The opposite is the case. If UG theory were right, languages "*varying radically from one another*" should not have come into being.

[29] The following paragraphs are based on Haider (2023).
[30] As for Africa, Ragsdale et al. (2023: 755) conclude: *"The earliest population divergence among contemporary populations occurred 120,000 to 135,000 years ago and was preceded by links between two or more weakly differentiated ancestral Homo populations connected by gene flow over hundreds of thousands of years."*

3.7 (Im-)Perfection

Generativist grammarians tolerate and in fact propagate an argument that scientists would not admit, namely the *argument from perfection*, based on a dogma called the "*the strong minimalist thesis*" (STM): "*We can therefore formulate SMT as the thesis that all phenomena of language have a principled account in this sense, that language is a perfect solution to interface conditions."* (Chomsky 2007: 5).

The argument from perfection has a long and scientifically infamous tradition. First, it was used in ancient astronomy in predicting the invariant trajectories of heavenly objects. "*The 'natural' expectation for ancient societies was that the heavenly bodies must travel in uniform motion along the most 'perfect' path possible, a circle*." (Jones 2022). The exceptional trajectories of planets were deemed to be 'recursive' epicycles. They allegedly move on circles whose centers move on circles, and so on.

The next permanent on-stage appearance of perfectivity is in theology, in the ontological proof of the existence of God (see Crittenden 1968)[31] and eventually, it has been revived in Generative Grammar: *"One useful way* [...] *is to entertain the strong minimalist thesis SMT* [= strong minimalist thesis]$_{HH}$, *which holds that* FL [= faculty of language]$_{HH}$ is perfectly designed." "*We can therefore formulate SMT as the thesis that all phenomena of language have a principled account in this sense, that language is a perfect solution to interface conditions." "Universal grammar (UG) is reinterpreted as the theory of the initial state of FL."* (Chomsky 2007).

For a purportedly empirical discipline in the 21st century, this is a remarkable statement since it presupposes divine omniscience. How could anyone be sure that a design of nature is "perfect" and what would "perfect" mean? Of course, anyone may claim anything, but then, this is the same type of claims as Dr. Pangloss' amusing claim of perfection in Voltaire's Candide (see Weiss & Dunsworth 2011). No biologist would claim that humans ought to have the perfect eye design of an octopus, given that this is a more perfect solution than our eye with a blind spot (due to the fact that the nerves in our eyes leave the retina cells on the side exposed to the incoming light rays). After all, the eye must meet various kinds of complex interface conditions.

Kinsella (2009) points out that evolution based on natural selection does not lead to 'perfect' outcomes. Adaptation by selection leads to outcomes that are preferable to other variants in the pool of variants. It leads to what Simon (1956: 129) calls 'satisficing':[32] "*Evidently, organisms adapt well enough to "satisfice"; they do not, in general, "optimize*".

Let's assume for the sake of argumentation UG were the perfect setting indeed. In this case, the numerous different grammars of human languages as instances of UG must be perfect, too. This makes perfection a vacuous claim since the grammar of any human language will qualify as

---

[31] "*This argument, which was formulated first by Anselm and elaborated by such thinkers as Descartes, Leibniz, and Hegel, is commonly known as the ontological proof." "God is by definition a perfect being and indeed – if we may speak of degrees in perfection – a supremely perfect being. But it is self-contradictory to regard a supremely perfect being as non-existent; for to lack existence must be an imperfection. Hence a perfect being must exist*." (Paton 1955, Ch. 12).

[32] The term *satisficing* is a blend of *satisfy* and *suffice*.

perfect and we rightly have to wonder why there are so many highly different 'perfect' grammars. The alternative is that nobody is perfect. Evolution is full of imperfect solutions, and so are grammars.

It is easy to find all kinds of imperfections in grammars (see also Section 10). Here is an example from English. English is the only Germanic language that cannot passivize intransitive verbs because of the unavailability of a suitable expletive (Haider 2019). As a consequence, there is no *perfect* match at the semantics interface in cases in which the subject argument is to be cancelled, as in many other languages. A general case of imperfection will be discussed in Section 3.10, namely (in)consistency and (in)completeness of grammars as formal systems.

3.8 Universal SVO architecture vs. adaptive landscape

In the evolution of complex systems, ultimate perfection is an irrelevant issue. Even an imperfect solution may gain a selectional advantage if it is more efficient than other imperfect solutions. A little bit of vision is better than no vision at all. Moreover, variation & selection does not guarantee permanent progress. Often, a system ends up with a globally suboptimal but locally maximal property (Kauffman 1993: 43). Wright (1932) formulated a powerful visualization of adaptive evolutionary changes, namely the concept of adaptive landscape, which has become a widely used model (Svensson & Calsbeek 2012).

> "*An adaptive landscape shows the relationship between fitness (vertical axis) and one or several traits or genes (horizontal axes). An adaptive landscape can therefore be viewed as a form of response surface, describing how a dependent variable (fitness) is causally influenced by one or several predictor variables (traits or genes). Evolution by natural selection in the context of an adaptive landscape can be viewed as a hill-climbing process, in which populations climb upwards to the trait or gene combination with the highest fitness, which are called "adaptive peaks*" (Svensson 2021).

The fitness landscape or adaptive landscape of human grammars has a similar topology as the fitness landscape of biological species. It is full of peaks and valleys, and crucially, it is not a plain with a single, high peak defined by UG.

The development of Indo-European languages is a handy example. The earliest accessible testimonies show that these languages coded the grammatical relations only morphologically (mainly by case and agreement inflection), with hardly any sharp word order restrictions. In theoretical terms, they were T3 languages (Haider 2023). None of these ancient languages was an SVO language, although SVO is the default in Generative Grammar. Within three millennia, a subset has developed a strictly topologically coding system (SVO), namely the continental Scandinavian languages with no nominal case and no finiteness agreement. The Romance family developed into SVO languages, too, but preserved verbal agreement morphology to a certain extent (see Italian). A sizeable number of languages did not change much in their makeup. They are still free-word-order languages (i.e. T3) and most of them have a morphological case and agreement system, such as the Slavonic languages (except for South Slavic languages, such as Bulgarian and Macedonian). What this indicates is that some systems have changed more than others, and that each system is sufficiently stable to be transmitted from generation to generation. Each of these grammar types represents a peak in the adaptive landscape of grammars.

Some of the peaks will get higher, some stay at the hight reached in the course of time. Grammars of [S[VO]] languages amount to steadily growing peaks, as agued in Haider (2023). Let me duly emphasize at this point that this language type is a *late outcome* of grammar changes and not the starting point, as the Generativist bias would have it.

Word order typology sketches a similar picture. Around the world, there are sizeable sets of SVO and SOV languages, a smaller group of VSO languages, and a large group of so-called free word order languages, which typologists more often than not misfile as SVO; see Haider (2023). Obviously, any one of these languages represents the present-day outcome of a potentially millennial developmental history since each one of these languages is a system that has successfully managed to replicate from generation to generation.

3.9 Changes – rule vs. exception

The UG-view would be fully compatible with a situation in which all languages have the same uniform and permanent grammar and differ only in their lexical inventory. Under an evolutionary perspective, this is virtually impossible, just as it is impossible that all birds look like ducks and differ only in the colour of their feathers and beaks and in the way how they chatter (except after a nearly total extinction catastrophe for birds). The same is true for grammars. If grammars are the product of CEG, they will be(come) different. Evolution is based on, and produces, variation.

The constant element is change, and individual changes are not predictable. All we see is that changes are vectored and divergent. This is what Sapir has recognized a century ago and termed "drift". "*The drift of a language is constituted by the unconscious selection on the part of its speakers of those individual variations that are cumulative in some special direction.*" (Sapir 1921: 166). What Sapir describes is the interplay between random *variation* and constant and therefore *directional* selection and *retention*. This is the essence of Darwinian natural selection operating on grammatical systems:

> "*It by no means follows that the general drift of language can be understood from an exhaustive descriptive study of these variations alone. They themselves are random phenomena. The linguistic drift has direction. In other words, only those individual variations embody it or carry it which move in a certain direction*." (Sapir 1921: 165)

In evolution by natural selection, changes are vectored because of the constant and blind sieving out of variants. Crucially, the preceding variants tend to end up in the set of sieved-out ones. Hence, there is no constant oscillation between the original variant and its successor variant. Eventually, evolutionary changes are expected to progress along lineages, and they do (Dunn et al. 2011).

These lineages are reflected in phylogenetic trees which linguists draw for grouping languages in terms of (historical) relatedness of their grammars, see Gray et als. (2011). The parallel to biology cannot be overlooked. We perceive this as the result of changes in language families, as for instance the emergence of V2-languages in the Germanic family.

This syntactic property materialised in the documented history, with a single exception, namely English. In English, the change shaped wh-clause formation but did not extend to declarative clause structure. It was counteracted by another, specifically English change, namely the immovability of main verbs. Cross-linguistically, V2 is a rare property, but in the Minimalist literature, V2 is treated as an option straightforwardly expected in the feature distribution along the functional spine of a clause. If this were so, V2 languages should have spread all over the world. ‘

Pandemic’ parameter-mis-setting could happen easily. The learner of an SVO language could easily confuse a minimal SVO sentence (3a) with a V2-sentence (3b), or vice versa. In English, this would be a mistake, but in all Scandinavian ‘sister’ languages each minimal SVO sentence, see (3b), is a V2-sentence:

(3) a. The object [$_{VP}$ *follows* the subject]
b. Objektet *följer*$_i$ [$_{VP}$ e_i subjektet]] Swedish
object$_{def}$ follow$_{present.}$ subject$_{def}$

This is just a sketch of the fact that UG theory does lend itself to a model of grammar change, which is the gist of the evolutionary approach, however. The theory of evolution is a theory of change. In the Generativist perspective, grammar changes are *mishaps*, due to mistakes in the fine-tuning of the learner data and the UG-geared grammar acquisition process. These mishaps are not part of the system and UG is expected to filter them out. Evolution theory, on the other hand, regards these ‘mishaps’ as a constant and unavoidable element of the way how grammars are transmitted from generation to generation. Without these ‘mishaps’, evolution could not take place.

Finally, since UG is the limiting frame for each language, with a ‘perfect’ fit between the actual grammar and UG, due to UG being the taskmaster during language acquisition, changes are predicted to be mishaps under constant repair. The original parameter setting ought to prevail in the population. The overall result would be a *seesaw pattern* of changes, oscillating between the original state, an innovation, and then ‘back to normal’. This is not what we learn from diachronic linguistics and not what we see in the history of grammar changes.

Finally, CEG opens the door to insourcing (mathematical) laws of evolution theory. *Fishers Fundamental Theorem*,[33] see Fisher (1930), for instance, accounts for the “island effect” in language change. Small and rather isolated communities are more resistant to grammar change. Well-known examples are Faroese, Icelandic, Logudurese Sardinian, Soqotri, to name just a few languages with conserved ‘archaic’ grammar features. On the other hand, the languages of small and isolated language communities often surprise linguists with their ‘exotic’ grammars, which means that they provide insights in earlier stages of related languages.

Another candidate for insourcing is the concept of “*Evolutionary stable strategies*“ (ESS), as detected by Maynard Smith & Price (1973). These are tools for modelling and understanding paths of irreversible grammar changes (see Sapir’s drifts), as for instance, the change from morphological to topological coding, with the emergence of an [S[VO]] clause structure an *evolutionary stable state*, see Haider (2023).

[33] $\Delta G/\Delta t = V(G)$ [with G = genetic fitness, t = time, V = variance] ≈ The rate of change of the mean fitness of any population is equal to the variance in genetic fitness. Linguistic version: The *rate of change* of the mean fitness of any population of grammar variants is equal to the *variance* in their neuro-cognitive fitness.

3.10 Consistency & completeness

Informally, the two basic notions are defined as follows: A formal system is *syntactically consistent*, if it is not possible to derive *p* as well as non-*p*. A formal system is *syntactically complete* if and only if for every well-formed sentence of the system, it or its negation is provable in the system. Generative grammar, resting on the regime of UG exerted on each individual's grammar, takes the competence of the competent speaker to be consistent and complete since UG guarantees a perfect grammatical calculus for each language and the language users' competence is modelled as acts of theorem proving (see Sect. 5).

This is not what we observe in the linguistic reality. In a picture matching study with test subjects with and without academic background (shelf-stackers, packers, assemblers, or clerical workers), the latter group scored at 43% (on possessive locatives with quantifiers), 78% (on locative quantification) and 88% (on passives) of the full scores, which were attained by the subjects in the academic group (Dąbrowska 2015: 8). In addition, there is evidence that even students of linguistics systematically fail in areas of grammar that, according to UG, must be part of their competence (see below).

Let us first focus on consistency. Bech (1963: 295, 297) was the first grammarian to name and explicitly analyse grammatical inconsistencies in the German grammar. "*Grammatical laws in contradiction*" is the translated title of his publication. He noticed, for instance, a systematic rule conflict in German in the interaction of infinitival syntax with the syntax of the clause-final verbal clusters. The two conflicting "laws" are the following ones. On the one hand, in IPP constructions (= infinitivus pro participio, aka Ersatzinfinitiv), the auxiliary *haben* ('have') must not follow a modal (4a) since this would trigger the participial form of the modal, which is obsolete. Therefore, the auxiliary is fronted (4b) across the modal.

(4) a.*ohne dass man das beendigen *gemusst* HAT
without that$_{C°}$ one it finish must $_{past\text{-}partic.}$ has ('without having to have finished it')
b. ohne dass man das beendigen HAT *müssen*
without that $_{C°}$ one it finish$_{Inf.}$ has must $_{Inf.}$ ('without having to have finished it')

On the other hand, the infinitival marker *zu* ('to') must occur on the final verb of a verb cluster. Thus, a clash of the two conditions is pre-programmed. If (4b) is transformed into the infinitival variant (5a), the infinitival marker would go together with haben ('*have'*), but the auxiliary is not in the final position in the cluster anymore. What is the grammatically consistent solution? There is none. Speakers either avoid this construction or they pretend to obey both rules and end up in a situation of "acceptable ungrammaticality" (Haider 2011). (11b) is accepted as the less deviant way out of the dilemma w.r.t. (11a).

(5) a. *ohne das beendigen *zu haben* müssen[34]
without it finish to have must$_{Inf.}$

[34] A note for the native German readers: When testing your judgements, you should make sure that you don't confound this example with a different cluster. The *base* order of (5) is $V_{Inf.}$ – modal – have. It must not be confused with the base order $V_{Partic.}$– have – modal, which is a different cluster, with a different meaning. Note that the main verb in (5) is infinitival, not participial, which shows that it depends on the modal and not on the auxiliary *haben*.

b. ?ohne das beendigen *haben zu* müssen
without it finish have to $must_{Inf.}$

In two production tests written in-class, one with 19 native German students of linguistics, the other with 17 native German students of German studies ("Germanistik"), 1 out of 19 and 3 out of 17, respectively, produced the 'correct' result, (5b),[35] that is, one of the variants of prescriptive grammar. The majority either ignored the IPP rule or produced gibberish. The details are described in Haider (2011: 233-236). This is an indication that even educated speakers of German get in trouble when abiding by their native competence. They get in trouble because the grammar underlying their competence is inconsistent. More cases of inconsistencies can be found in Haider (2011) and in Reis (2017).

4. On the epistemological status of UG

Already from its beginning, UG has been more like a brainteaser than a theory of an empirical substance. UG is the narrative needed for bridging the abyss between a highly abstract Generative grammar of a given language and the psycho-linguistic reality of children acquiring a grammar of their language during language acquisition. In the P&P model, UG was the collecting basin of all unlearnable properties accrued in the development of the continuously more and more abstract modelling of grammars. With the switch to the Minimalist program, UG has been turned upside down but only by reversing the perspective of describing it. Although Chomsky (2007: 7) refers to his view as "*approaching UG from below*" and "bottom-up", it is still an entirely top-down perspective:

> "*The MP seeks to approach the problem 'from bottom up': How little can be attributed to UG while still accounting for the variety of I-languages attained.*" And then he adds: "*One useful way to approach the problem from below is to entertain the strong minimalist thesis SMT, which holds that FL is perfectly designed.*" *"A particular language is identified at least by valuation of parameters and selection from the store of features made available by UG, and a listing of combinations of these features in LIs (the lexicon)."*

Chomsky's essay is a discussion of assumptions of proponents of the Minimalist Program and of how they relate to the core axiom, the SMT (strong minimalist thesis) of a perfectly designed language faculty. Strikingly, the paper does not touch any cross-linguistically ascertained specific linguistic facts at all. The promised approach "*from bottom up*" is not part of the paper.

Apart from all details, a cardinal defect of this theory is its isolation from neighbouring fields. UG is dissociated from (behavioural) genetics, from (cognitive) anthropology, from (human) neuro-science, from psycho-linguistics, from cognitive psychology, from field linguistics and linguistic typology. Biologists do not take seriously what nativists speculate about. Here is an outspoken statement from a biologist who contributes to a volume on language universals:

> "*Bemusement is this biologist's response when straying into cognitive territory, regarding its denizens prospecting for the universals of language and cognition. What could they be looking for, and what would the demonstration of a universal feature of language learning signify to them? If the language prospectors believe the world to be unstructured, the vehi-*

[35] Or a variant with 'haben' (have) in front of the main verb,

cles of perception and production unlimited, the content of communication, and the evolutionary possibilities of the brain relevant to communication unconstrained, then the appearance of "language universals" in independent language learners would be a remarkable and illuminating finding. [...] *But if any aspect of the world is structured, if available information has predictable content or history, or if information-processing capacities were limited, universals could arise from any or all of these sources, if we may draw parallels with other biological information-transmission devices*." (Finlay 2009: 261)

Biologists are familiar with all kinds of 'UGs,' e. g. the UG of aquatic life forms, with fins and streamlined bodies, or of the volant life forms, with wings and hollow bones, and of all other kinds of life forms, all of which are adapted to their habitat. They do not have to invoke a mysterious UG since they know since more than a century how these life forms have developed. UG linguists, however, completely eclipse the fact that the human language capacities rest on a developmental history of the partially recruited mental capacities[36] (memory functions, pattern detection and analysis, etc.) and bodily organs (see articulation). In Gould's (1980: 20) words "*they are jury-rigged from a limited set of available components*". Strangely enough, UG theory is ahistorical. The fact that grammars owe their present form and function to a multi-millennial history of cognitive evolution of human grammar systems has no relevance for it.

It is this moment that is completely lacking in the Generativist UG theory although it is the only scientific and empirically founded answer to the cavalier statement that "*FL may indeed be well-designed to satisfy CI interface conditions*" (Chomsky 2007: 28). There is a designer, indeed, and this designer is Dawkins' (1991) "blind watchmaker", that is, the (cognitive) evolution (of grammars).

Chinese and English have grammar systems of reduced complexity (on the side of the declarative neuro-cognitive network). Cognitive evolution favours the reduction of complexity that would strain the declarative memory system. In general, it rewards shifting the processing load from the 'costly' declarative network to the 'cheaper' procedural network. It is an open question how much of the overall complexity is reduced, but it is evident that Kolmogorov complexity is reduced in languages like English, Chinese, or Afrikaans, in comparison to Russian, Sanskrit, or Warlpiri. If you are in doubt, compare the case and agreement systems in these languages and count the number of relevant pages with their full descriptions in the respective grammar books. This will yield a rough estimate of the Kolmogorov complexity in each language.

5. UG viewed from inside and outside

From an outside perspective, UG is just a scientific hypothesis and its merits are judged like those of any scientific hypothesis. The dominant view from neighboring disciplines ranges between indifference and disbelief, see for instance Edelman & Christiansen (2003). The view from the inside is the opposite. Generative grammar without UG would be but a highly abstract way of looking at grammars, packed with empirically insufficiently supported, complicated conditions and lengthy derivations. Without UG as scaffold, any Generative grammar of a given language, with its numerous covert elements and covert operations, is speculative and unacquirable. Even if a given analysis may be highly implausible, precarious or ultimately

[36] Bates (1999: 244): „*Language is a new machine built out of old parts, reconstructed from those parts by every human child.*"

wrong,[37] it is accepted if it can be argued to follow from, or add to, some of the already accepted axioms and 'theorems' with UG status, or if it is needed for maintaining empirically challenged core assumptions. If someone dares ask how such an abstract system with its overly complex derivational procedures could possibly be acquired, the baffling answer is (a polite paraphrase of) "*It's all innate, stupid*!"

Generative grammar characterizes grammatical well-formedness as a result of theorem proving, with the speaker as introspective observer of mental computations. A given expression is grammatically well-formed if it can be derived in a well-formed way. The 'proof' is the derivation. If an expression is ungrammatical, the proof fails.[38]

> "*Intuitively, the proof "begins" with axioms and each line is added to earlier lines by rules of inference or additional axioms. But this implies no temporal ordering. It is simply a description of the structural properties of the geometrical object "proof." The actual construction of a proof may well begin with its last line, involve independently generated lemmas, etc*." (Chomsky 2007: 6).

Nobody has bothered to find out whether this is a psycho-linguistically feasible[39] task at all and whether our brain really supports a mental capacity of higher 'grammatical algebra,' with an effectively working theorem-proving component. As research on vision has revealed, our brain is excellent in pattern processing (representation, feature extraction, classification, matching, storage, and retrieval) but much slower and less effective in rule following. The former is a 'geometric' capacity, not an 'algebraic' one, as the latter. Generative grammar insists on a rule-following algebra, for reasons far away from the empirical underpinning.

For Lasnik & Uriagereka (2002: 149), UG is not more than "*a rational conjecture*" that "*children come equipped with a priori knowledge of language*". Let us grant a reading of 'a priori knowledge' that Pullum & Scholz (2002: 17) and Scholz & Pullum (2002: 187) characterize as "innately-primed learning", even if it may be unclear how it works. The mere possibility of feasibility does not relieve the proponents of such a hypothesis from demonstrating what this knowledge consists of and how it is put to use in real time. Without compelling facts it's just fantasy.

6. Summary

In a direct confrontation of the two hypotheses – CEG (cognitive evolution of grammar) and UG (universal, innate grammar) – CEG proves to be superior. The explanatory power of CEG is the explanatory power of the theory of evolution. The relevant linguistic evidence supports CEG rather than UG. In particular,

---

[37] Here is a taster: *Obligatorily covert pied-piping* of wh-phrases [sic!], but only for in-situ *argumental* wh-expressions. (Choe 1987, Pesetsky 1987, Nishigauchi 1990). The innate UG is supposed to somehow bring it about that argumental wh-items in extraction islands are covertly pied-piped together with the whole island. It would apply only in English-like languages, since in German, the *arguments-only restriction* clearly does not hold [see Haider (2018: 9-13) for data and discussion].

[38] And if a competent speaker fails to arrive at the proof of a grammatical expression, (s)he is probably entangled in a garden-path situation. What is entirely missing is the proof of the operational efficiency of the general assumption. Why can we be sure that our brain effectively supports a theorem-proving device of this complexity given our well-demonstrated lack of talent in other, but similar, situations of theorem proving?

[39] Labelle (2007) argues that human (and even more so: juvenile) short-term memory capacities are by far too limited for computing the complex processes that current Minimalist model presupposes.

- CEG provides a straightforward solution of the problem of *grammar acquisition* since it is the essence of evolution that a system is and remains self-replicating. Grammars as cognitive apps adapt to the neuro-cognitive environment that is operative in acquisition and use.
- CEG predicts the principally dissipative course of *diachronic changes* while UG entails that in the long run, all changes converge to the default values of UG.
- CEG explains the vectored quality of *grammar changes* and excludes seesaw-type changes that UG would admit and predict.
- CEG explains the tension between *cross-linguistic invariants* and *cross-linguistic variation* as the result of divergence by variation and convergence channelled by the selection environment (viz. convergent evolution).
- CEG captures variation without parametrization and without the concomitant problems of parameter fixation.
- CEG provides room for economy effects in the shaping of grammars.
- CEG ties *grammar theory* to the most successful scientific theory of dynamically developing self-replicative systems, namely the general theory of evolution.

The linguistic version of Theodosius Dobzhansky's (1973) well-known maxim "*Nothing in biology makes sense except in the light of evolution*" is this: Nothing in grammar theory makes sense except in the light of the cognitive evolution of grammars.

Bibliography

Akhtar, Nameera 1999. Acquiring basic word order: evidence for data-driven learning of syntactic structure. *Journal of Child Language* 26: 339–356.

Atkinson Elizabeth G, Audesse Aamanda J., Palacios Julia A. et al. 2018. No evidence for recent selection at FOXP2 among diverse human populations. *Cell* 174(6):1424-1435.e15.

Aunger, Robert (ed.) 2001. *Darwinizing culture: The status of memetics as a science*. Oxford: Oxford University Press.

Bates, Elizabeth. 1999. On the nature and nurture of language. In Emilio Bizzi, Pietro Calissano, & Virginia Volterra (eds.). *Frontiere della biologia. The brain of homo sapiens*. 241–265. Rome: Istituto della enciclopedia Italiana fondata da Giovanni Trecanni.

Bech, Gunnar. 1963. Grammatische Gesetze im Widerspruch. *Lingua* 12: 291–299

Berger, Lee R. & Hawks, John. 2023. Revisiting the age of the hominin material. *Homo* 74(1): 55–60.

Boeckx, Cedric & Evelina Leivada. 2013. Entangled parametric hierarchies: Problems for an overspecified universal grammar. *PLoS ONE* 8(9): e72357.

Bynon, Theodora. 1986. August Schleicher: Indo-Europeanist and general linguist. In Bynon, Theodora & F. R. Palme (eds.) *Studies in the History of Western Linguistics*. 129-149. Cambridge: Cambridge University Press.

Chater, Nick & Morten H. Christiansen 2008. Language as shaped by the brain. *Behavioral and Brain Sciences*. 31: 489–558

Choe, Jae W. 1987. LF Movement and Pied-Piping. *Linguistic Inquiry* 18: 348-353.

Chomsky, Noam. 1981. Lectures on government and binding. Dordrecht: Foris.

Chomsky, Noam. 2007. Approaching UG from Below. In: Sauerland, Ulli and Gärtner, Hans-Martin (eds.). *Interfaces + Recursion = Language? Chomsky's minimalism and the view from syntax-semantics*. 1-29. Berlin: Mouton de Gruyter.

Chomsky, Noam 2011. Language and other cognitive systems. What is special about language? *Language Learning and Development*. 7: 263–278.

Chomsky, Noam. 2017. The Galilean challenge: Architecture and evolution of language. *Journal of Physics: Conference Series,* 880(1): 012015,

Collins, Chris. 2022. The complexity of trees, Universal Grammar and economy conditions. *lingbuzz*/006774

Crittenden, Charles. 1968. The argument from perfection to existence. *Religious Studies* 4 (1): 123 – 132.

Dąbrowska, Ewa. 2015. What exactly is Universal Grammar, and has anyone seen it? *Frontiers in Psychology*. 6:852. doi: 10.3389/fpsyg.2015.00852

Davidson, Donald. 1980. *Essays on actions and events*. Oxford: Clarendon Press

Dawkins, Richard. 1991. *The blind watchmaker*. London: Penguin.

De Boer, Bart & Thompson, Bill & Ravignani, A. & Boeckx, C. (2020). Evolutionary dynamics do not motivate a single-mutant theory of human language. *Nature Science Reports,* 10, 451.

Dirlam, David. 1972. Most efficient chunk size. *Cognitive Psychology* 3: 355-359.

Dobzhansky, Theodosius G. 1973. Nothing in biology makes sense except in the light of evolution. *American Biology Teacher*, 35: 125-129,

Dryer, Matthew S. 2013. Position of interrogative phrases in content questions. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) The World Atlas of Language Structures Online. http://wals.info/chapter/93. Accessed on 2022-08-19.

Dunn, Michael, Simon J. Greenhill, Stephen C. Levinson & Russell D. Gray. 2011. Evolved structure of language shows lineage-specific trends in word-order universals. *Nature* 473. 79-82.

Edelman, Shimon und Morten H. Christiansen 2003. How seriously should we take Minimalist syntax? *Trends in Cognitive Science*. 7(2): 60-61.

Finlay, Barbara L. 2009. Evolution, development, and emerging Universals. In Morten H. Christiansen, Christopher Collins, and Shimon Edelman (eds.). *Language Universals*. 261-265. Oxford: Oxford Univ. Press.

Fisher, Ronald A. 1930. *The genetical theory of natural selection*. Oxford: Clarendon Press.

Foote, Andrew D. Yue Liu, Gregg W. Thomas, Tomáš Vinař, Jessica Alföldi, Jixin Deng, Shannon Dugan, Cornelis E. van Elk, Margaret E. Hunter, Vandita Joshi, Ziad Khan, Christie Kovar, Sandra L. Lee, Kerstin Lindblad-Toh, Annalaura Mancia, Rasmus Nielsen, Xiang Qin, Jiaxin Qu, Brian J. Raney, Nagarjun Vijay, Jochen B. Wolf, Matthew W. Hahn, Donna M. Muzny, Kim C .Worley, Thomas P. Gilbert & Richard Gibbs. 2015. Convergent evolution of the genomes of marine mammals. *Nature Genetics* 47: 272-275.

Fritzenschaft, Agnes, Ira Gawlitzek-Maiwald, Rosemarie Tracy & Susanne Winkler. 1990. Wege zur komplexen Syntax. *Zeitschrift für Sprachwissenschaft* 9: 52-134.

Gaatone, David 1998. *Le passif en français*. Paris and Bruxelles: Duculot.

Goldberg, Adele 2013. Constructionist approaches. In Thomas Hoffmann,& Graeme Trousdale (eds.). *The Oxford handbook of Construction Grammar*. 15-31. Oxford University Press.

Gould, Stephen J. 1980. *The panda's thumb*. New York: W. W. Norton.

Gray, Russell D., Quentin D. Atkinson & Simon J. Greenhill. 2011. Language evolution and human history: what a difference a date makes. *Phil. Trans. R. Soc. B*. 366: 1090-1100.

Haider, Hubert. 1986. Affect alpha. *Linguistic Inquiry* 17: 113-126.

Haider, Hubert. 1987. Expletives pro - eine Phantomkategorie. [Expletive pro- a phantom category]. In Dressler, Wolfgang, Corrado Grassi, Rositta Rindler-Schjerve & Martin Stegu (eds.) *Parallela III.* 237-247. Tübingen: Narr.

Haider, Hubert. 1989. Matching projections. In: Anna Cardinaletti & Guglielmo Cinque & Giuliana Giusti (eds.). *Constituent Structure*. 101-121. Dordrecht: Foris

Haider, Hubert. 1992. Branching and Discharge. *Working papers of the SFB 340*. #23. Univ. Stuttgart, Univ. Tübingen, IBM Heidelberg. [2000 in Peter Coopmans & Martin Everaert & Jane Grimshaw (eds.) *Lexical specification and insertion*. 135-164. Amsterdam: Benjamins.

Haider, Hubert. 1997. Typological implications of a directionality constraint on projections. In Artemis Alexiadou & Tracy A. Hall (eds.) Studies on Universal Grammar and typological variation. 17-33. Amsterdam: Benjamins.

Haider, Hubert. 1999. On the survival of the fittest grammar (theory). *Zeitschrift für Sprachwissenschaft* 18(2): 216-218. [Open access: https://www.degruyter.com/document/doi/10.1515/zfsw.1999.18.2.216/html]

Haider, Hubert. 2000a. Towards a superior account of superiority. Lutz, Uli & Gereon Müller & Arnim von Stechow (eds.) *Wh-scope marking*. 231-248. Amsterdam: Benjamins.

Haider, Hubert. 2000b. OV is more basic than VO. In Peter Svenonius (ed.). The Derivation of VO and OV. 45-67. Amsterdam: Benjamins.

Haider, Hubert. 2004. Pre- and postverbal adverbials in VO and OV. *Lingua* 114(6): 779-807.

Haider, Hubert. 2011. Grammatische Illusionen. *Zeitschrift für Sprachwissenschaft* 30: 223-257.

Haider, Hubert. 2013. *Symmetry breaking in syntax.* Cambridge: Cambridge University Press

Haider, Hubert. 2014. The VO-OV split of Germanic languages - a T3 & V2 production. *Interdisciplinary Journal for Germanic Linguistics and Semiotic Analysis*.19(1): 57-79.

Haider, Hubert. 2015a. "Intelligent design" of grammars – a result of cognitive evolution. In: Aria Adli & Marco García García & Göz Kaufmann (eds.). *Variation in language: System- and usage-based approaches*. 205-240. Berlin/New York: de Gruyter.

Haider, Hubert. 2015b. Head directionality – in syntax and morphology. In: Antonio Fábregas, Jaume Mateu, Mike Putnam eds. *Contemporary linguistic parameters*. 73-97 London: Bloomsbury Academic.

Haider, Hubert. 2018. *On Minimalist theorizing and scientific ideology in grammar theory.* Ling-buzz/ 004967

Haider, Hubert. 2019. On absent, expletive, and non-referential subjects. In Peter Herbeck, Bernhard Pöll, and Anne C. Wolfsgruber (eds.) *Semantic and syntactic aspects of impersonality*. 11-46. Hamburg: Helmut Buske.

Haider, Hubert. 2021a. Grammar change – a case of Darwinian cognitive evolution. *Evolutionary Linguistic Theory* 3(1): 6-55.

Haider, Hubert. 2021b. Responses to the responses. 2021. *Evolutionary Linguistic Theory* 3(1): 109-121.

Haider, Hubert. 2022a. The Left-Left Constraint – A structural constraint on adjuncts. In Freywald, Ulrike & Simon, Horst J. (eds.) Headedness and/or grammatical anarchy? 177-210. Berlin: Language Science Press.

Haider, Hubert. 2023. SVO – Attractor in the declarative-to-procedural shift in grammar evolution. *Acta Linguistica Academica*. 70. DOI: https://doi.org/10.1556/2062.2023.00642. [*lingbuzz*/006747]

Haider, Hubert & Luka Szucsich. 2022a. Slavic languages – "SVO" languages without SVO qualities? *Theoretical Linguistics* 48: 1-39.

Haider, Hubert & Luka Szucsich. 2022b. Slavic languages are Type 3 languages – Replies. *Theoretical Linguistics* 48: 113-142.

Hale, Ken. 1983. Warlpiri and the grammar of non-configurational languages. *Natural Language and Linguistic Theory* 1: 5-47.

Harding, Sandra. 1976. Introduction. In: Harding, Sandra (ed.) *Essays on the Duhem-Quine thesis. Can theories be refuted*? ix-xxi. Dordrecht: Reidel.

Hauser, Marc D., Chomsky, Noam, Fitch, W. Tecumseh. 2002. The faculty of language: What it is, who has it, and how did it evolve? *Science* 298: 1569-1579.

Haworth, Claire M., Wright MJ, Luciano M, Martin NG, de Geus EJ, van Beijsterveldt CE, et al. 2010. The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Molecular Psychiatry*. 15: 1112-1120.

Holland, John H. 1992. *Adaptation in natural and artificial systems*. Cambridge: MIT Press.

Jenkin, Fleeming. 1867. The origin of species. *North British Review*. 46: 277-318.

Jermias, Johnny. 2001.Cognitive dissonance and resistance to change: The influence of commitment confirmation and feedback on judgement usefulness of accounting systems. *Accounting, Organizations and Society*. 26 (2): 141–160.

Jin, Lifeng &William Schuler. 2020. Memory-bounded neural incremental parsing for psycholinguistic prediction. *Proceedings of the 16th International Conference on Parsing Technologies and the IWPT 2020 shared task on parsing into enhanced universal dependencies*. 48-61. Association for Computational Linguistics. URL: https://aclanthology.org/2020.iwpt-1.pdf [Aug. 2022]

Jones, Raymond A. 2022. Ptolemaic system. *Encyclopaedia Britannica.* https://www.britannica.com/science/Ptolemaic-system. [August 2022].

Karlsson, Fred. 2010. Review of A.R. Kinsella, "Language evolution and syntactic theory". *Moderna språk*, 2: 47-50.

Kauffman, Stuart A. 1993. *The origins of order, self-organization and selection in evolution*. Oxford: Oxford University Press.

Kayne, Richard. 2000. *Parameters and universals*. Oxford: Oxford University Press.

Kinsella, Anna R. 2009. Language evolution and syntactic theory. Cambridge: Cambridge University Press

Kinsella, Anna R. & Gary F. Marcus. 2009. Evolution, perfection, and theories of language. *Biolinguistics* 3: 186–212.

Kirby, Simon. 2001. Spontaneous evolution of linguistic structure: An iterated learning model of the emergence of regularity and irregularity. *IEEE Transactions on Evolutionary Computation* 5 (2): 102-110.

Kohl, Patrick L, Thulasi Neethu, Rutschmann Benjamin, George E.A, Steffan-Dewenter I, Brockmann A. 2020. Adaptive evolution of honeybee dance dialects. *Proceedings of the Royal Society*. B287:20200190.

Kriengwatana, Buddhamas, Paola Escudero and Carel ten Cate. 2015. Revisiting vocal perception in non-human animals: a review of vowel discrimination, speaker voice recognition, and speaker normalization. *Frontiers in Psychology* 5, article 1543.

Labelle, Marie. 2007. Biolinguistics, the Minimalist Program, and psycholinguistic reality. *Snippets* 14. http://www.ledonline.it/snippets/.

Lakatos, Imre. 1970. Falsification and the methodology of scientific research programmes. In Lakatos, Imre & Alan Musgrave (eds.). *Criticism and the growth of knowledge*. 91-196. Cambridge: Cambridge University Press.

Lakatos, Imre. 1978. *The methodology of scientific research programmes*. Vol. 1. *Philosophical papers*. Cambridge: Cambridge University Press.

Lashley, Karl S. 1951. The problem of serial order in behavior. In L. A. Jeffress (ed.). *Cerebral mechanisms in behavior. The Hixon Symposium*. 112-146. New York: Wiley.

Lasnik, Howard & Juan Uriagereka. 2002. On the poverty of the challenge. *The Linguistic Review* 19: 147-150.

Liao, Dana A. & Brecht K.F. & Johnston, N. & Nieder, A. 2022. Recursive sequence generation in crows. *Science Advances.* 8(44): eabq3356. doi: 10.1126/sciadv.abq3356.

Martins, Maurício D., Martins, Isabel P. & Fitch, W. Tecumseh. 2016. A novel approach to investigate recursion and iteration in visual hierarchical processing. *Behavioral Research Methods* 48: 1421-1442.

Maynard Smith, John. 1986. *The problems of biology*. Oxford: Oxford University Press.

Maynard Smith, John & George Price. 1973. The logic of animal conflict. *Nature* 246, 15-18.

Mollon, Josephine, Knowles, Emma E.M., Mathias, Samuel R. et al. 2021. Genetic influence on cognitive development between childhood and adulthood. *Molecular Psychiatry* 26: 656–665.

Nichols, Johanna. 2011. Monogenesis or polygenesis: A single ancestral language for all humanity? In Tallerman, Maggie and Kathleen R Gibson. (eds.) *The Oxford handbook of language evolution.* 558-572. Oxford: Oxford Univ Press.

Nishigauchi, Taisuke. 1990. *Quantification in the theory of grammar*. Dordrecht/Boston/ London. Kluwer.

Paton, Herbert J. 1955. *The modern predicament – A study in the philosophy of religion*. London: Allen & Unwin.

Pesetsky, David. 1987. Wh-in-situ: Movement and unselective binding. In Eric J. Reuland & Alice ter

Meulen (eds.), *The representation of (in)definiteness*, 98-129. The MIT Press, Cambridge.Pollard, Carl and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.

Prell, Heinz Peter. 2003. Typologische Aspekte der Mittelhochdeutschen Prosasyntax. Der Elementarsatz und die Nominalphrase. In: Anja Lobenstein-Reichmann und Oskar Reichmann eds. *Neue historische Grammatiken*. 241-256. Tübingen: Niemeyer.

Pullum, Geoffrey K. & Barbara C. Scholz. 2002. Empirical assessment of stimulus poverty arguments. *The Linguistic Review* 19: 9-50.

Ragsdale, Aaron P., Weaver, Timothy D., Atkinson, Elizabeth G., et al. 2023. A weakly structured stem for human origins in Africa. *Nature* 617: 755-763.

Reis, Marga. 2017. Grammatische Variation und realistische Grammatik. In Konopka, Marek & Wöllstein, Angelika (eds.). *Grammatische Variation. Empirische Zugänge und theoretische Modellierung*. 255-282. Berlin: de Gruyter.

Richter Daniel, Rainer Grün, Renaud Joannes-Boyau, Teresa Steele, Fethi Amani, Mathieu Rué, Paul Fernandes, Jean-Paul Raynal, Denis Geraads, Abdelouahed Ben-Ncer, Jean-Jacques Hublin, Shannon Mcpherron. 2017. The age of the hominin fossils from Jebel Irhoud, Morocco, and the origins of the Middle Stone Age. *Nature* 546 (7657): 293-296.

Rinderer, Thomas E. & Lorraine D. Beaman. 1995. Genic control of honey bee dance language dialect. *Theoretical and Applied Genetics* 91: 727–732.

Rivière, Nicole. 1981. La construction impersonnelle en français contemporain. *Documents de Linguistique Quantitative.* no. 41. St. Sulpice-de-Favière: Éditions Jean-Favard.

Roberts, Ian & Anders Holmberg. 2010. Introduction: parameters in minimalist theory. In Biberauer, Theresa, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.) *Parametric variation: Null subjects in Minimalist Theory*. 1-57. Cambridge: Cambridge University Press

Rogers Timothy T. & James L. McClelland. 2014. Parallel distributed processing at 25: further explorations in the microstructure of cognition. *Cognitive Science* 38. 1024-1077.

Sandler, Wendy & Irit Meir & Carol Padden & Mark Aronoff 2005. The emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences* 102(7): 2661-2665.

Sapir, Edward. 1921. *Language, an introduction to the study of speech.* New York: Harcourt, Brace & Co.

Schoenemann, P. Thomas. 2012. Evolution of brain and language. In Michel A. Hofman and Dean Falk (eds.) *Progress in brain research.* 195: 443-459. Amsterdam: Elsevier.

Scholz Barbara C. & Geoffrey K. Pullum. 2002. Searching for arguments to support linguistic nativism. The Linguistic Review 19, 185–223.

Scholz Barbara C. & Geoffrey K. Pullum. 2006. Irrational nativist exuberance. In Robert Stainton (ed.), *Contemporary debates in cognitive science*. 59–80. Oxford: Blackwell.

Simon, Herbert A. 1956. Rational choice and the structure of the environment. *Psychological Review,* 63: 129-138.

Smolin, Lee 1992. Did the universe evolve? *Classical and Quantum Gravity*. 9:173.

Son, Minjeong & Peter Svenonius. 2008. Microparameters of cross-linguistic variation: Directed motion and resultatives. In Abner, Natasha & Jason Bishop (eds.). *Proceedings of the 27th West Coast Conference on Formal Linguistics.* 388-396. Somerville, MA: Cascadilla.

Stockwell, Robert P., Paul Schachter & Barbara Hall Partee. 1973. *The major syntactic structures of English*. New York: Holt, Rinehart and Winston.

Svensson, Erik & Ryan Calsbeek (eds.) 2012. *The adaptive landscape in evolutionary biology*. Oxford: Oxford University Press

Svensson, Erik. 2021. *Adaptive Landscapes*. DOI: 10.1093/obo/9780199941728-0073

Vidal, Céline. 2022. Age of the oldest known Homo sapiens from eastern Africa. *Nature*. DOI: 10.1038/s41586-021-04275-8. www.nature.com/articles/s41586-021-04275-8

Waldmann, Christian. 2012. Moving in small steps towards verb second: A case study. *Nordic Journal of Linguistics.* 34(3): 331-359

Weiss, Kenneth M. & Holly M. Dunsworth. 2011. Dr. Pangloss's nose – In evolution, cause, correlation, and effect are not always identical. *Evolutionary Anthropology* 20: 3-8.

Wexler, Ken. 1998. Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua* 106: 23-79.

Wright, Sewall. 1932. The roles of mutation, inbreeding, crossbreeding, and selection in evolution. In Donald F. Jones (ed.) *Proceedings of the sixth international congress on genetics*. 1: 355-366. Menasha: Brooklyn Botanic Garden.

Yu, Kristine M. 2021. Tonal marking of absolutive case in Samoan. *Natural Language and Linguistic Theory* 39: 291-365

Xu, Shuquin & Liu Pei & Chen Yuanxing, et al. 2018. Foxp2 regulates anatomical features that may be relevant for vocal behaviors and bipedal locomotion. *Proceedings of the National Academy of Science*. 115(35): 8799-8804.