A farewell to UG and a welcome to CEG

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

Abstract

In this paper, predictions of UG theory and the theory of the cognitive evolution of grammars (CEG) are contrastively tested. For 10 core areas of grammar theory it is shown how the predictions of UG fail while CEG is successful. A prime testing ground for any scientific theory is the predictive power, that is, the ability to generate testable, subsequently confirmed empirical predictions and explanations. The hypothesis of an innate universal grammar (UG) turns out as 'underpowered'. The central predictions fall short. The competing hypothesis presented in this paper – grammars as target of evolution on the level of cognitive systems – is well in line with the empirical situations in language acquisition, diachronic grammar change and cross-linguistic invariants. For ten issues of grammar theory, the competing hypotheses are evaluated, with a clear-cut outcome. The theory of the *cognitive evolution of grammar* (CEG) stands the test and avoids the problems associated with UG. Moreover, it connects grammar theory to the most powerful theory of system development and change, namely the theory of evolution, in a new field of application outside of genetics, namely cognition.

1. Background: Innate UG vs. emergent invariants

An authorised definitory statement 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). However, even if the grammar of each human language meets all the constraints of UG "*languages appear to be extremely complex, varying radically from one another*." (Chomsky 2017: 2).

The crucial ingredients of the two quotes are "*genetically based*" and "*varying radically*". That humans are equipped with a cognitive capacity for language processing (perception, production, acquisition) is undisputed and trivial, since we experience it daily. The disputed issue is the claim that UG is the theory of this capacity and that it postulates a genetically coded blueprint for something that nevertheless "varies radically". In other words, our genome is supposed to code the entire 'computational programmes' of our grammars, somewhere and somehow. This is a much stronger claim than the idea that our 'talent' for language is part of our human nature and thereby species-specific and ultimately somehow genetically conditioned, and it is very likely wrong, as the discussion will show.

As Scholz & Pullum (2002, section 2) explicate in detail, "innate"[1] is an ambiguous term. In one reading – the common reading within the school of Generative Grammar – even details of the make-up of grammars are taken to be innate, as for instance properties that will be examined below, namely the principle that rules out some in-situ wh-adverbials, or the "Extended Projection Principle", that is, the (empirically wrong) generalization (Haider 2019), that in *every* language, a sentence contains a subject, or the *that-trace*-Effect. This is not the place for itemising all allegedly innate items of UG, and it is dispensable since it is not the empirically adequate

[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)

approach anyway as will be argued below. Not only the three exemplary candidates just mentioned do not warrant any appeal to innateness.

For capturing the linguistic reality, the appropriate reading of 'innateness' seems to be this. It is agreed that there is a 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 are not spontaneously able to do so either. This is the exact point where the Generative 'short circuit' happens, however. 'Not learnt' is equated with innate and this is equated with genetically coded. This is a faulty way of tackling the problem, since it overlooks the real thing.

The real thing is the fact that grammars are part of a cognitive system and that they are put to use. 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 only source of 'innate' properties. Second, in dynamic systems, the parts of the system interact in complex ways. What we observe as properties of the output of the system are on the one hand properties *directly determined* by the system and on the other hand, and to a large extent, *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 directly determined ones may be perceived as "programmed in", which they are not. They are reflexes of a *domain-specific* assembly of *domain-general* brain routines.

In grammar theory, we are always dealing with equations involving at least these *two unknowns*. Generative Grammar oversimplifies and decrees that any property of grammar that cannot be acquired bottom-up is "programmed in" and genetically coded. Both, the processual variable as well as the contribution from the recruited routines, is ignored. Apparently unlearnable properties of particular constructions (e.g. that-t-effect in interrogatives, but not in relatives; see below) are elevated to innate properties of an innate UG. Let us note in passing that the focus on specific details of constructions mimics construction-grammar in so far as allegedly innate universals are properties broken down to the level individual constructions and variants thereof.

Let us change the perspective now. First, it is the processual variable and the properties of recruited routines that are responsible for a number of properties that tend to be wrongly analysed as *directly* contributed by the UG-enforced *innate* grammar format. Second, and importantly, grammars themselves are products of a process, namely the process of cognitive evolution, 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 a multi-faceted and powerful general cognitive capacity that has allowed for 'subcontracting' 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 – third – in its present structure and function a result of *cognitive evolution*, that is, of variation & selection. This is an important but overlooked aspect in the UG debate. Grammars are learnable because unlearnable aspects of grammars necessarily have to 'die 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. Every grammar since the beginning more than

300.000 years ago[2] is in continuity with a preceding and a following grammar variant, each of which must have been able to be acquired. This guarantees that grammars are 'purified' of unlearnable and 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 processing routines. 'Easier' is the cover term for more economic and better adapted features of a grammar. 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.

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 result 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 recruited processing routines of our brain. As known from biological evolution, this process is able to produce highly complex, self-reproductive systems. One such family of cognitive systems is grammar systems (Haider 2015a, 2021a,b).

A negative trait of UG theory is its unproductivity on the level of grammar theory when it comes to generate testable "stunning predictions"[3] that are empirically confirmed. On the one hand, UG provides so many degrees of freedom that there is room for lots of auxiliary hypotheses to block counterevidence. On the other hand, 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) are regarded as deviations that have 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 do? 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 is a master grammar in the form of UG.

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 that philosophers of science tend to 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.[4] Some of the pudding is better than none of the pie and the longer a wrong theory has been retained and bolstered, the stouter it tends to be defended. Loss aversion is a strong cognitive bias.

A hindering 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

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

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

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

dependent on an innate UG; for opponents (not only construction-grammarians, functionalist typologists, cognitive psychologists, 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 weaker cognitive abilities. Both 'parties' try to arrive at the grammar of a language by data analysis (and not by genetic scrutiny). Half a century of dedicated research in grammar theory has not produced an accepted compendium of UG or a grammar based on it.[5] How would children manage a task with much less elaborate problem solving capacities than professional linguists? Patently, general problem solving does not seem to be the whole story either.

The present situation is one of denial without alternatives. The 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 a language learner. But on the other hand, they do not offer a compelling demonstration based on empirical evidence that the grammar of a seven-year-old can be acquired by domain general problem solving. Problem solving is known to be much more intelligence-dependent than language acquisition. If the UG thesis and the general problem solving capacity thesis are mistaken 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. If a position is wrong, there usually is counterevidence, and if two positions are wrong, there is counterevidence for both of them. But, even if there is massive counterevidence for one of these positions, this does not warrant the conclusion that the opponents' position is right. It just means that giving up a wrong position is the right thing to do.

A truly amazing fact about the UG conjecture is the long-lasting framing of research as research on something seemingly resembling a computer program for language processing in the spirit of the late fifties, when the first powerful computers came into sight. Generative grammar appears to ignore the advancements in developing cognitively more adequate models of computing over decades, namely parallel and distributed processing (Rogers & McClelland 2014) and continues to model grammars by outdated sequential rule application. Nevertheless, such a type of UG is claimed to be the master program of any mental computation program for human language.

At the same time, an already established sub-field of research thrived and continued to develop further one of the most successful scientific theories that deals with the same kind of problems that linguists deal with, namely 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

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

subfield of population genetics.[6]

It is a law of nature that complex systems do not materialize in a flash.[7] Complex systems develop over time, and they typically develop in a variation & selection way, 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). This is not only true for all forms of life (Darwinian evolution), but seems to be true also for the cosmological dynamics of the universe (Smolin 1992), just as well as for the quantum theory of the propagation of photons (Feynman 1985), to pick out just a few examples. That the dynamics of grammar systems is another field of evolution has been overlooked by mistake. Previous approaches to language evolution either mistake the evolution of language(s) to be *biological* evolution or miss the point completely.[8]

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 our cognitive 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 applied to a neuro-cognitively based system.

2. 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 scientific hypothesis. The dominant view from neighboring disciplines ranges between indifference and disbelief, see Edelman & Christiansen (2003). The view from inside is 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.[9] Without UG as scaffold, any Generative grammar of a given language, with its numerous covert elements and operations, is empirically overly underdetermined. Even if a given analysis may be highly implausible, precarious or ultimately wrong (see fn. 8), 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 an abstract system with its complex derivational procedures could possibly be acquired, the baffling answer is (a paraphrase of) "*It's all innate, stupid*!"

---

[6] It "*is intimately bound up with the study of evolution and natural selection, and is often regarded as the theoretical cornerstone of modern Darwinism*." Okasha (2016).

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

[8] "*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.)

[9] 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].

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.[10]

> "*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 bothers to check whether this is a psycho-linguistically feasible[11] task at all and whether our brain really supports a mental capacity of 'higher grammatical' algebra, with an effectively working theorem-proving component. What Chomsky and followers permanently fail to put on the table is the computational program of the derivational system and the empirical confirmation of the theorem proving abilities of adults and of children (during the parameter fixing phases). They prefer to discuss bits and pieces in isolation that cannot be experimentally checked. To date, nobody has produced any evidence for the (neuro-)cognitive reality of the *theorem proving* capacity. In the above quote, Chomsky denies a *timed* succession of steps in a proof although the Minimalist Program is presented as an entirely derivational system with ordered and thereby timed rule applications, whose timing may even be cross-linguistically variant (see 'spell-out positions').

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) characterize as "innately-primed learning", even if it is unclear how it works.

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, allegedly unlearnable *that-t*-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.

2.1 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-

[10] 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 ability for 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?

[11] Labelle (2007) argues that human short-term memory capacities are far too limited for computing the complex structures that current the Minimalist theory presuppose.

120), and all T3 languages, such as the Slavonic languages (Haider and Szucsich 2022a: 14-16), as illustrated in (1c.-f.).

(1) 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?[12] German
which materials has he _why_ used
d. Dit hangt dus af [van [wie het _hoe_ definieert]].[13] 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[14] that reference grammars hardly ever mention them, and if they do, they do not dwell on constraints.[15] 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 the manner adverbs in (1a,b) and in (2) by a bare wh-pronoun, although this is a reliable recipe for multiple question formation otherwise?

(2) 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, 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,[16] with the well-known mantra: English speakers know the grammar of their language better than they could have inferred from their 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

[12] https://www.wohnklamotte.de/trends/creative-mind-tim-labenda-ueber-sein-brandspace-im-stilwerk/
[13] https://www.waterkant.net/suriname/2020/02/01/adhin-100-miljoen-gebruikt-om-ondermeer-aardappelen-en-uien-te-kopen/
[14] 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 still smaller.
[15] 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_." ('Note – There can occur also more than a single question word in a single sentence.'). Haeseryn et als. (1997: 1430).
[16] = the problem of explaining how we can know so much, given our limited experience. (Encyclop. Britannica).

to break up sentences such as (2) in two, as in (3), whenever the preceding wh-item is the subject, which 'competes' for the clause-initial position, too.

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

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 the Generative account, the allegedly innate parameter is hopelessly complex:

According to Huang (1982) and Lasnik & Saito (1984),[17] '*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.

Let us start with Jaques Monod's topos of 'chance and necessity'. It is not accidental that English or Swedish[18], 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 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* (wh)-pronominal wh-quantifiers. Second, adverbials are syntactically combined with the phrase that matches their semantic type. This is a property of the incremental construction of semantic representations. *Sentence adverbials* are combined ('merged') with something that heads an expression with a denotation that is compatible with the denotation of the adverbial (Haider 2004). For sentence adverbials, the minimal domain is the domain that contains the main verb with its

[17] These accounts presuppose the Principles & Parameters model. The Minimalist Program is silent on these facts and this 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'.

[18] 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	

event variable (Davidson 1980). '*Why'* and '*how'* are *interrogative quantifiers* of the semantic type of sentence adverbials. 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.[19] Syntactically, the post-verbal position is available for any adverbial[20] but the efficient semantic-construction fails at the structure-semantic interface if a sentence adverbial as a higher-type adverbial is ill-positioned.[21]

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 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 are well-formed in 'low' positions.[22]

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 (4a) and (4b). 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 (4a) is adjacent to the VP, the head of the PP in (4b) is not. This is the superficial indicator of a violation of the Left-Left-Constraint (LLC), explicated in Haider (2022a).

(4) a. Hollywood has [much 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 parenthetical expressions, 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 for 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

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

[20] 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.

[21] 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).

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

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.[23]

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 but no valid argument, and I dare profess that I suspect UG to be full of walking-on-hands restrictions that exist only in theory but not in the linguistic reality.

In fact, this is not astonishing. A theory developed mainly on the basis of SVO languages has to be augmented with all kinds of auxiliary hypotheses to make it work for OV.[24] 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

[23] 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*."

[24] Some seriously meant and widely accepted proposals read like a perfect 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. All checks fail (Haider 2013, ch. 9).

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' system is UG.[25]

Here is a non-linguistic analogy. Everyone is familiar with this problem who has ever played a ball game in which you have to judge the trajectory of a ball in motion in the air. Someone throws or hits a ball and you judge where it will come down in order to be there in advance. Check on physics and you will find out that you are judging a ballistic motion, determined by a highly complex interaction of initial speed, launch angle, spin of the ball, temperature, wind, air friction, Magnus force, etc. (Asai et als. 1998). It is obvious that no soccer or tennis player is able to mentally *calculate* the equations for the trajectories but nevertheless, most players manage to solve the task when hurrying to the putative landing site. Would this be evidence for a ball-playing UG? The physicist calculates differential equations, but what does the player do when heading towards the descending ball? Does he rely on an innate DESC ('differential equation solving capacity')? Doubts are appropriate. All that a player has to do is trying to *keep constant the angle* above ground at which (s)he observes the ball (Craig et als. 2009). The player has to speed up if it gets more acute and slow down if it gets wider. In our linguistic example, syntacticians implicitly assume that language learners behave like 'little linguists', but of course they don't, just like ball players (unlike trained artillerists) do not behave like 'little physicists' and do not depend on an innate DESC.

2.2 Innate or emergent (II) – Layered VPs in OV vs. VP-shells in [S[VO]]

Let us turn now to an apparently much simpler case, namely the head-initial vs. head-final architecture of phrases. Do we need UG for a pattern property that is impossible to infer from primary data? Yes, we seem to need it, because the linear array of items needs to be assigned the appropriate *phrase structure*. The learner has to detect this structure. An oversimplified way has been suggested some time ago by Wexler (1998: 29). For first-language acquisition, he proposes "*very early parameter-setting*".[26] According to this conjecture, 'basic parameters' include among others the following: word order, e.g. VO versus OV (e.g. English and Swedish versus German or Dutch) and V2 or not (e.g. German and Swedish versus French or English).

At the age around two, German speaking children figure out that e.g. 'koch-t' (*cook* 3rd sg.), 'koch-st' (*cook* 2nd sg.) or 'koch-en' (*cook* infinitive; or 1st or 3rd pl.) are different forms of the same verb. The differences suffice for telling apart the V_{finite} pattern from the non-finite V-end

[25] "*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).

[26] Wexler (1998: 25) postulates that basic parameters are set at the earliest observable stages, that is, beginning with the onset of multi-word stage. In terms of age this is around 2 years of age, plus minus half a year.

pattern in main clauses. The pattern correlation is simple and a pattern-detecting mind is expected to uncover this early. A context-free probabilistic procedure is sufficient, as Heuser & Tsvilodub (2021) have demonstrated, based on corpus data of child directed speech.

However, Wexler's early parameter fixing is contradicted by data from *embedded* clauses. Here, children use unacceptable main-clause patterns even until the age of four (Tracy 2002:11, sect. 4.5). Obviously, the children have not yet grasped the complementary distribution of clause-initial complementizers and the fronted finite verb. This fact is incompatible with early parameter-setting, which is implausible anyway since the working memory of children at an early age is insufficient for processing utterances with embedded clauses in an adult style. At least by the age of 3-4, when children typically start to master and use embedded clauses, they are ready to note that the complements of verbs of saying and believing are either C-introduced V-end clauses (5a), or complementizerless embedded V2 clauses (5b). So, the primary input is enough for grasping the complementary patterns.

(5) a. Sie glaubt, [*dass* er einen Fehler gemacht *hat*]
she thinks [*that* he a mistake made *has*]
b. Sie glaubt, [er *hat* einen Fehler gemacht]
she thinks [he *has* a mistake made]

Why would English-speaking children not be misled by the verb-second wh-pattern in English and wrongly infer that in the answer to (6a), the wh-item could be replaced by the answer (6b)? After all, this pattern was available until the beginning of the Early Modern English period (Bækken 2002:15), as exemplified by (6c), and it is a standard V2 pattern in Scandinavian languages.

(6) a. *What* has she done?
b.**Nothing* has she done!
c. And *all this* saw and marked the emperour's doughter. [1502, Robert the Deuyll]

Again, the frequency-based reasoning works in finding out that a specific instance of a highly frequent general pattern, viz. fronting of objects in declaratives that is parallel to wh-constructions, is missing. The respective answer item cannot simply replace the question pronoun in (6a,b). What is the primary evidence for the learning child? There is a lot of data available involving PP-fronting, followed by the subject, that provides a sufficient basis for avoiding the V2-pattern (6b), notwithstanding the fact that this pattern is standardly but infrequently in use with downward entailing quantified expressions.[27]

In Generative grammar, the concept of micro-parameters is contemplated for such cases. Kayne (2000) prophesized a privileged role in the future for "microparametric syntax". He does not invoke innateness, though. What "microparametric variation" is meant to refer to is exactly the kind of variation that is characteristic of complex systems and that feeds the gradual progression of evolution by variation & selection. There is no urge for invoking parameterization, however. Variation is an *open-ended* process; parametrization would be a re-positioning in a *closed* system space. Those who favour parametrization do not get tired of proposing any odd parameters. What they fail to do, although this would be the essential part of their task, is to work out a

[27] i. "And *nothing* would he do, *nothing* would he say." (https://mojim.com/usy107863x9x10.htm)
ii. "*Never in my life* would I dare dream that." (source: *Trials & Minor Suffering* by Dawna DeCorby)

theory of grammatically possible and grammatically impossible parameters and parameter values, since this is the gist of UG. Without a theory of parameters and their valuation, anything goes in the inflationary world of parametrization. In the absence of a theory of parametrization, a parameter claim is just a technical paraphrase of the data description.

Let us return to the superficially simple case of head positioning. In the one-dimensional array of items enforced by the phonetic interface, the head either obligatorily precedes its dependents, or it obligatorily follows, or it comes in any of the two orders. There is no other logical possibility. It is not surprising that in human languages these three possibilities are attested, that is, head-initial phrases, head-final phrases and Type-3 (T3) phrases. In other words, there are phrases with the head obligatorily preceding ('VO') or obligatorily following ('OV'), or optionally preceding, following or sitting in between its dependents, respectively (T3). The situation is complicated by the possibility of head-positioning being sensitive to the lexical category of the head. In numerous languages, the verb and the adjective phrase is head-final, while NPs and PPs are head initial. Chinese is an example of a converse setting since it is head-initial for verbs and head-final for nouns.

English is a uniformly head-initial language, that is, any lexical head is an initial head. German is a language with category-dependent positioning. VPs and APs are head-final, NPs and PPs[28] are head-initial. You may call this a parameter but there is no reason to do so. These are simply the admissible patterns in the grammar of German as options in the relevant system space and a language learner won't have problems in identifying a clear pattern in linear order. The real task for the learner is to figure out the phrase structure which a grammar superimposes on the linearly ordered items.

In the *stepwise* combinatorics of the head of a phrase with its dependents, the head-final pattern is the least complex structure. Starting from the linear order, there are only two options for binarily structuring, namely right-branching (7a) or left-branching (7b). The latter structure can be dismissed right away, given that the head of phrase is the element the combination starts with (see the discussion below) since it is the locus of the dependency informations.

(7) a. [α [β [γ h°]]]
b. [[[α β] γ] h°] (inexistent!)

More complex but richer in information for the receptive side of processing (i.e. the parser's side) is the head-initial phrase structure, since it presents the head of the phrase early. The scientific observer, but not the learner, can be misled and assume a left-branching structure (8) for a complex, head-initial phrase, as Chomsky did (1981: 171). (8) would be a kind of mirror image of (7a). A superficial inspection of the word order pattern is enough to realize that a perfect mirror image structure of (7a) would also entail the mirror image order of the dependent elements in (8). However, the base order of the arguments of a verb is identical in head-initial and head-final verb phrases.

(8) [[[h°α] β] γ] (inexistent!)

[28] 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).

Natural languages do not employ the structure-type (8) for complex lexical phrases. This is a cross-linguistically invariant property and therefore, this property is a potential case for UG. How could language learners infer from the primary input that (8) is the wrong structure for an expression such as "*send them the document to their private mail boxes*" 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. Here is a sketch of the explanation.

Structure (8) is a *left-branching* structure, that is, the node on the projection line of the head is a left sister of a merged phrase and the dependents of the head are layered up on the right. Left-branching structures are known to be difficult to process for an incremental (memory-bound left-corner) parser (Jin & Schuler 2020) since the deeper embedded items always would precede the less embedded ones in the input. So, the parser would be bound to constantly guess the number of brackets and then backtrack. Since there is an alternative, more processing-friendly ways of structuring an array consisting of [h°-α-β-γ], a grammar based on right-branching structures (8) would not be able to outrival alternatively available options in the process of language acquisition and the CEG. Here is the structure that arguable is the empirically adequate one (9):

(9) [h°[α [~~h~~° [β [~~h~~°γ]]]

The theory behind (9) is easy to sketch [see Haider (2015b) and earlier work for details], but here, what we go for is an evolutionary explanation for the preference of (9) in the course of language acquisition. Let us start with the theory. Right-branching structures guarantee that a head-dependent phrase, when it is parsed, is higher in the structure than any other phrase-mate that follows. Thus the simultaneous top-down (grammar-driven) and bottom-up (data-driven) flow of information is guaranteed for the parser. The immediately dominating node on the projection line is always the node that dominates the rest of the phrase. For details please consult Haider (2010, 2013, 2015b)

Cross-linguistically, we observe a structural invariant with a basic asymmetry. Phrases are universally right-branching. Left-branching projections of lexical heads do not exist. This is the invariant. The build-up of head-initial structures in (10) follows directly from the interaction of directional licensing and the universal invariant, formulated as the *Basic Branching Constraint* (BBC), see Haider (1992), (1997), (2000b), (2013:3). Given that phrases are universally right-branching (BBC) and, as it is the case for VO, the canonical licensing directionality of the (verbal) head is converse to the direction of merger, a shell structure is the predicted outcome of structuring. (10a) to (10c) are the respective steps of merger:

(10) a. [*explain*$_{\rightarrow}$ everything]
b. [everyone [*explain*$_{\rightarrow}$ everything]]
c. [$_{VP}$ explain$_{\rightarrow}$ [everyone [~~*explain*~~ everything]]]

First in (10a), the lowest argument joins the verb and receives the canonical directional licensing by the verbal head. Then, the indirect object is merged with the structure (10a) in the right-branching structure (10b). Since the canonical licensing direction is to the right, neither the verb nor V' are able to license to the left. So, the verb needs to be re-instantiated for licensing the

indirect object directionally in (10c). This is how the shell structure emerges in complex head-*initial* phrases, and only in head-initial ones. In OV languages, the VP has no shell structure because the verb remains in the foot position of the VP, with all arguments in the directionality domain of the verbal head or its projection nodes.

In OV structures, the head *and* each projection node, that is V', are licit directional licensers and the directionality of licensing is identical with the direction of merger. In VO languages, the projection nodes of the head are on the 'wrong' side, directionality-wise, and this is why the verb has to be re-instantiated. Here are the principles behind these circumstances (Haider 2010: 29, 2015b):

(11) i. Projection lines are universally *right-branching*[29] and *endocentric*.
ii. A dependent phrase is licensed in the *canonical direction* of the head.
iii. The position of a dependent phrase P is *licensed* $=_{\text{Def.}}$ the (projection of the) phrase head h and P *minimally and mutually* c-command each other.

It is the *minimal & mutual* c-command condition (11iii) that is directly causal for the numerous OV/VO contrasts. It is worth emphasizing that it is the *very same* principle (viz. merger is universally to the left) applied under parametric directionality that produces the different outcomes for OV and VO:

(11ii) and (11iii) amount to a *specific locality condition* for the head of the phrase and the dependent phrases. The domain of merger is directionally constrained and in this domain, the merged phrase and head (or a projection node of it) always c-command each other. In OV, this is trivial. The merged phrase is always on the directionally appropriate side of a sister node that is a directionally licensing node of the projection spine of the head. In OV, the directionality of merger is the converse of the licensing directionality. The only way to achieve canonical directional licensing and mutual c-command is the re-instantiation of the verb head in the resulting shell structure (9).

This is the simple, theoretical side. It is simple because of the huge amount of empirical data available for checking and re-checking (11). But what is the corresponding side in the brain of a language-acquiring child? First, we don't know, and second, we have to find it out. If our information processor seems to prefer particular data structures, this means that other data structures are less easy to process. And this means furthermore, that the way the processor works is constrained in a particular way. It arguably works as a left-corner parser. So, the parser is the source of priming for structure assignment by the learner.

Take for example the inversion of lists of words and try to repeat a given sentence backwards: "*Modest doubt is called the beacon of the wise*" ⇒ "*Wise the of beacon the called is doubt modest*". Evidently, our computing mind does not provide a handy routine for it. So it is not part of any grammar. On the other hand, our computing mind ecells in manipulating structures by replacing parts of it. A German-speaking person has no problems at all to produce *many*

[29] In other words, the direction of merger *within* a phrase is universally to the left. This particular claim has been presented first at an international conference in Utrecht (Dec. 1991) and published first in Haider (1992).

alternative beginnings for a sentence such as (12a), that is, for replacing "*er*" in (12a), as for instance (12b) or (12c), with all the tricky[30] collateral effects.

(12) a. Er glaubt anscheinend, sie *damit beeindrucken zu können*, wenn er …
He thinks apparently you *it-with impress to can* if he ...
'Apparently he thinks to be able to impress her if he ...'
b. *Damit* glaubt er anscheinend *sie* …..
with-it thinks he apparently her ....
c. *Beeindrucken zu können* glaubt er *sie* anscheinend …
impress to be-able believes he her apparently

If in doubt, ask a computational linguist and (s)he will confirm that list inversion is technically a much simpler task than generating variants of a German clause with different beginnings. Technically it may be simpler, in 'brainy' terms however it is not. This example is meant to point out that we must duly consider a factor that is not properly honoured in grammar theory, namely the restrictions on human information processing routines. This knowledge is essential since these routines act as the selection filter for grammar variants in the course of the cognitive evolution of grammars. As a consequence, grammars turn out to be adapted to the processing conditions.

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. It is this way that 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 tasks and thereby have become essential for language processing. They have shaped grammars in the ongoing cognitive evolution of language structures.

2.1 Innate or emergent (III) – *That-t* Effect

A summary in a nutshell is Cowart & McDaniel's (2021: 258) „*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*." What the *that-trace* constraint is meant to rule out is the shaded configuration in (13a), illustrated in (13b) and originally observed by Perlmutter

[30] The variant (12c) involves a 'tricky' moment of German infinitival constructions. In this example, the entire infinitival verbal cluster is fronted. This is possible only in the optional clause-union variant, but not in the construction with an embedded infinitival clause constituent. The telling reflex is the positioning of the pronoun, *preceding* the matrix sentential adverbial 'anscheinend' (apparently).

(1971). The constraint is operative only in the lowest CP (13c), and it is overridden by interacting effects. One overriding moments is the effect illustrates by (13d), as analysed by (Browning 1996: 238), the other is the effect in relative clauses (14e, 15b).

(13) a. [Wh$_i$ .... [$_{CP}$ e$_i$ [complementizer$_{C°}$ [e$_{i\text{-subject}}$ [...]]]
b. What do they claim (*that) is covered by UG?
c. What do you think (that) they claim (*that) is covered by UG?
d. What do they claim (that) *in the Minimalist Program* is covered by UG?

The interacting effect in (13d) is the fact that the complementizer is not dropped if the subject is preceded by a fronted adverbial; see Doherty (2000: 15). The case (13d) is the combinative effect of wh-subject-extraction and adverb fronting. The latter requires a complementizer (14) and this requirement evidently overrides the that-t-Effect. The same kind of overriding effect emerges with wh-movement in a C-introduced relative clause (14e). These data have been raised against a *that-t* filter-condition already in the 70ies by Joan Bresnan.

(14) 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.
d. John claims *(that) [during the party] Ted squirted water at Eric.
e. This food is still cooked in the way^i [e^i [*(that) [e$_{i\text{-subject}}$ [is prescribed in ancient books]]]]

In both instances – with a fronted adverbial as well as in relative clauses – dropping the complementizer is banned independently of wh-movement and the ban is obviously a parsing effect. In (14a-d), the complementizer systematically disambiguates the domain of the adverb (matrix or dependent clause), and the same happens with relative clauses (15b). Without the complementizer, '*the solution*' in (15b) gets mapped on the structural template for subjects and produces an intolerable garden path effect.

(15) a. This is the solution (that) they propose
b. This is the solution *(that) surprises everyone

In both cases – with adverb preposing and with subject-relating relative clauses – the *that-t*-Effect is absent because *that*-drop would be unacceptable in these structures in general. This fact has a crucial implication: If the ban on *that*-drop is a parsing effect, then the that-t-effect must involve a parsing component as well, otherwise it could not be overridden.

In many languages, such as German, the *that-t*-Effect is absent. In his collection of examples on "Satzverschlingung" ('sentence intertwining'), Paul (1919: 321-323) lists excerpts that involve *that-t* configurations[31] (Haider 2010: 88). A broad experimental study by Schippers et als. (2020: 128) concludes:

"*Although COMP-trace violations in German are not categorically ruled out, they are associated with decreased acceptability*. [...]. *We argue that our data is best explained by assuming COMP-trace effects are in essence processing related, which has led to the exclusion of COMP-trace configurations in certain languages, including English*."

[31] Google searches prove positive as well: Wer$_i$ glaubt er denn, [[dass$_{C°}$ [e$_{i\text{-subject}}$ ihm […] sehr viel Geld im voraus für eine Luxusuhr überweisen wird]]]? [Aug. 30th 2022: https://forum.watchtime.ch/viewtopic.php?t=69845]

The "*which-has-led*" conjecture is noteworthy. If this is correct, it is an example of a tipping-point constellation in the evolution of grammar. This is a constellation in which a *quantitatively* characterized distribution reverts into a *qualitative* one, in the form of a discrete grammatical constraint. What we see is that speakers avoid the complex structure *t-that-t* in (16b) and prefer the simple structure[32] with traces in the base positions only (16a). In (16c) however, the simple structure without a complementizer is dispreferred, since it produces a strong garden path configuration, in which the preceding DP would get analysed as the subject of a finite clause.

(16) a. Who$_i$ did he tell everyone [e$_i$ [will win]]
b. Who$_i$ did he tell everyone [e$_i$ [(*that) [e$_i$ will win]]]
c. This is the design [e$_i$ [*(that) [e$_i$ will win]]]

Taken as a whole, there is the following alternative. Either the English grammar (and UG) is assumed to contain an extremely complex constraint[33] against "[e$_i$ [that [e$_i$ [..." configurations (see Douglas 2017) or the avoidance of these structure is a processing effect, with a strong preference for the simpler structure. If, however, the *that-t*-Effect is indeed deemed to be part of that portion of English grammar that is implicated by UG, then the proponents are urgently summoned to put up or shut up. What they should put up is the exact UG constraint and a plausible way of how its genetically coded version is expressed in the brain of a child who is about to acquire English. Everything short of it is 'hot air'. If, on the other hand, 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 neuro-computational selection effects of varying degrees. Müller (2015) justly emphasizes that processing properties do not play any role in UG, and genuine economy conditions that are grounded in processing are not part of the UG theory, as underscored by Collins (2022), too.

In sum, the cases discussed above are sketches but even these sketches are less imprecise than the UG conjecture with respect to the innate quality of these three cases. We do not need to invoke UG if we are confronted with complex details of grammars. Grammars have not been implanted into our brains by a programmer.[34] They are products of cognitive evolution and thereby 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 do acquire a grammar system which has been tailored to the brain capacities in a millennial[35] process of cognitive evolution of grammar (Haider 2022b).

---

[32] String vacuous movement structures are superimposed (Haider 1989). Instead of a CP-TP-VP structure, a CP/TP structure is projected, that meets both requirements. In passing, this explains the absence of *do*-support.
[$_{CP}$Who$_i$ [will$_k$ [$_{TP}$ e$_i$ [e$_k$ [$_{VP}$ e$_k$ win]]]]]? in superimposed structure: [$_{CP/TP}$ Who$_i$ [$_{C'/T'}$ will$_k$ [$_{VP}$ e$_i$ [e$_k$ win]]]]?

[33] "*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).

[34] 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 appears to be custom-fitted to our human cognitive capacities in the course of a multi-millennial *cognitive* evolution of grammars.

[35] It is a fact that homo sapiens has reached the boundary of the Mediterranean Sea at least 300.000 [sic!] years ago, as the excavations of Jebel Irhoud prove (Richter et als. 2017). This amounts at least to 25.000 generations of language using and language acquiring brains and at least 25.000 opportunities of cognitive evolution by variation & selection of grammar systems, starting with "*Me Tarzan, you Jane*" and gradually arriving at "*Then Tarzan picked up the carcass of Horta and came up the slope to Danny, who knelt openmouthed and petrified.*"

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.[36]

3. Innate UG versus ongoing CEG

In this section, a number of characteristic claims and implications of the UG theory will be confronted with counter positions in the theory of CEG, the cognitive evolution of grammars. They are listed in Table 1 and will be discussed in the given order.

[Table 1]	Universal Grammar	Evolution of grammar
1.	genetically grounded	cognitively grounded
2.	universals by innateness	universals by convergent cognitive evolution
3.	parametric	parameter-free
4.	closed system space	open system space
5.	parameter valuation required	no parameter valuation
6.	mono-genetic grammars	potentially poly-genetic grammars
7.	perfection	imperfection (‘bricolage’)
8.	‘SVO’ by default	adaptive landscape (SOV, T3, 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

The UG hypothesis is a take-it-or leave-it parcel. One school uses it as defence shield against accusations of invoking fictitious and unlearnable qualities in the complex machinery of Generative grammar. The other schools have joined in attacks but not in devising a more productive alternative position. Here is a position that has not been considered in the debate. This position is the evolutionary perspective on grammars as *cognitive systems*. 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 detailed presentation. The resulting ‘balance sheet’ is unequivocal. In contrasting the two approaches, the fundamental deficits of the UG hypothesis appear in their unveiled gestalt.

In the past, the debate on UG resembled a combat, with attackers and defenders. The Duhem-Quine thesis has taught us that a theory cannot be defeated like a mediaeval castle or refuted like a mathematical proof.[37] Theoretical claims cannot be confirmed or falsified on their own, in isolation from surrounding hypotheses. These surrounding hypotheses can be modified in order to avert particular counter evidence in each case. Defenders of an empirical theory *always*

[36] Apparently simple things may be complicated, also in the case of dressing, see *dressing apraxia* (=def. “*A group of cognitive disorders characterized by the inability to perform previously learned skills that cannot be attributed to deficits of motor or sensory function*.” (USArad Medical Dictionary).

[37] (Harding 1976: x): “*The physicist can never subject an isolated hypothesis to experimental test, but only a whole group of hypotheses.”*

can block each attack with (ad hoc) auxiliary assumptions. Theories are generally not defeated or refuted; they are left behind for a better theory.[38]

3.1 Why are neighbouring fields disinterested in an innate UG?

If UG were innate, it would be genetically coded and hereditary. UG allegedly covers all aspects of language, that is, phonology, morphology, syntax, semantics and pragmatics. The respective language networks involve 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 since decades.

It is amazing in this context to read Haworth et als. (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*."

It is not so much the content of the message that is amazing but the fact that not a single linguist has been involved. Language abilities evidently do not play a role in behavioural genetics. If I were a nativist, it would be my prime concern to adduce positive genetic data for the fundamental claim that UG is hereditary and seek the cooperation with behavioural geneticists. Instead, office doors have been re-labelled with 'biolinguistics'. Linguistics is not primarily "bio", it is primarily "cognitive". In the past 50 years, genetics has made impressive headway, but linguists have not been part of it, although they should have, since an innate UG would have to be a prime target of behavioural genetics.[39]

Even if it is a fact that many innate capacities[40] are involved in language acquisition and use, the UG claim is much stronger. It claims that the *whole* neuro-cognitive grammar system with all its details and conditions of operation is innate and governs the child's build-up of grammar by innately-primed learning. This is a bold and fruitless claim since this 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 in their evolution. During this time, sub-species develop into species and bud into new species. For homo sapiens, it is a biological fact that such a complex system as UG cannot be the product of a singular fulguration event[41] in a single[42] species. Whoever boldly claims the opposite deserves credibility only if (s)he can produce solid, compelling and

[38] Even if a new position later turns out to be true, Max Planck's dictum remains true, too: "*A new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up that is familiar with it",* which is paraphrased as "*Physics advances one funeral at a time*". And so does linguistics.

[39] "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

[40] 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).

[41] 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.

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

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 (Pullum & Scholz 2002). Presently, UG as explanation for cross-linguistic invariants and effective acquisition of grammar is founded on speculations and not on facts, and it fatally suffers from the *confirmation bias*.[43]

In the evolutionary perspective, the grammar systems we see within our very narrow time horizon of about 3000 years are the products of an ongoing process of cognitive evolution. This means that they are products of a 'steadily grinding' 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 transmission, even more so than in the biological counterpart. Each generation has to acquire the grammar from being exposed to the linguistic outputs of their environment. Selection is the effect of sieving out. The variants that survive 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 in language acquisition and processing.

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 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, the theory is seriously in default.

3.2 Universals by innateness vs. by evolution

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

Here is a biological example. Mammals have re-entered the sea as a habitat at several occasions. If animals change their habitat, the factors of selection change and we note the effects of adaptation to the new habitat, ranging from fish-like forms (see dolphins), to transformations of limbs into paddles (see seals), or merely webbed feet, as in the case of water hounds or otters. For example, 50 millions of years ago, in the course of roughly 15 million years of evolution, the processes of evolution-by-selection have produced life forms that look like fish (see dolphins or whales). The message of this excursion is that nobody would call for a "universal

[43] 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).

grammar" of sea life in order to explain 'universals' of marine life forms. The theory of evolution is sufficient. The cross-species invariants are emergent. Biologists call this "*convergent evolution*" (Foote et als. 2015).

In fact, in grammar theory, invariants are rare but it is full of cross-linguistic *variation,* just like in the case 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 interrogative clauses with room for 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 footprint 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. A *single* invariant structure would suffice and be condign for an innate blueprint of a human grammar.

3.3 Parametric

Parametrization seems 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.

It has soon turned out that the number of parameters grows with every language adduced. Macro-parameters got supplemented with micro-parameters. How many parameters would UG have to provide?[44] In Generative Grammar, there has never stuck out a strong ambition of presenting a theory of possible and impossible parameters. Still, 'parameter' is just 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 the open slot in an otherwise invariant, universal principle. The UG principles were deemed to be universal and the cross-linguistic differences were characterized as differences in terms of the valuation of the parameter of a given universal principle. In the Minimalist view, 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 (17):

[44] Let us assume, just for the sake of concreteness, 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 variant out of 1 267 650 600 228 229 401 496 703 205 376 variants.

(17) 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 zero, 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 general null subject language (18a), but like many other languages, it uses a null-subject construction in some particular type of finite clauses, namely in imperative clauses (18b). In Minimalist terms, there must be available a D-feature for T in German but it must not be used in declarative and interrogative clauses. It identifies a second-person null pronoun, in singular or plural. This analysis is questionable.[45]

(18) a. Jetzt wundere *(ich) mich. Heute regnet *(es) nicht.
now wonder (I) myself. Today rains (it) not
b. Jetzt denk (du) einmal nach! Jetzt denkt (ihr) einmal nach!
now think$_{sg.}$ (you$_{sg.}$) once about! Now think$_{pl.}$ (you$_{pl.}$) once about!

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

The 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 be 'swamped'. 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*." The problem disappeared only when the mechanism for the retention of mutations was understood.

For UG, the problem situation is the inverse. UG is a mechanism of retention, hence changes do not have a chance to spread. The original parameter values would prevail in the population since any nascent change would be swamped in the speech community. Diachronic linguistic would be a boring profession. Language changes would have a chance only in cases such as massive bilingualism in the course of massive migration.

UG theory will once be shelved on the same sub-compartment as Sir Fred Hoyle's steady-state theory for cosmology, which has misled Einstein to commit "*the greatest blunder of my life*". Neither the astronomical universe nor the ensemble of languages spoken on earth is in a steady state defined and kept within narrow limits by a universal set of structures and 'virtually necessary' constraints.

[45] If an imperative clause would contain a null subject that surfaces under focus (i), then (ii) would be acceptable:
i. Erklärt *ihr$_{pl.}$ zwei/beiden* mir das Problem! – ii. *Erklärt *pro$_{pl.}$ beiden/zwei* mir das Problem!
explain$_{Imp.}$ you two/both me the problem explain$_{Imp.}$ both/two me the problem

3.4. Closed vs. open system space

The history and future of a 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 of the set of invariant properties with the set 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 on this planet. 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, which are significantly involved in linking the 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 some of the attested means of alignment.

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. The Chinese grammar is negatively set for the bunch of parameters that define a split-ergative language, and the parameter setting in English 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.

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 exact opposite.

The evolutionary view of CEG is confirmed by what we observe. There is no pre-specified closed system of variability.[46] Grammars change in unpredictable ways and they are not always streamlined but they may develop in unexpected ways, sometimes and in rare cases. Not only nature is full of unexpected creatures, evolutionary grammar theory provides room for 'exotic' grammars, too, but they are rare, just like exotic life forms in biology.

[46] 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.

3.5 Parameter valuation during language acquisition

In Generative Grammar, language acquisition, as far as the syntactic component is concerned, means parameter valuation. The learner has to identify the value for each parameter and set it, whatever this may mean. 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. 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 goes on 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 the children will give 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 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, parameter-setting is untraceable for a child.

It is this cognitive untraceability that makes parameter identification and setting an intractable problem 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 master 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 only influence behaviour but it cannot communicate with the outside.

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 the children utter at that time, but not a proof of parameter setting. Today we know it better.[47] The acquisition of V2 in Germanic languages

[47] Akhtar (1999) present 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 instantaneous but a lengthy process. Waldmann (2012) reports a long phase of V3 for an otherwise typically developing case in Swedish. Fritzenschaft et als. (1990) report and analyze data from 5 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.

First, children primarily meet short utterances in child directed speech. Such utterances frequently contain only a single 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, it will wrongly set the parameter as if it were 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 addition to the 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 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 serves as a filler for a position that must not be empty is a poetic idea and 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 (Haider 2019) since there exists no such language at all, see Haider & Szucsich (2022a sect. 2.1) and (2022b, sect. 4). A null expletive exists only in the Generative universe (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 phonological nor semantic. It does not interact with anything. How does it betray its existence? Its existence could be ascertained only by inference. In the absence of it, another strange innate ingredient of Generative Grammar could not be upheld, namely the "*Extended Projection Principle*" (EPP), according to which a grammar must ensure that every clause have a subject. A child has to 'know' this and will kindly infer from parent's utterances such as (19) that the subject of such sentences must be a silent expletive, since their parents would presumably not have committed a grammar mistake by uttering a subjectless sentence. (19a,b) are passives of intransitives.

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

Children are able to describe their bodily feelings already at an early age and they would convey this grammatically correct by (20), which are very frequent utterances. Again, they instinctively

'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.

(20) 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, namely *es* ('it'), as in (21c). Why would they not overgeneralize it in (20) and in (21)? 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.[48]

(21) a. Vorhin hat jemand angerufen.
a-little-while-ago somebody has called
b. *Es* hat vorhin jemand angerufen.
EXPL a has -little-while-ago somebody called
c.*Vorhin hat *es* jemand angerufen.
a-little-while-ago has EXPL somebody 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 completely ignore[49] the fact that in *no Romance null subject language*, intransitive verbs can be passivized with the standard aux + participle passive (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 for intransitive passive. However, they would clearly have to permit it if the ominous expletive null subject existed. (22) and (23) present instructive minimal pairs. In fact, Romance null-subject languages ought to be a showcase of expletive null subjects. The expletive pronoun in French (22) would be subject to the null-subject parameter and end up as null expletive in (23).

(22) 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

[48] Proponents of the EPP seem to be unaware of the fact that obligatory expletive subjects are not part of the grammar of SOV languages. *Not a single* SOV language abides by the EPP and therefore, EPP-proponents apparently condone that *every* SOV or T3 language obligatorily employs a null expletive.

[49] 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).

c. qu'il a été procédé à cette arrestation — Le Figaro, Sept. 7, 2016
that it has been proceeded to this detention

(23) 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 Venetó[50].

(24) a. *Z'è* stà parlà de ti — Venetó (= Venetian)
there has been spoken about you
b. *Gh'è* stà parlà de ti — [variant of (24a)]

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

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 legs, seals have developed paddles, and primates developed thumbs. 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 evolution are the explanatory background for the cross-linguistic variants and invariants.

3.6 Mono- or poly-genetic grammars

When Ken Hale shared his field-work results with MIT colleagues, a new parameter was born, namely the (non)-configurationality parameter (Hale 1983). This has been a truly capitulating move. Warlpiri was too 'exotic' for an English-based UG, and so it was quarantined away. The new starting point for the learner in the decision tree of parametrization was [± configurational], with [- configurational] being similar to an Australian wallaby compared to an English Shetty.[51]

For human languages, Indo-European studies have confirmed the evolutionary concept of a common origin followed by dissipative evolution. The research time depth is necessarily shallow. It is about 3 millennia and this is 1% of the time that has elapsed since the attested appearance of homo sapiens in the Mediterranean region (Richter et als. 2017). So 99% of this time

[50] I am grateful to Cecilia Poletto, who is native in this language, for checking the data.
[51] Biologists will tell us that there is a very distant common ancestor for Macropodidae and Equus, but also a very long history of evolutionary differentiation.

and presumably much more before is inaccessible. But homo sapiens did not start using language simultaneously with inventing script on clay tablets in the middle 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. Since highly specific mutations have their starting point in a single genome, just like the first case of HIV. An innate UG would have come into being and resided in the head of a *single* human in a *single* clan. This person's innate UG would be the 'mother of all grammars', that is, the first ancestor of all grammar-processing brains thereafter. Unfortunately, there is no reliable empirical evidence at all for Chomsky's evolutionary conjecture, namely a 'single-mutant' theory of the human language capacity (De Boer et al. 2020).

However, even if there had been such a person, (s)he would not be able to take advantage of the newly acquired linguistic talent since the others could not follow. Moreover, the genome of the over-talented person would be 'swamped' in the course of transmission to his descendants in the speech community and disappear.

The evolutionary perspective[52] 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. Polygenesis of grammars during and after such a "*Me Tarzan – You Jane*" epoch is more probable than monogenesis of UG, given the scarcely populated African and Eurasian continents at that time. It is no outrageous assumption that modern grammars are the result of cognitive evolution starting with a one- and two-word phase.

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. Who would prefer a UG with lots of parameters if a UG without parameters would be easier learnable? UG theory predicts a highly convergent way of diachronic changes, with UG as the constantly strong attractor. The opposite is the case. Moreover, if UG theory were right, languages "*varying radically from one another*" should not exist.

3.7 (Im-)Perfection

Generative Grammar admits and in fact propagates an argument that no scientist would 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

[52] These paragraphs draw on Haider (2022b).

appearance is in theology, in the ontological proof[53] of the existence of god (see Crittenden 1968) and now it has infected 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 or 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". The answer is too obvious. And by the same token: Why does nobody ask himself why the earth is not inhabited by a singular kind of living being speaking a single language that results from the most minimal assumptions on a possible life form and language?

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, on the other hand, 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 because of the unavailability of a suitable expletive (Haider 2019). There is no *perfect* match at the semantics interface for cases in which the subject argument is to be cancelled.

3.8 Single SVO peak 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).

[53] "*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).

> "*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 2022b). 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 strict 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 (2022b). [S[VO]] languages are outcomes of grammar change and not the 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 (2022b). Obviously, 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 a 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. 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 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 als. 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. V2 became established is the documented time, in the so-called 'middle' periods, 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 be 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:

(25) a. The object [$_{VP}$ *follows* the subject]
b. Objektet *följer*$_i$ [$_{VP}$ e$_i$ subjektet] Swedish

This is just a sketch of exemplifying that the theory of UG 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. 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.

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 the infinitival construction, the infinitival marker *zu* ('to') must be in the final position of the verbal cluster (26a). Second, in IPP constructions (= infinitivus pro praticipio, aka Ersatzinfinitiv), the auxiliary haben ('have') must not follow a modal (26b), since this would trigger the participial form of the modal, which is not admitted. Therefore the auxiliary is fronted (26c).

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

The clash of the two conditions is easy to predict. If (26c) is transformed into the infinitival variant on the model of (26a), the infinitival marker would go together with haben ('*have'*), which is the finite form transformed into the infinitival form, but the auxiliary is not in the final position in the cluster anymore. What is the grammatically consistent solution? There is none. Speakers avoid this construction or they pretend to obey both rules and end up in a situation of "acceptable ungrammaticality" (Haider 2011). An ungrammatical outcome (27b) is accepted as the less deviant way out of the dilemma w.r.t. (27a).

(27) a. *ohne das beendigen *zu* haben müssen[54]
without it finish to have must$_{Inf.}$

[54] 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*.

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

In two production tests written in-class, one with 19 students of linguistics, the other with 17 students of German studies ('Germanistik'), 1 out of 19 and 3 out of 17 produced the 'correct' result, that is, the acceptable-but-ungrammatical 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. Since the swivel to the Minimalist Program, UG has been turned upside down, but apparently only by looking up from the inverted position, since 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 a long list of assumptions in the present 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 a single cross-linguistically ascertained fact and it does not discuss any bottom-up elements as have been produced in syntactic typology.

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, and most absurdly so, from field linguistics and linguistic typology. Biologists do not take seriously what nativist Minimalists speculate about. Here is an outspoken statement of a biologist's contribution to a volume on language universals:

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

> *cles of perception and production unlimited, the content of communication, and the evolutionary possibilities of the brain relevant to communication unconstrained, then the appearance of "language universals" in independent language learners would be a remarkable and illuminating finding. Some special hardware in all the language learners or users must have been installed. But if any aspect of the world is structured, if available information has predictable content or history, or IF INFORMATION-PROCESSING CAPACITIES WERE LIMITED* [emphasis mine]HH, *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, 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 for 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[55] (memory functions, pattern detection and analysis, etc.) and bodily organs (see articulation). Crucially, 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 UG theory.

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 indeed a designer, and this designer is Dawkins' "blind watchmaker", namely 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 premise, as I have tried to show in this paper. *Reduction* of the complexity of grammars is an evolving variable. It is the effect of adaption to the 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 shouldn't be judged too harshly, 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).

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

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' neither does, as assumed in the old days, language change mean continuous decay, which was a common opinion in the 19th century.[56]

The grammar of English is arguably *less complex*[57] 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 2022b). 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. 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 space complexity (i.e. memory load and retrieval costs) 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 the full description in the respective grammar books. This will yield a rough estimate of the Kolmogorov complexity in each language.

Summary

First, in the direct confrontation of the two hypotheses, the hypothesis of CEG (cognitive evolution of grammars) and UG, the CEG turns out as superior in all aspects of comparison. Second, the predictive power of CEG is the predictive power of the theory of evolution and it is positive. The predictive power of UG theory is negative.

In particular, CEG

- entails a trivial solution of the problem of *grammar acquisition* since it is the essence of evolution that a system is self-replicating. Unlearnable properties of grammars are sieved out and grammars adapt to their neurocognitive environment
- predicts the principally dissipative course of *diachronic changes* while UG wrongly predicts developments that converge to the default values of UG,
- explains the vectored quality of *grammar changes* and excludes seesaw-type changes that UG would admit,

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

[57] 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.

- explains the tension between *cross-linguistic invariants* and *cross-linguistic variation* as the result of divergence by variation and convergence by the selection environment,
- provides room for economy conditions in the shaping of grammars
- opens and ties *grammar theory* to one of the most successful scientific theories of dynamic self-reproductive systems, namely the theory of evolution.[58]

Bibliography

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

Asai Takeshi & Takao Akatsuka & Steve Haake. 1998. The physics of football. *Physics world*. 11(6): 25.

Bækken, Bjørg. 2002. Word order in different text types in Early Modern English. *Studia Neophilologica,* 74(1): 15-29.

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

Browning, Marguerite A. 1996. CP recursion and that-t effects. *Linguistic Inquiry* 27(2). 237-255.

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

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. 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

[58] 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.*

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.

Feynman, Richard. 1985. QED – The strange theory of light and matter. Princeton: Princeton University Press.

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, Jessica Alföldi, Jixin Deng, Shannon Dugan, Cornelis E van Elk, Margaret E Hunter, Vandita Joshi, Ziad Khan, Christie Kovar, Sandra L Lee, Kerstin Lindblad-Toh, Annalaura Mancia, Rasmus Nielsen, Xiang Qin, Jiaxin Qu, Brian J Raney, Nagarjun Vijay, Jochen B W Wolf, Matthew W Hahn, Donna M Muzny, Kim C Worley, M Thomas P Gilbert & Richard A Gibbs. 2015. Convergent evolution of the genomes of marine mammals. *Nature Genetics,* 47(3), 272–275.

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

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

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. 2015a. "Intelligent design" of grammars – a result of cognitive evolution. In: Aria Adli & Marco García García & Göz Kaufmann (eds.). 205-240. *Variation in language: System- and usage-based approaches*. 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.* Lingbuzz/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. 2022b. *SVO – Attractors in the declarative-to-procedural shift in grammar evolution*. lingbuzz/006747. (submitted to a volume edited by K.É. Kiss)

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.

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.

Heuser, Annika & Polina Tsvilodub. 2021. Modeling German Word Order Acquisition via Bayesian Inference *Proceedings of the Society for Computation in Linguistics* (SCiL) 2021, pages 420-424.

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*. New York: Oxford University Press.

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

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. Cambeidge: Cambridge University Press.

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

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

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

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.

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

Okasha, Samir. 2016. Population Genetics. In Edward N. Zalta (ed.). *The Stanford Encyclopedia of Philosophy* (Winter 2016 ed.). URL = <https://plato.stanford.edu/archives/win2016/entries/population-genetics/>.

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.

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

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.

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. & Pullum, Geoffrey K. 2002. Irrational Nativist Exuberance. *The Linguistic Review* 19: 59-80

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

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

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

Tracy, Rosemarie. 2002. Deutsch als Erstsprache: Was wissen wir über die wichtigsten Meilensteine des Erwerbs? *Informationsbroschüre 1/2002 der Forschungs- und Kontaktstelle Mehrsprachigkeit*. Mannheim: Universität Mannheim.

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.