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, and language acquisition. The theory of cognitive evolution of grammar (CEG) proves its worth and avoids problems associated with UG. Moreover, it connects grammar theory to the most successful theory of system development and change, namely the theory of evolution, in a new field of application outside of genetics, namely cognition.

1. Background

UG is an elusive subject whose content has changed in unison with the theory of grammar of which it is an axiomatic part. An authorised characterization 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 human genome and Generative linguists have gradually refrained from presenting a testable version of UG over the past years.

That humans are equipped with a cognitive capacity for language processing (in perception, production, acquisition) is undisputed and trivial. We experience it daily. The disputed issue is the claim that UG is the theory of this capacity 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).

In other words, our genome is supposed to somehow code the computational programme of grammars. This 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 therefore ultimately somehow genetically conditioned, and it is very likely mistaken, as the discussion will show.

The hypothesised existence of *innate* and thereby universal principles of the organization of human grammars is not so much an empirical discovery than a theoretical postulate. On the one hand, it is the answer to an induction gap. The Generative grammar that L1 learners apparently manage to acquire is underdetermined by the input. On the other hand, UG is meant to account for cross-linguistic invariants of grammar systems. These two aspects are antagonistic, however. Innate constraints that enhance acquisition 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.

"Innate"[1] is an ambiguous notion, as Scholz & Pullum (2002: 189-191, 2006: 66) explicate in detail. For capturing the linguistic reality, an appropriate reading of 'innateness' seems to be this. It is agreed that there is a small but non-negligible set of grammatical properties that are not 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 by their general problem-solving abilities since linguists have not been able to spontaneously do so either.

This is the point where the Generative 'short circuit' has happened. "*Not acquired bottom-up*" is equated with "*innate"* and this is equated with "*genetically coded"*. This is a precarious way of tackling the problem since it underestimates the real thing. The real thing is the fact that (mental) grammars are put to use and part of our human cognitive system. They run on a neuro-cognitive 'computational platform'. This is a network of computational routines of our brain recruited for language processing and this is the locus of innate properties. Second, in dynamic systems, the parts of the 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. The 'programmer' is – just as in biological evolution – the ongoing process of evolution operating on cognitive programmes. Grammars are products of an evolutionary process, namely a multi-millennial process of cognitive evolution of grammars, as explicated in Haider (2021a,b).

The cognitive capacity for language, with grammar as its core part, is species-specific because only our species has multi-faceted and powerful general cognitive capacities that have allowed for 'subcontracting' parts of these capacities and cross-linking them into a network that – in modern diction – we may call the 'language app'. This app shows properties of genetically-based capacities (due to the involved, genetically based general cognitive routines), but the app itself is not genetically coded although it is domain-specific. It is 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).

This is an important but overlooked aspect in the UG debate. Grammars are learnable since variants of grammars that are difficult to acquire would immediately be 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,[2] each grammar is in continuity with a preceding and a following grammar variant, each of 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

[1] "*Innate behavior is behavior that's genetically hardwired in an organism and can be performed in response to a cue without prior experience*." (https://www.khanacademy.org/science/ap-biology/ecology-ap/responses-to-the-environment/a/innate-behaviors)

[2] As excavations in Jebel Irhoud have revealed (Richter et al. 2017), homo sapiens already settled in North Africa 300.000 years ago.

various processing routines. 'Easier' is the cover term for better adapted features of grammars. Economy[3] is part of it.

The selection mechanism that is constantly at work is the learners' brain 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 resultant adaptation to the selecting environment, that is, to the processing routines of our brain. Simon Kirby justly 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*." (Kirby 2001: 110). As known from biological evolution, the process of variation & selection is able to produce highly complex self-replicative systems and one such family of cognitive systems is grammar systems (Haider 2015a, 2021a,b).

A negative trait of UG theory has been its unproductivity when it comes to generate testable "stunning predictions"[4] that are empirically confirmed. On the one hand, UG provides so many degrees of freedom that there is room for lots of auxiliary hypotheses to block counterevidence. For historic reasons, it is closely tailored to the characteristics of SVO languages. Consequently, other structure systems such as SOV or free-word-order languages (T3 languages) appear as deviations that need to be recaptured by a lot of auxiliary means, but with no answer to the essential question: Why is there so much cross-linguistic variation with additional grammatical ado if a much simpler, UG-streamlined SVO grammar would suffice? The right answer is that this is exactly the picture one expects if evolutionary processes are at work, but not what one expects if there existed a master grammar in the form of UG. Evolution is dissipative since it is fed by random variation.

In each case – the innateness claim and the cross-linguistic universality claim – the sustaining moment has been the absence of a rival theory. There are opponent positions but they are equally underpowered in their predictive potential. This is a situation of the kind that philosophers of science judge as a crisis. There is an acknowledged problem, there are many alternative and incompatible theoretical attempts of removing it, without any broadly recognized consent though. That the UG conjecture can still be upheld is not so much a fact about human insufficiency but an excellent example of scientific conservativism. Scientists do not dump an unproductive theory if the alternative is a situation without any theory at all.[5] Some of the pudding is better than none of the pie and the longer a theory has been retained and bolstered, the stouter it tends to be defended. 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 regulatory principles* of grammars are directly or indirectly dependent on an innate UG; for opponents (not only construction-grammarians, functionalist

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

[4] „*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).

[5] "*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).

typologists, construction grammar, etc.) no *grammatical principle* is innate. This is an unrewarding controversy since the crucial point is not innate vs. not-innate but the explanation of the conditions of the possibility of efficient language acquisition and use.

Opponents tend to overlook an evident fact. If language acquisition were nothing but 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 efficient cognitive abilities. Both parties try to arrive at the grammar of a language by data analysis (rather than genetic guidance). Half a century of dedicated research in grammar theory has not produced an accepted compendium of UG or a grammar based on it.[6] How would children individually manage a task with much less elaborate problem-solving capacities than cooperating professional linguists? Patently, general problem solving does not seem to be the whole story either.

The present situation is one of 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 maze; 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 available to a child. After all, problem solving is known to be much more dependent on general intelligence than language acquisition. If the UG thesis and the general problem-solving capacity thesis are inappropriate approaches, what is a more promising explanation? 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 established branch of research flourished and kept on developing one of the most successful scientific theories that deals with the same kind of problems that linguists deal with, namely the problem how complex systems emerge, how they are represented in an individual, 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. It is a law of nature that complex systems do not materialize in a flash.[7] 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).[8]

Previous approaches to language evolution often seek the evolution of language(s) in *biological* evolution or miss the point completely.[9] It is evolution, but evolution in a *neuro-cognitive* setting (Haider 2021a,b, 2015a), that is, *cognitively based* evolution. The target of cognitive evolution is the "language app", which is a domain specific program package in the ensemble of

[6] The first and last attempt has been Stockwell & Schachter & Hall-Partee (1973).

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

[8] This is not only true for all forms of life but seems to be true also for the cosmological dynamics of the universe (Smolin 1992).

[9] "*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, a very different matter*." (Chomsky 2011: 265.)

our cognitive computational capacities. The structure and much of the content of this package is a product of cognitive evolution, that is, the present result of processes of variation & selection (plus drift and the analogue of gene transfer). The selective environment is the subset of general brain routines recruited for language processing, that is, for the actual usage of the app. Nonetheless, the evolutionary perspective on grammars is just the general Darwinian evolutionary perspective on dynamic systems, but applied to a neuro-cognitively based system.

Expressed in computer terminology, the acquisition of language & grammar is not a program upload or an auto-installation process. It is learning-by-doing of something that has been custom-fitted to our human cognitive capacities in the course of a multi-millennial cognitive evolution of grammars.

2. How much UG?

Over the years, appropriately adjusted conceptions of UG have shadowed the various models of Generative Grammar. In the time of the Principles & Parameters model, UG has been regarded as a full-fledged master grammar. Individual grammars are seen as particular 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 differences. In the absence of a theory of parameters, the continuously growing number of parameters eventually amounted to the very same kind induction-gap dilemma that UG theory was meant to solve (see section 3.5).

The Minimalist Program, as a complete relaunch, goes together with efforts of arriving at 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. However, there is evidence 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 Minimalist Program is truly minimalist but also a total about-turn. It sacrifices the original motivation for its postulation, namely the explanation of the induction gap and the 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).[10] Evidently, the Minimalist UG theory does not answer the question 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.

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

3. UG versus CEG

In this section, a number of characteristic implications and deducible predictions of the UG theory will be confronted with counter positions in the theory of CEG, the cognitive evolution of grammars. Table 1 lists and contrasts ten areas in which the hypothesis of CEG clearly stands out against the hypothesis of an innate universal grammar. The table serves as a menu for the following presentation.

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, ...)
9. changes = parameter re-valuation	changes = evolution
10. complete & consistent grammar	partially incomplete & undefined grammar

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

3.1 Genetically versus cognitively grounded

UG is supposed to be genetically coded, as (Chomsky 2017: 3) claims. Immediate evidence is missing, however. 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. "*FOXP2 might be important for anatomical features contributing to derived human traits, including speech and bipedalism*." (Xu et al. 2018: 8799). Atkinson et al. (2018: 1424) note that "*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.*"

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. Language abilities evidently do not play a role in

behavioural genetics.[11] In the past 50 years, genetics has made impressive headway, but linguists have not been part of it, although they would have, if an innate UG had qualified as a prime target of behavioural genetics.

Even if it is a fact that many innate capacities[12] 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 since it is not the way complex systems come into being. If they reach 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[13] in a single[14] species (see De Boer et al. 2020). Whoever boldly claims the opposite deserves credibility only if (s)he is able to produce solid, compelling and immediate 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 explanation for cross-linguistic invariants and effective acquisition of grammar suffers from the *confirmation bias*.[15]

In the evolutionary perspective, the grammar systems we see within our very narrow time horizon of less than 4k years are the products of an ongoing process of cognitive evolution. This means that they are products of a process 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 languages we could not learn could not have come into being. The evolutionary approach to grammars as products of a process 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 owe to

[11] "*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

[12] For instance, categorical perception is an innate capacity involved in phonetics-to-phonology mapping. It is not primate-specific, though. Chinchillas, monkeys, chicken or rats dispose of it, too (Kriengwatana et al. 2015).

[13] 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.

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

[15] 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 possibilities. (Jermias 2001).

Charles Darwin. UG theory claims to be genetically grounded, otherwise the appeal to innateness could not be upheld. Unfortunately, this claim has never been empirically validated. Therefore, this theory is in default.

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 perceive as cross-linguistic invariants aka language universals. Is this a satisfactory explanation for language universals? It would be satisfactory if we were able to trace a causal chain between a genetically determined neuro-cognitive property and a linguistic property. This is not the case however. We do not know the neuro-cognitive processes underlying the computing of language nor do we know their genetic genesis. UG-based 'explanations' lack empirical substance. They appeal to an explanatory background whose existence is merely postulated. No causal chain is known and no experiment has confirmed any link of it.

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 its linguistic applications. Crucially, universals are emergent, they are not programmed in. The processing brain evolutionarily shapes grammars by sieving out less adapted variants. The major traits of variants that pass are the traits we observe cross-linguistically and construe as language universals.

Here is a biological example. Mammals have entered the sea as a habitat at several occasions. 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. 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) insists 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 is better than its factual accuracy. 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 poor when it has to apply operations to linearly

structured representations. The inversion of a list, for instance, is a computation that is easy to program[16] but hard to carry out mentally. No grammar uses list inversion as a grammatical rule.

Structure dependence is just one side of the medal, with recursion as the other side. Structure dependence is recursive chunking. Those who regard recursion as a typical feature of human grammars should not be surprised that grammatical rules are structure-dependent, rather than linearization-dependent. 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, this is a domain-general capacity that is also operative in all modalities of human information processing, such as vision, action planning, event perception (Lashley 1951, Martins et. al. 2016), and also in language processing. So, we should not be surprised at all that grammatical rules operate on categorized chunks (aka phrases) rather than on serial properties of sequences of terminals.

In the meantime, it has been confirmed that the handling of recursion in symbol processing is not an exclusive achievement of homo sapiens. Contrary to what is claimed by proponents of UG, other species are able to identify recursive patterns, too. Crows do it, as Liao et al. (2022) have shown (see sect. 2). Structure-dependency is just another way of describing the fact that rules of grammar operate on the level of chunks and not on the level of serial terminal elements. This is not surprising and does not justify invoking an innate UG.

3.3. Closed vs. open system space

The history and future of a given grammar is dynamic, open ended, and vectored. This is what the theory of evolution predicts and this is what we see. UG theory, as a closed system, predicts small, oscillating changes, with swift returns to the UG defaults 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 and a subset of properties stemming from parametrized conditions. UG theory necessarily comprises 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 seem to exhaustively partition the system space. There is the purely *structural* 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 structural & 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

[16] 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.

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 (e.g. Is there a genitive of negation in Mandarin? Is there a Dative subject in Afrikaans? Is Dative a lexical case in intransitive passive in English?) since the learner has no chance to meet decisive data. So, there must be defaults for 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 opposite.

The evolutionary view of CEG is confirmed by what we observe. There is no pre-specified closed system of variability.[17] 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, but they are rare, just like 'exotic' life forms in biology.

3.4 Parametric

In grammar theory, 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 question. In Slavic languages, *multiple* wh-fronting is the rule. Germanic languages provide room for only a *single* wh-expression in the clause initial position. In yet other languages, wh-items remain 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 of different outcomes of grammar evolution. For UG, all these variations (macro- and micro-parameters) are an embarrassment. If it were true that human grammars closely follow an innate blueprint, a single invariant structure would suffice and be appropriate.

Parametrization seemed to be an elegant way of capturing the embarrassing amount of variation that has to be covered by a theory of universal invariants. 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 match 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:

[17] This is the place for an aside on Cartographic Syntax, which maps UG 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 vertebrates, with parametrized features. Looking back at grammar cartography in this perspective should be an eye opener.

"*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?[18] 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.*"

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 set to 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 the 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 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):

(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 whose function is the licensing of the null pronoun "*pro*" in Spec T. The situation is not so simple, however. It is true that German is no null-subject language (18b), but like many other languages, it uses a null-subject construction in a particular type of finite clauses, namely in imperatives (18a). In Minimalist terms, there must be a D-feature for T in German but it cannot 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. Jetzt denkt (ihr) einmal nach!
now think$_{2\text{-pl.}}$ (you$_{pl.}$) once about!
b. Jetzt wundert *(ihr) euch, weil *(es) heute nicht regnet.
now wonder (you) yourselves because (it) today not rains

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

[18] Let us assume, just for the sake of a concrete estimate, UG would involve exactly 100 binary parameters (which is a modestly low 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.

history of homo sapiens. Grammars develop and they develop in dissipative ways. However, the direction of change is not the direction towards a UG-based invariance or uniformity.

UG theory predicts exactly the opposite, for the following reason: The UG system would be the conservative moment for variation and changes since it is the attractor that corrects 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 once, 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*."

For UG, the problem situation is the inverse. 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. What appears to be parametric is merely 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 two different languages, as in the case of language families that are continuations of a common 'mother' language. The other source is *convergent evolution*. It makes grammars of 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.[19] The learner has to identify the appropriate value for each parameter and set it, whatever this may mean for a 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 complexity 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, however, since before a child can set the parameter, (s)he must be able to trace its effect in the data in order to determine its valuation. However, 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.

[19] Not only would it be fully consistent with the innate hypothesis, but it would also be expected that parametric differences are innate. The grammar of the waggle dance of honeybees is innate and the parametric differences too. (Rinderer & Beaman 1995; Kohl et al. 2020).

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 end up in cluelessness. A sentence is not labelled for its parameter values. A well-formed input sentence is the 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 with a wrong setting. In short, complex parameter-setting is intractable for a child.

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.

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.[20] The acquisition of V2 in Germanic languages is not instantaneous but a lengthy process. 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' patterns just become fewer over time and eventually disappear. Parameter-setting would be an inappropriate description for this kind of development. It is a lengthy elimination process. This

[20] 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.

is not parameter-setting but pattern identification and use.

In concluding, a particularly remarkable parameter idea deserves to be called up by its oxymoronic name, the "*expletive null subject*", as sub-parameter of the null-subject parameter. The idea that something empty has to serve as a filler for a position that must not be empty sounds like scientific gobbledygook. Nevertheless, it is a much-discussed concept in Generative textbooks. "*Some languages apparently allow expletive null subjects, but not referential ones. German is one such language*" (Roberts & Holmberg 2010: 8). German is certainly no such language since there exists no such language at all, see Haider (2019), Haider & Szucsich (2022a sect. 2.1) and (2022b, sect. 4). A null expletive exists only in the Generative universe of ideas (see below).

How would a child find out that there is something where there apparently is nothing? An "expletive null subject" is a grammatical 'neutrino' without any content, neither lexical nor phonological nor semantic. It does not interact with anything. How does it betray its existence? Its existence could be ascertained only by inference from another allegedly innate ingredient of Generative Grammar, namely the "*Extended Projection Principle*" (EPP), according to which a grammar must ensure that every clause has a subject. A child has to 'know' this and will kindly infer from parent's utterances such as (3) that the subject of such sentences must be a silent expletive since their parents would not constantly commit a grammar mistake by uttering a subjectless sentence. (3a,b) are commands in the form of passives of intransitives.

(3) a. Jetzt wird geschlafen!
now is slept
b. Hier wird nicht herumgetanzt!
here is not danced-around

Already at an early age, children are able to describe some of their feelings and they would convey this grammatically correct for instance by (4), which are frequent utterances. Again, they instinctively have to 'know' that they *must not* use the expletive form they have learnt to use with weather verbs, namely 'es' (it), although this would satisfy the EEP.

(4) a. Mir ist kalt
me$_{Dat}$ is cold ('I am cold')
b. Mir ist schlecht
me$_{Dat}$ is sick ('I am sick')

Moreover, children do not overgeneralize an overt expletive they encounter in the sentence initial position of V2-clauses (5a), namely *es* ('it'), as in (5c). Why would they not overgeneralize? They would if they 'knew' the EPP. But in reality, they do not overgeneralize because they are never confronted with an obligatory, purely structural subject position in the language they learn, viz. the OV language German. A clause structure with an obligatory structural subject is the marked case and it is the hallmark of SVO languages. In SOV languages and in T3 languages there is no such position. Proponents of the EPP seem to be unaware of the fact that not a single SOV language is known that abides by the EPP.

(5) a. *Es* hat vorhin jemand angerufen.
EXPL a has -little-while-ago somebody called

b.*Vorhin hat *es* jemand angerufen.
a-little-while-ago has EXPL somebody called
c. Vorhin hat jemand angerufen.
a-little-while-ago somebody has called

Roberts & Holmberg (2010), and in fact many other syntacticians, even those who are native or specialized in a Romance null-subject language, continue to 'keep their eyes wide shut' to the fact that in *no Romance null subject language*, intransitive verbs can be passivized with the standard aux + participle passive[21] (see Haider 2019). In total and permanent disregard of this fact, Roberts & Holmberg (2010: 13) present the following putative inclusion relation for the sub-parameters:

"*We can range them along a scale of "liberality" as follows: expletive null subjects ⊃ partial null subjects ⊃ consistent null subjects ⊃ discourse pro-drop*.

Any Romance null-subject language is a *consistent* null-subject language, but none of these languages would allow an intransitive passive. However, they would clearly have to admit it if the ominous expletive null subject existed. (6) and (7) are instructive minimal pairs. In fact, Romance null-subject languages ought to be a showcase of expletive null subjects. The expletive pronoun of French (6) would be subject to the null-subject parameter and end up as null expletive in Italian and Spanish (7).

(6) a. Il a beaucoup été fumé dans cette salle — Gaatone (1998: 124)
it has much been smoked in this room
b. Il a été dormi dans ce lit — Rivière (1981: 42)
it has been slept in this bed
c. qu'il a été procédé à cette arrestation — Le Figaro, Sept. 7, 2016
that it has been proceeded to this detention

(7) a. *È stato ballato in questa sala — Italian
has been danced in this hall condone
b.* È stato dormito in questo letto
has been slept well in this bed
c. *Fue trabajado duro aquí. — Spanish
was worked hard here

There are Romance varieties that have managed to recruit a subject expletive of the adverbial type and consequently they passivize intransitives, as expected, as for instance in Vèneto.[22]

(8) a. *Z'è* stà parlà de ti — Vèneto
there has been spoken about you
b. *Gh'è* stà parlà de ti — [variant of (8a)]

The idea of an "expletive null subject" is a theoretical fiction (Haider 1987). It is the correct inference from a wrong premise, namely the premise that every sentence contains a subject. The EPP is not a universal of sentence structure but the type-defining and predictable property

[21] This is apparently ignored because English (as a singular exception among the Germanic languages) does not passivize intransitives (because of the unavailability of a suitable expletive) either.
[22] I am grateful to Cecilia Poletto, who is native in this language, for checking the data.

of [S[VO]] languages (Haider 2015). For SOV languages and T3 languages, a structural subject position is not an obligatory part of the clause structure. They allow for genuinely subjectless clauses, for instance when intransitive verbs are passivised.

In the evolutionary perspective, parameter is a meta-theoretical concept. The linguistic observer who compares grammars will notice related patterns across languages and identify bifurcation points in the evolution of grammars. This is qualitatively the same situation as the situation of a zoologist who notes that bats have developed wings out of their arms and seals have developed paddles. Evidently, these are not parameters of a UG of body shapes but anatomical homologies that developed in the course of evolution. The analogous processes of convergent evolution are the explanatory background for cross-linguistic invariants.

3.6 Mono- or poly-genetic grammars

When Ken Hale shared his field-work results, a new parameter was born, namely the (non)-configurationality parameter (Hale 1983). This was a capitulating move. Warlpiri was too exotic from the perspective of an English-centered UG, and so it was quarantined away. The new starting point for the learner in the decision tree of parametrization was [± configurational], where [-configurational] is the complement set of English-like languages.

Indo-European studies, on the other hand, have confirmed the evolutionary concept of a common origin followed by dissipative evolution. The research time depth is necessarily shallow. It is 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. But homo sapiens did not start using language simultaneously with inventing script on clay tablets in the south-eastern neighborhood of Europe. So, for most of the relevant time span we do not have any data at all.

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 perspective[23] 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[24] continent, polygenesis of grammars during and after such a "*Me Tarzan – You Jane*" epoch is more probable than monogenesis of UG. It is no outrageous assumption that modern grammars are the result of cognitive evolution starting from a one- and two-word phase. In Nichols (2011: 572) words "*language originated gradually over a diverse population of pre-languages and pre-language families*".

The cross-linguistic invariants of modern languages are the expected reflex of *convergent* cognitive evolution by constant cognitive selection of grammar variants by the invariant neurocognitive processing resources that constitute the human language-processing facility. 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.

[23] The following paragraphs are based on Haider (2023).

[24] 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."*

What circumstances would give rise to a UG with lots of parameters if a UG without parameters would do the same job? UG theory predicts highly narrow channels of diachronic change, with UG as the constantly strong attractor and gate keeper. The opposite is the case. If UG theory were right, languages "*varying radically from one another*" should not exist.

3.7 (Im-)Perfection

Generative 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 general trajectory 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 next permanent on-stage appearance is in theology, in the ontological proof[25] of the existence of god (see Crittenden 1968) and now 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 an allegedly empirical discipline of 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 can claim anything, but then, this is the same type of claims as Dr. Pangloss' claim of perfection in Voltaire's Candide (see Weiss & Dunsworth 2011). No biologist would have claimed 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. Nobody is keen to find out why biologists do not assume a "strong minimalist thesis".

Let's assume for the sake of the argument that UG is 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 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. Here is an example from English. English is the only Germanic language that cannot passivize intransitive verbs

[25] "*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).

because of the unavailability of a suitable expletive (Haider 2019). As a consequence, there is no *perfect* match at the semantics interface for 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 gains a selectional advantage if it is more efficient than other imperfect solutions. A little bit of vision is better than no capacity of vision at all. Moreover, variation & selection does not guarantee permanent progress. Often, a system ends up in a globally suboptimal, local maximum (Kauffman 1993: 43). Wright (1932) formulated a powerful visualization of adaptive evolutionary changes, namely the concept of adaptive landscape, which became 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, extremely 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 structurally 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 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 lower in the course of time. [S[VO]] will be the steadily growing peak, as agued in Haider (2023). [S[VO]] languages are outcomes of grammar changes and not their starting point.

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 usually misfile as SVO; see Haider (2023). Ob-

viously, any one of these languages represents the present-day outcome of a potentially millennial developmental history. Each 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 morphology. 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). 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 variant ends up as a member of the sieved-out ones. Hence, there is no constant oscillation between the original variant and its successor variant.

Finally, changes are expected to progress along lineages, and they do (Dunn et al. 2011). In linguistics, we perceive this as clusters of changes in language families, as for instance the emergence of V2-languages in the Germanic family. This syntactic property emerged in the documented history, with a single exception, namely English. In English, the change shaped only the wh-clause formation but did not reach the declarative clause structure. It was counteracted by another, specifically English change, namely the immovability of the 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 mistake an SVO sentence as a V2-sentence. In English, this would be a mistake, but in all Scandinavian 'sister' languages, this is the rule:

(9) 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 not embody a theory of grammar change, which is the gist of the evolutionary approach. The theory of evolution is a theory of change. In the Generative 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 should 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 proceed.

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.

3.10 Consistency & completeness

Informally, the two basic notions are 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, based on the regime of UG in each 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 theorem proving (see sect. 2).

This is not what we observe in 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.

Let us begin with consistency. Bech (1963: 295, 297) was the first grammarian to name and explicitly analyse a grammatical inconsistency in the grammar of German. "*Grammatical laws in contradiction*" is the translated title of his publication. He noticed a systematic rule conflict in German in the interaction of infinitival syntax with the syntax of the verbal complex. The two conflicting "laws" are the following ones. First, in IPP constructions (= infinitivus pro participio, aka Ersatzinfinitiv), the auxiliary *haben* ('have') must not follow a modal (10a) since this would trigger the participial form of the modal, which is not admitted. Therefore, the auxiliary is fronted (10b).

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

Second, the infinitival marker *zu* ('to') must occur on the final verb. A clash of the two conditions is easy to predict. If (10b) is transformed into the infinitival variant (11a), the infinitival

marker would go together with haben (‘*have’*), which is the finite form of (10b) transformed into the infinitival form (11a), 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).

(11) a. *ohne das *zu haben* beendigen müssen[26]
without it finish to have must$_{Inf.}$
b. ?ohne das *haben* beendigen *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, (11b), that is, the variant of the 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. It 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 continuously more and more abstract modellings of grammars. With the switch to the Minimalist program, UG has been turned upside down but only by reversing the perspective. Although Chomsky (2007: 7) describes 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).”*

The essay is a discussion of assumptions in present-day grammar analyses 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-

[26] 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 (27) is V$_{Inf.}$ – modal – have, not the base order V$_{Partic.}$– have – modal, which is a different cluster, with a different meaning. Note that the main verb in (27) is infinitival, not participial, which shows that it depends on the modal and not on the auxiliary *haben*.

linguistically ascertained specific linguistic facts. The promised approach "from bottom up" is missing in 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) neuroscience, 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 vehicles 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[27] (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 Generative 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.

In passing, one more dogma will have to be ditched, namely the dogma of the equi-complexity of grammars: "*To put it briefly, languages are different forms, historically modified, of a single cognitive capacity, language."* (Mendívil-Giró 2020: 2). This capacity is UG, and therefore there are no more- or less developed languages.

This reasoning rests on a wrong axiom, as I have tried to show in this paper. *Reduction* of the declarative complexity of grammars is an evolving variable. It is the effect of adaption to the

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

limits and strengths of the recruited computational capacities of the human brain. The consequent but mistaken conclusion drawn from the wrong axiom is this:

> "*We cannot accept that the evolution of language can be explained as a succession of linguistic changes that would convert supposed primitive languages or prehuman languages into the human languages that we speak now*." (Mendívil-Giró 2020: 3).

This statement is reminiscent of Max Müller's notorious Rubicon-attack on Darwin, which may be judged leniently, given what was known about evolution and about grammar systems in the 19th century:

> *Where, then, is the difference between brute and man?* [...] *I answer without hesitation: the one great barrier between the brute and man is language.* [...] *Language is our Rubicon, and no brute will dare to cross it.* (Müller 1862: lecture IX).

How could one know for sure what has been the linguistic ambience of homo sapiens 300.000 years and more ago? Mendívil-Giró merely rephrases the fear that the assumption of a series of evolutionary steps in the evolution of grammar would be equal to labelling previous stages as more primitive. They are indeed more 'primitive' in the etymological sense of the term (i.e. first or earliest of its kind), but they are not languages of less intelligent or old-fashioned people. 'More primitive' may be abused by racists, but then it is abuse. 'Less complex grammar' does not mean 'less complex thinking' nor would, as assumed in the old days, language change mean continuous decay, which was a common opinion in the 19th century.[28]

The grammar of English is arguably *less complex*[29] than the grammar of Latin. But of course, reduced grammatical complexity does not negatively influence thinking or writing. Ironically, dogmatic believers in equi-complexity seem to fear that modern languages are considered to be more prestigious because of *increased* complexity than ancient languages, or languages spoken by people who live in the Brazilian or North Australian tropical forests. In fact, it is exactly the other way round (Haider 2023).

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 declarative network to the 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

In the view from outside, UG is just a hypothesis and its merits are judged like those of any

[28] *"It is true that both the Schlegel brothers and Humboldt subscribed to the growth-decay model*." (Bynon 1986: 132)

[29] A serious discussion would need an agreed definition of complexity. One could adduce a combination of space complexity (*How much memory space does it take to compute*?), time complexity (*How much time does it take to compute*?) and Kolmogorov complexity (*What is the length of the shortest binary program that represents the given grammar*?) Needless to mention that nobody knows how to precisely measure these factors for languages.

scientific hypothesis. The dominant view from neighboring disciplines ranges between indifference and disbelief, see Edelman & Christiansen (2003). The view from inside is the converse. Generative grammar without UG would be but a highly abstract way of looking at grammars, packed with empirically insufficiently supported or tested, complicated conditions and derivations. Without UG as scaffold, any Generative grammar of a given language, with its numerous covert elements and operations, is empirically overly underdetermined and speculative. Even if a given analysis may be highly implausible, precarious or ultimately wrong,[30] it is accepted if it can be argued to follow from, or add to, some of the already accepted axioms and 'theorems', or if it is needed for maintaining empirically challenged core assumptions. If someone dares ask how such an abstract system with its complex derivational procedures could possibly be acquired, the baffling answer is (a 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.[31]

> "*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 check whether this is a psycho-linguistically feasible[32] 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 in 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 Lasnik & Uriagereka (2002: 149), UG is not more than "*a rational conjecture*" that "*children come equipped with a priori knowledge of language*". Merriam Webster (on-line) defines "knowledge" as *the fact or condition of being aware of something*. What would children be aware of, given the a priori knowledge attributed to them by UG advocates? Let us nevertheless grant a reading of 'know' that Pullum & Scholz (2002: 17) and Scholz & Pullum (2002: 187) characterize as "innately-primed learning", even if it may be unclear how it works.

---

[30] Here is a taster: *Obligatorily covert pied-piping* [sic!] of phrases, 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* does not hold [see Haider (2018) for data and discussion].

[31] 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?

[32] Labelle (2007) argues that human (not to mention juvenile) short-term memory capacities are far too limited for computing the complex structures that current the Minimalist theory presupposes.

Since the UG debate more often than not focuses on highly abstract issues, let's be concrete and consider tangible data, as for instance some of the *complex* cross-linguistic wh-in-situ constraints in languages with wh-fronting (see Cheng 2003), and next, the apparently *simple* head-complement serialization patterns of verbs and other phrasal heads, and finally the notorious, often-invoked *that-trace*-Effect. It will be shown that we do not have to resort to an innate UG for understanding the possibility of acquisition of even such fairly abstract and fine-grained grammatical properties.

5.1 Innate or emergent (I) – In-situ '*why*' and '*how*' in [S[VO]] languages

In languages like English, that is, in uniformly head-initial SVO languages, the adverbial wh-elements *why* and *how* display an exceptional behaviour (1a,b). They are not tolerated in situ. This contrasts with all OV languages, such as German, Dutch, or Japanese (Haider 2010: 118-120), and all T3 languages, such as the Slavonic languages (Haider and Szucsich 2022a: 14-16), as illustrated in (12c.-f.).

(12) a. Which materials did he use therefore/for this reason/**why*?
b. It is easy to guess *which words* they will spell incorrectly/**how*.
c. Welche Materialien hat er *weshalb* verwendet?[33] — German
which materials has he *why* used
d. Dit hangt dus af [van [wie het *hoe* definieert]].[34] — Dutch
this depends thus on [of [who it *how* defines]]
e. Dare-ga naze soko-ni itta no? — Japanese
who$_{Nom}$ *why* there-to went Q$_{-PARTICLE}$ — (Saito 1994: 195)
f. Mne interesno, kakuju poezdku Maša kak dolgo planirovala — Russian
me interests which journey Mary *how long* planned

Multiple wh-constructions are so rare[35] that reference grammars hardly ever mention them, and if they do, they do not dwell on details.[36] Nevertheless, adult speakers of English avoid utterances such as (1a,b), but German, Dutch, Japanese or Russian speakers do not hesitate to accept the corresponding patterns. Obvious questions arise. First, could the situation be contrariwise (i.e. ok in English but unacceptable in the non-SVO languages), and second, how do English speakers find out that they must not replace "therefore*"* or "incorrectly" in (12a,b) or in (13) by a bare wh-pronoun, although this is a reliable recipe for multiple question formation otherwise?

(13) a. Who would therefore/**why* remain without a job in Formula 1 next season?
b. Who has stupidly/**how* tripped an alarm?

"*Speakers have never encountered such a pattern* before" would not be an acceptable answer,

[33] https://www.wohnklamotte.de/trends/creative-mind-tim-labenda-ueber-sein-brandspace-im-stilwerk/

[34] https://www.waterkant.net/suriname/2020/02/01/adhin-100-miljoen-gebruikt-om-ondermeer-aardappelen-en-uien-te-kopen/

[35] Grebenyova (2006:160-161) checked child-directed speech in the Russian CHILDES corpus and counted 697 single wh-questions but only a single double question. For English, she found three multiple wh-questions in the set of the first 5000 wh-questions. This means a *type* probability of 0,001 and 0,0006, respectively, in toto. The per-child ratio is even smaller.

[36] Here is a Dutch example. The following quote is all you find in the Dutch reference grammar ANS: "*Opmerking – In één zin kan ook meer dan één vraagwoord voorkomen.*" (Translation: Note – There can also be more than one question word in one sentence.). Haeseryn et al. (1997: 1430).

of course. Multiple wh-questions are rare and therefore the chance to encounter them without the help of search engines is minimal. Rare constructions would 'die out' by and by if using them presupposed previous encountering. Such a situation is a typical situation for invoking UG and appealing to Plato's problem,[37] with the well-known mantra: English speakers know the grammar of their language better than they could have inferred from input. They rely on innate constraints with the consequence that in-situ *why* and *how* are identifiable as ill-formed.

Things would be much simpler if all languages were like English, but they aren't. So, why don't Dutch or German speakers eschew constructions that English speakers shy away from? Moreover, English speakers are even trapped in ineffability since the only way out of the dilemma is to break up sentences such as (13) in two, as in (14), whenever the preceding wh-item is the subject, which 'competes' for the clause-initial position, too.

(14) a. Who would remain without a job in Formula 1 next season and *why*?
b. Who has tripped an alarm and *how*?

Obviously, the constraint that operates in English is not a universal constraint, and this aggravates the problem. How could an English speaker 'know' that (s)he may leave "*what*", "*where*" or "*when*" in situ but not "why" or "how", and why is this constraint absent in many other languages? And, finally, we would like to understand how language learners arrive at this level of competence. This is the kind of put-up-or-shut-up situation for proponents of an innate UG. It contains all ingredients of the UG debate, from Plato's problem to no-negative evidence. Nevertheless, the parameter-based UG account is in a forlorn position here. The distribution is parametric, but in a Generative account, the allegedly innate parameter is hopelessly complex:

According to Huang (1982) and Lasnik & Saito (1984),[38] '*why'* and '*how'* would have to be covertly fronted after spell-out and this type of fronting is constrained in a particular way. Even if this were true, it is far beyond the limits of anything plausibly innate. Moreover, it does not apply to languages like Dutch, German, Japanese, or any Slavonic language. Fortunately, we do not have to belabor this point since there is an account that works and that does so without appeal to any innate UG principles.

It is not accidental that English or Swedish[39], to call up another SVO language, behave alike in this respect, and that Slavic languages, German, or Japanese are different, and that it could not be the other way around. What is the UG-based explanation? Is the English situation the default or the exception?

In the "a priori knowledge" perspective of UG, languages like English or Swedish are the default which means they are taken to be closer to UG, because Generative Grammar has been modelled mainly on the evidence of English, Romance and North-Germanic languages. But, if English is the default, why don't speakers of other languages show due respect? Or is English the 'marked' situation, and what would this mean? Evidently, UG must be compatible with

[37] = the problem of explaining how we can know so much, given our limited experience. (Encyclop. Britannica).

[38] These accounts presuppose the Principles & Parameters model. The Minimalist Program is silent on these facts and the phenomenon is not covered since ECP and 'head-government' are not part of the MP model anymore. A *minimal link condition* on covert movement would not work because of the many languages that 'opt out'.

[39] Evidence from corpus search is easy to adduce, and it shows that (i) is well attested but (ii) is absent:

i. Vem ringde och varför?	ii.*Vem ringde varför?	(Swedish)
Who called and why	who called why	

both situations, but how do the learners find out what applies to their own language? They cannot find it out directly. There must be something in the input from which a learner can infer it. But what is the triggering input, given that multiple wh-constructions are extremely infrequent (with the exception of Generative syntax textbooks)?

The reality is complex but not inscrutable. First, '*why'* and '*how'* are *adverbial* pronominal wh-quantifiers. As described by Szabolcsi and Zwarts (1993), '*why'* and '*how'* are interrogative quantifiers of a *higher-order* type (i.e. higher than first-order).[40] Second, adverbials are syntactically combined with the phrase that matches their semantic type (Haider 2004). This is a property of the incremental construction of semantic representations.

For higher-order wh-adverbials, the minimal domain is the domain that contains the main verb with its event variable (Davidson 1980). Therefore, for a language with head-initial VPs, the smallest phrase they merge with is the V°-initial VP, as we see in languages such as English or Swedish, since the wh-adverbials need to c-command the main verb as the donor of the event variable. Consequently, *post-verbal* in-situ positions are inadequate for adverbial wh-*quantifiers* of this semantic type[41] in SVO languages. Syntactically, the post-verbal position is available for any adverbial but the efficient semantic-construction fails at the structure-semantic interface if a quantifier of the type of a sentence adverbial is ill-positioned.[42]

Second, the pre-VP position in head-*initial* phrases is no position for *phrasal* adjuncts of a head-initial VP or any other head-initial phrase, unless the head of the adjoined phrase is adjacent. (Haider 2022a). Pronouns are not heads of phrases but lexical pro-forms for phrases. Consequently, the preverbal position is blocked for '*why'* and '*how'* and any other wh-words. In VO languages, the only licit position for *why* or *how* is the sentence-initial position, and this is what we see in English or Swedish, and any other [S[VO]] language. In OV languages, for principled reasons, the VP is not compact (Haider 2004; 2010: 12). Hence, adverbials are free to occur VP-internally, and each position c-commands the verbal base position. Hence, sentence adverbials and higher-order quantifiers are well-formed in 'low' positions.[43]

If this sketch is correct, where exactly is UG involved and how can UG become operative here? The crucial property of the primary input is the difference between patterns such as (15a) and (15b). In the 15,5 milliard NOW corpus, one does not find a single token of "*has with much greater regularity*", followed by a verb, although the phrase "*with much greater regularity*" is well represented otherwise. A Google search (August 25, 2022), restricted to 'books', produces the same result: not a single token of "*has with greater regularity*" but 293 tokens of "*has much more regularly*" and 2430 hits with "*has more regularly*". The head of the adverb phrase in (15a) is adjacent to the VP, the head of the PP in (15b) is not. This is the superficial indicator

[40] Note that technically, "*for which reason*" or "*which way*" are *first*-order quantifiers since they range over e-type variables (viz. particular reasons and ways). As first-order-order quantifiers, they are well-formed in-situ (Reinhart 1998: 31).

[41] Here is an example: „*When did they start what/where/*how/*why* (**unfortunately*)?"

[42] It is easy to check the interactions between position and interpretation as a domain-effect. In (i), the final position and the immediate pre-verbal is inadequate for the interpretation "*It was clever that ....*". The interpretation as sentence adverbial must have the (trace of the) finite verb in its domain.
i. (Cleverly) they would (cleverly) have (cleverly) placed their adverbs (cleverly).

[43] i. Er hat [manches *vielleicht* salopp formuliert] – he has [somethings *perhaps* casually phrased]
Source: https://www.vol.at/offener-brief-fluechtlingshelfer-wehren-sich-gegen-mateschitz-kritik/5230886

of a violation of the Left-Left-Constraint (LLC), explicated in Haider (2022a).

(15) a. Hollywood has [more *regularly*] depicted images of ....
b.* Hollywood has [*with* greater regularity] depicted images of ....

Viewed from this angle, the learner has enough information in the primary input for finding out that *phrases* do not *freely* occur in immediately pre-VP positions (except for parentheticals, which can occur in virtually any phrasal interspace). So, wh-phases cannot appear there either.

Let me duly emphasize that the argument from '*no negative evidence'* does not apply here. It is applicable only to details of *infrequent* constructions. For otherwise highly frequent cases, the absence of a specific pattern *does* legitimate the inference on the learner's side that the absence is due to a restriction. This is a matter of frequency-based inferencing: If a phrase type is frequent, as in the case of PP adjuncts, and it occurs alternatively in various places, as e.g. in clause-initial and in post-verbal positions, then the probability is very low that its total absence in a particular position is purely accidental. Therefore, for the learner, it safe to interpret such a lacuna as the effect of a barred position for phrases that are not properly attached to the VP. It is this property that immediately applies to wh-pronouns since they are phrasal pro-forms, void of head-specific lexico-*syntactic* information associated with heads.[44]

Let us recapitulate: An apparently complex distribution of in-situ wh-items reduces to two simple facts. First, adverbial *quantifiers* are placed in positions that are amenable for the appropriate incremental semantic-construction plus variable binding. This is an interface condition of syntactic structure and semantic interpretation. Second, the position of adjuncts of head-*initial* phrases needs to be licensed like any other position. Since the immediately pre-VP position is outside of the directionality domain of the head of a head-initial VP, an adjunct needs to be 'properly' attached (s. "proper attachment" in Haider 2022a), which is the background of the LLC constraint.

The situation in OV and T3 languages is noticeably different. Here, the positions of adjuncts are within the canonical directionality domain and therefore, no special means of licensing is necessary. This correctly accounts for the facts. The '*why & how'* phenomenology is one of uniformly head-initial languages. These languages, however, provide enough direct evidence for the pre-VP *void* of phrasal adjuncts. The bottom line is that an effective learning strategy does not need to be overly sophisticated in order for the leaner to be able to grasp and understand this pattern. Crucially, there is no need at all to appeal to a *UG principle* and in particular to something like "*ECP for covert movement*".

Let us invert the perspective now. It is very easy to propose an unlearnable principle for covering such patterns and to claim that, given that such a principle cannot be inferred from the data, it must be part of an innate UG. This is just a walking-on-hands approach and no valid argument

[44] Another, often overlooked source of feedback is repeated but slightly reformulated utterances, as Chouinard & Clark (2004: 667) have found out: "*Our findings show that adults reformulate erroneous child utterances often enough for learning to occur. Their reformulations are found for all kinds of child errors – errors of phonology, morphology, syntax and word choice*."

but understandable. A theory developed mainly on the basis of SVO languages needs to be augmented with all kinds of auxiliary hypotheses to make it work for OV.[45]

The really surprising moment, though, is the fact that advocates of UG unceasingly defend principles that have been postulated on an affirmative-only basis. In Generative grammar, in contrast to the established scientific practice, consequent falsification checks are very rare. Principles are proposed, bolstered with selected affirmative evidence, and maintained without thorough falsification testing. Even if counterevidence for the mainstream analysis of wh-constructions and their constraints has been 'on the market' for quite some time (e.g. Haider 1986, 2000a, 2010), it tends to be neglected or explained away in a case-by-case manner with ad-hoc auxiliary assumptions. The epistemological parallels to Ptolemaic epicycles are very close.

The above discussion deals with an example of a frequent fallacy, namely the argument from complexity, well-known from theology (Barash 2013). The logical fallacy materializes when the lack of understanding of a complex matter is adduced as sufficient evidence for the claim that the complex fine tuning of grammar is impossible without intelligent design. In our case, the intelligently designed[46] system is UG.

5.2 Innate or emergent (II) – The structure of (verb) phrases

Let us return to the superficially simple case of head positioning in a phrase. In the one-dimensional array of items at the phonetic interface, can only precedes its dependent or follow. This is what we see cross-linguistically. There are phrases with the head obligatorily preceding ('VO'), obligatorily following ('OV'), or optionally preceding, following or coming in between its dependents, respectively (Type-3).

In many languages, heads of all categories are serialized uniformly. On the other hand, there are also languages in which the order relation is sensitive to the lexical category of the head. English is a uniformly head initial language. German is a language with category-dependent positioning. VPs and APs are head-final, NPs and PPs[47] are head-initial. In Mandarin Chinese, noun phrases are head final but verb phrases head-initial. In any case, the difficult task for the learner is not so much the detection of the head-complement order but the identification of the phrase structure that a grammar superimposes on the linearly ordered items. Here is an example.

What is the *internal* structure of an English VP with two objects? It is easy to realize that depending on the theoretical background, different structures are assigned to a phrase like (16a), namely (16b-d). This shows that the task of identifying the phrase structure of a double-object VP is not trivial.

(16) a. send someone a present

[45] Some seriously meant and widely accepted proposals read like a satire: Not only native SVO syntacticians are happy with the assumption that *every* language allegedly is basically SVO. Other clause structures are mere 'distortions' of the SVO base. In uniformly head-final languages, for instance, all phrases are deemed to be obligatorily fronted across the head, but nobody is alarmed that syntactically, they do not behave like phrases in derived positions, since these properties are not critically checked. Relevant checks fail. (Haider 2013, ch. 9).

[46] "*The arguments for INTELLIGENT DESIGN* [emphasis mine]$_{HH}$ *from irreducible complexity bear an uncomfortable similarity to that originally posited for the necessity of a genetically specified universal grammar."* (Finlay 2009: 262).

[47] Less than a handful of relational particles are genuine post-positions in German. Most of them are re-analyzed lexemes, as for instance "*des Geldes wegen*" ('the money due-to'; lit. in the ways of the money) or 'der Prawda *zufolge'* ('the Prawda according-to'; lit. in-pursuance).

b. [$_{VP}$ send someone a present] n-ary branching, 'flat' VP [Pollard & Sag 1994]
c. [$_{VP}$ [send someone] a present] binary left-branching VP [Chomsky 1981: 171]
d. [$_{VP}$ send$_i$ [someone [e$_i$ a present]]] binary right-branching VP, with 'VP-shell'

How is it that children do a task better than professional grammarians? An appeal to UG would merely be a paraphrase of "*Don't ask, it's innate*." This is uninformative since the implicit argument is closely circular (17):

(17) a. Premise 1: Structure Ψ is determined by UG.
b. Premise 2: If UG determines structure Ψ, language acquisition converges on Ψ.
c. Therefore: Language acquisition converges on structure Ψ.

But how can we be sure that premise 1 is true? We assume it since there is evidence that for any native learner, acquisition converges on structure Ψ. Evidently, this is no explanation but only a circular paraphrase. The evolutionary perspective is more revealing. In this perspective, Ψ is the result of processes of variation and constant sieving out (aka selection). The structure of complex, head-initial VPs that is compatible with the relevant evidence is a version of (16d). Here is a sketch of the evolutionary path towards the syntactic structuring of verb phrases that eventually leads to (16d).

The essential step of processing is that linear arrays of terminals are exhaustively[48] chunked. In principle, there would be two ways of *uniformly and exhaustively* chunking a phrase, namely (18a) or (18b). One of these options is sieved out since it is inferior in on-line processing. It is easy to understand why this is (18b). The bracketing betrays it.

(18) a. ... [x [y [z [(...)]]]] (right-branching)
b. ... [[[[x [y [z [(...)] (left-branching)

In (18a), the structure of the already processed part does not change if additional material has to be integrated while proceeding. In (18b), however, the processor would have to look ahead already at the beginning of the phrase since any additional material at the end changes the structure at the beginning w.r.t. embedding. Processing is incremental. It does not wait. So, under (18b), structure assignment either had to halt until the end of the phrase or backtrack. These problems are avoided in structure (18a).[49]

The next step is the efficient identification of the head of a phrase. In the right-branching structure (18a), there are three potential positions for a head, namely (19a-c). If a language makes available all these positions simultaneously, it is a Type-3 language (Haider 2015b). Slavonic languages are representative of this type (Haider & Szucsich 2022a,b) and head-positioning within the phrase has been perceived as free in such languages.

(19) a. [$_{VP}$ x [y *h*°]]
b. [$_{VP}$ x [*h*° y]]
c. [$_{VP}$ *h*$_i$° [x [e$_i$ y]]]

[48] In syntactic terminology, this characterizes a binary branching phrase structure.
[49] The claim that left-branching is the universal way of phrasal structuring and evidence for it have been presented first in Haider (1992) and in subsequent publications (1997, 2010, 2013, 2015b) under the heading *Basic Branching Constraint* (BBC).

A fixed head position is an advantage for the predictive accuracy in processing. Evolution of grammar has led to grammars with fixed head positions. The initial evolutionary nudge of a fixed position favors (19a). The dependent elements are step-by-step joined with the head or its projection. The result is a head-final structure in which the head marks the end of the phrase. Many languages are head final. As simple as the structure is, it has one small downside for the processor. The head comes late. Processing would benefit from an early presentation of the head since it provides access to all the top-down information associated with the head of the phrase. The mirror image of a head-final structure would be left-branching.[50] So, it is excluded.

The early presentation of the head in a structure with more than one object carries a price, namely a more complex structure. The head must be instantiated twice (19c). Only in this structure are both objects within the appropriate directional range of the head. Diachronically, this structure is a late development, as the diachrony of English and Scandinavian languages tells us (Haider 2014). It is the successor of a T3-structure or a head-final structure with secondary head-movement (as for instance the verb-second phenomenon of Germanic languages).

Let us return to the initial question, namely the learnability of the structure of complex head-initial phrases. Do we have to invoke UG? Does UG impose these particular shell structures and would the learner arrive at totally different structures without UG? The answer is "no" in both cases. The learner has to rely on the available processing routines, and they work in a particular way. This way is reflected in the resulting structures. The technical details of phrase structure processing with their particular consequences have to be elucidated by neuro-computational linguists. Presently, we know hardly anything about our brain software and how it is put to work on the general neuro-computational 'platform' of our brain. What we can safely assume is that is has particular properties and these will be reflected one way or the other in grammars. The properties of the processing routines are not specifically linguistic, even if linguistic structures are the main domain of their application. These are routines of our pattern processing brain that have been recruited for this task and thereby have become essential for language processing. They have shaped grammars in the ongoing cognitive evolution of language structures.

5.3 Innate or emergent (III) – *That-trace* Effect

Depending on the theoretical background, the that-trace effect tends to be viewed either as a grammatical restriction or as a by-product of sentence processing.[51] "*The core pattern, though easily stated, remains a conundrum. It is evident in numbers of unrelated languages, but also appears to vary even in closely related languages*." Cowart & McDaniel (2021: 258). The "easily stated" core pattern in English (20a) is a constraint on wh-subjects extracted from a C-introduced clause, with an exception noted already by Bresnan (1977: 194 fn.), namely the intervention-effect of adverbials (20b). The effect applies to long-distance relations as in (20a,c)

[50] Originally, when the author detected the "*no-left-branching*" universal (Haider 1992), he declared it a UG property (Haider 1997, 2000b, 2010, 2013, 2015a). In the evolutionary perspective (Haider 2021a,b; 2022), however, this finding turns out to be gratuitous. It is an *emergent effect* in CEG, the cognitive evolution of grammars.

[51] Schippers et al. (2020: 128) conclude: "*We argue that our data is best explained by assuming COMP-trace effects are in essence processing related.*" Sobin (2002: 542) reports acceptance rates up to 64% for that-t-cases in acceptance tests.

and is absent in local relative clause formation (20d).

(20) a.*What$_i$ does he think [*that* [e$_i$ affects everyone]]?
b. What$_i$ does he think [*that* [in this situation e$_i$ affects everyone]]?
c.*This is the incident [0$_i$ [that I think [*that* [e$_i$ affects everyone]]
d. This is the incident [0$_i$ [*that* [e$_i$ affects everyone]

The *that-trace* effect is a restriction on extraction from a special structural position, namely the *functional* subject position of a clause. It is the very position that matters. For this reason, the effect is observed only in [S[VO]] languages like English and in other, structurally similar languages, also outside of the IE-family.[52] These are languages with an obligatory structural position for the subject outside of the VP. In OV languages as for instance Dutch (21a) or German[53] (21b), an empty subject position following a complementizer does not affect grammaticality,[54] as the following book-corpus excerpts illustrate. The VP-external, obligatory functional subject position is the defining property of the [S[VO]] clause structure. No other clause-structure type (SOV, VSO, T3) has it, whence the absence of the that-t-effect in such languages (Haider 2010: 88, 128; Haider 2015b).

(21) a. Wie$_i$ denk je [dat [e$_i$ hem gestuurd heeft]]?
who think you [that him sent has]?
b. Wer$_i$, glauben Sie denn, [dass [e$_i$ den Ausführenden sagt, was sie tun müssen]]?
who think you PARTICLE [that the performers$_{Dat}$ tells what they do must?
'Who do you think tells the performers what to do?

What is the relevant accessible evidence for native learners of English? First of all, a learner finds out that for a large class of verbs, the complementizer '*that'* is optional for clausal objects. Second, the learner finds out that the complementizer is obligatory for subject relative clauses (22a) and strongly preferred[55] for object clauses with a fronted PP (22b). Consequently, the complementizer is retained in subject relative clauses with a local antecedent and in clauses with a fronted PP.

(22) a. This is an incident *(*that*) affects everyone.
b. He thinks ??(*that*) in this situation such an incident affects everyone.

So, what remains as the heart of the problem is this: In SVO languages, but not in SOV languages, long-distance wh-extraction of a subject is avoided in the presence of the complementizer. The relevant structures in (23) provide the cue. For the avoided structure (23a) there is a simpler alternative, namely (23b). Instead of stacking two functional projections, each with a chain link in its spec-position, there is only one functional projection with one link in (23b).

[52] Pesetsky (2017) refers to Nupe (Nigeria) and Wolof (Gambia and Senegal).
[53] Corpus data for this construction can be found already in Paul (1919: 321f.), in the subsection named "Satzverschlingung" (sentence convolution) of his German grammar.
[54] This does not exclude that there is a preference for the complementizer-less variant in general in German (and also in English, as (Cowart 1997: 19) observed), but this preference does not differentiate between subject- versus object-extraction.
[55] Doherty (2000: 15) is even stricter and stars the following examples:
a. She prayed *(that) [next Wednesday] the check would arrive.
b. They concluded *(that) [in the future] he should be closely watched.
c. They maintain *(that) [in Dublin] good coffee is hard to find.

In SOV languages, there is no functional layer for the subject, hence no functional projection can be saved.

(23) a.*[Who$_i$ ..... [$_{FP2}$ e$_i$ [*that* [$_{FP1}$ e$_i$ [$_{VP}$ V° .....]]]]] [56] *that-t* configuration in SVO
b. [Who$_i$ ..... [$_{FP}$ e$_i$ [$_{VP}$ V°.....]]] acceptable variant in SVO
c. [Who$_i$ ..... [$_{FP}$ e$_i$ [that [$_{VP}$ e$_i$ .....V°]]]] that-t configuration in SOV

At this point, we are confronted with the following alternative. Either the grammar of English (and UG) is assumed to contain a complex constraint against "[e$_i$ [that [e$_i$ [..." configurations (see Douglas 2017)[57] or a strong preference[58] for the less complex structure (23b) is the effect of processing economy. If we are dealing with a processing effect, then this is what we expect when a grammar as a product of cognitive evolution is put to use. It is subject to computational selection effects of varying degrees. Müller (2015) justly emphasizes that processing properties do not play a significant role in UG, and genuine economy conditions that are grounded in processing are not part of the UG theory, as underscored by Collins (2022).

In sum, the three cases discussed above indicate that there are bottom-up ways of arriving at the relevant properties without appeal to an innate UG and its alleged guidance. Grammatical properties may appear to be inaccessible for a learner if viewed in the top-down perspective of a Generative grammar of a given language, but this is not the perspective of the learner and user.

Grammars are not implanted into our brains by a programmer and language acquisition is not a program upload. Grammars are products of cognitive evolution and thereby necessarily well-adapted to grammar-learning and grammar-using human brains. The UG conjecture is but a deus-ex-machina solution for a hitherto ill-understood problem. Children do not acquire a complex software package for language processing, neither by guidance of an innate universal blueprint of grammars nor in a trial-and-error behavior guided only by commonsense problem-solving capacities. Children acquire a grammar system which has been tailored to the brain capacities in a millennial[59] process of cognitive evolution of grammar (Haider 2023). At the end, the task may be as simple or complex for a child as selecting the glove suiting the left or the right hand when presented with a pair of them.

6. Summary

In the direct confrontation of the two hypotheses, that is, the hypothesis of CEG (cognitive evolution of grammars) and UG, 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,

[56] "FP" stands for *functional projection*.

[57] "*These effects arise from Spec-to-Spec Anti-locality interacting with systematic variation in the degree of articulation of the C-domain in clauses and RCs with and without* that*."* (Douglas 2017: 22).

[58] Corpus data show that *that*-t-clauses are in use, especially in written language, but they are infrequent.
i. Who does he think *that* will primarily benefit, other than big business?
ii. What does he think *that* makes him what he is?

[59] It is a fact that homo sapiens has reached the boundary of the Mediterranean Sea 300.000 years ago, as the excavations of Jebel Irhoud prove (Richter et al. 2017). This amounts at least to 25.000 generations of language using and language acquiring brains, with evolution by variation & selection for grammars.

- CEG entails a trivial solution of the problem of *grammar acquisition* since it is the essence of evolution that a system is self-replicating. Properties of grammars are sieved out if they cannot be passed on to the next generation during language acquisition. Grammars adapt to the neurocognitive environment that is operative in acquisition and use.
- CEG predicts the principally dissipative course of *diachronic changes* while UG wrongly predicts developments that converge to default values of UG.
- CEG explains the vectored quality of *grammar changes* and excludes seesaw-type changes that UG would admit.
- CEG explains the tension between *cross-linguistic invariants* and *cross-linguistic variation* as the result of divergence by variation and convergence by the selection environment.
- CEG provides sufficient room for economy conditions in the shaping of grammars.
- CEG opens and ties *grammar theory* to the most successful scientific theory of dynamic self-replicative systems, namely the theory of evolution.[60]

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.

Barash, David, P. 2013. Atheism and Darwinism. In Stephen Bullivant & Michael Ruse (eds.) *The Oxford handbook of atheism*. 414-430. 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

Bresnan, Joan. 1977. Variables in the theory of transformations. In Peter W. Culicover, Thomas Wasow and Adrian Akmajian (ed). *Formal syntax*. 157-196. New York: Academic Press.

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.

Cheng, Lisa Lai-Shen. 2003. Wh-in-situ. *Glot International* 7(4): 103-109.

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,

Chouinard, Michelle M. & Eve V. Clark. 2004. Adult reformulations of child errors as negative evidence. *Journal of Child Language* 30: 637-669.

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

[60] The linguistic match of Dobzhansky's well-known "*Nothing in biology makes sense except in the light of evolution*" is this: *Nothing in grammar theory makes sense except in the light of cognitive evolution of grammars.*

Cowart, Wayne & Dana McDaniel 2021. The That-Trace effect. In Goodall, Grant (ed.). *The Cambridge handbook of experimental syntax.* 258-277. Cambridge: Cambridge University Press.

Craig, Cathy & Goulon, Cedric & Berton, Eric & Rao, Guillaume & Fernandez, Laure & Bootsma, Reinoud. 2009. Optic variables used to judge future ball arrival position in expert and novice soccer players. *Attention, perception & psychophysics*. 71: 515-22.

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.

Doherty, Cathal. 2000. *Clauses without "that": The case for bare sentential complementation in English.* New York: Garland.

Douglas, Jamie. 2017. Unifying the that-trace and anti-that-trace effects. *Glossa: a journal of general linguistics* 2(1): 1–28 (Article #60).

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.

Foote, Andrew D. & Yue Liu & Gregg W. C. Thomas & Tomáš Vinař et al. 2015. Convergent evolution of the genomes of marine mammals. *Nature Genetics* 47: 272-275. https://doi.org/10.1038/ng.31982015.

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.

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

Grebenyova, Lydia. 2006. *Multiple interrogatives: syntax, semantics, and learnability*. PhD-Dissertation. Washington D.C.: University of Maryland, College Park,

Haeseryn, Walter & Guido Geerts & Kirsten Romijn & Jaap de Rooij & Maarten C. van den Toorn. 1997. *Algemene Nederlandse Spraakkunst*. Groningen: Wolters-Noordhoff.

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

Haider, Hubert. 1987. Expletives pro - eine Phantomkategorie. In Dressler, Wolfgang U. & Corrado Grassi & R. Rindler-Schjerve & M. 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.

Huang, James C.-T. 1982. Move wh in a language without wh-movement. *The Linguistic Review* 1: 369-416.

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].

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.

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 & Mamoru Saito. 1984. On the nature of proper government. *Linguistic Inquiry* 15: 235-289.

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.

Mendívil-Giró, José-Luis. 2020. On language evolution. *Inference* 5(3): 1-6.

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

Müller, Friedrich Max. 1862. *Lectures on the science of language*. London: Longman, Green and Roberts.

Müller, Stefan. 2015. The CoreGram project: theoretical linguistics, theory development, and verification. *Journal of Language Modelling* 3(1): 21–86.

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.

Paul, Hermann. 1919. *Deutsche Grammatik*. Vol. III, part IV: Syntax. Halle (Saale): Niemeyer.

Perlmutter, David 1971. *Deep and surface structure constraints in syntax*. N. Y.: Holt, Rinehart & Winston.

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

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.

Reinhart, Tanya. 1998. Wh-in-situ in the framework of the Minimalist Program. *Natural Language Semantics* 6: 29–56.

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.

Saito, Mamoru. 1994. Additional wh-effects and the adjunction site theory. *Journal of East Asian Linguistics* 3: 195-240.

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

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.

Schippers, Ankelien & Margreet Vogelzang & David Öwerdieck. 2020. COMP-trace effects in German: the role of processing. *Nordlyd* 44(1): 117-132.

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.

Stepanov, Arthur & Wei Tien D. Tsai. 2008. Cartography and licensing of WH-Adjuncts: A cross-linguistic perspective. *Natural Language & Linguistic Theory* 26: 589-638

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

Svensson, Erik I. & 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

Szabolcsi, Anna & Frans Zwarts. 1993. Weak islands and an algebraic semantics for scope-taking. *Natural Language Semantics* 1: 235–85.

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.

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.