“Intelligent design” of grammars – a result of cognitive evolution

Hubert Haider

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

June 2013

2015 in: Aria Adli & Marco García García & Göz Kaufmann (eds.): *Variation in Language: System- and Usage-based Approaches.* 205-240. Berlin/New York: de Gruyter.

Abstract. [1] It is astounding how closely the short history of linguistics replicates intellectual hurdles that *evolutionary biology* had to overcome in its "childhood". Until the end of the 19th century, the major divide in biological theorizing was the divide between *structuralism* and *functionalism*. The historical arguments are strikingly parallel to present day disputes between linguistic *functionalists* and generative *structuralists*. The terminology and argumentation is nearly identical. This is not surprising since the basic problem is identical, too, namely the question of how to explain the "intelligent design" of complex, well-adapted, self-replicating systems in interaction with their environment.

In biology, the dispute has never been resolved. In hindsight we understand that the intellectual combat between functionalists and structuralists could not have been won by either party because neither party was right. The dispute turned out to be completely irrelevant after Darwin's theory of evolution gained ground. This theory – adaptation as an emergent property of natural selection – explained how form and function are entangled. However, it is not "form *in order* to function" but "form that *happens to* function".

Darwin (1871:59) had already realized that evolution is not *substance-bound* and that there is a parallel between the biological evolution of species and the diachronic development of languages in terms of adaptation as a consequence of random variation and non-random but "blind" selection. It is undeniable that human language grammars are adaptive, but this is neither a product of biological evolution nor of social engineering. It is the result of evolution on the level of cognitive, self-replicating systems. Grammars are neither merely biologically grounded nor merely human artefacts. They are results of an on-going process of cognitive evolution.

Linguistics has not arrived at firm scientific grounds yet. Functionalists ("form follows function") rival with structuralists (autonomy of structures; innate programs). The functionalist schools tend to down-play the strong formal system boundaries while the structuralist schools diligently de-emphasize the role of adaptive properties (in acquisition, production and perception) in language design. Neither of the two qualities must be ignored, however, since they are indispensable for an understanding of the properties of natural language grammar systems.

The main claim of this paper is the following: Evolution by natural selection is substance-independent, as Darwin (1871:59) already saw. A process of evolution is at work not only for biological organisms but also for cognitive „organisms". In other words, the adaptive properties of grammars are a consequence of cognitive evolution in the variation plus selection game of a (substance-neutral) Darwinian evolution, with the autonomous processing capacities of the brain as the selection filter for variants of linguistic structures (and in the long run, variants of the grammar) in the variation pool of a given language.

Cognitive evolution is evolution on the level of cognitive structures (rather than on the level of biological structures, viz. the genome in biological evolution). Grammars are self-replicating systems (that replicate themselves in the course of grammar acquisition).

[1] I am very much indebted to Fritz Newmeyer and Jon Ringen for a careful and critical reading of a draft version of this paper (previous title: "*Cognitive evolution – why language systems are* …"). Their criticism, encouragement and suggestions have been very instrumental. Thanks galore to Göz Kaufmann for re-checking the final version. Any remaining shortcomings are to be blamed on the author, of course.

This process is prone to generate variants („mutations“). The variants „compete“ for restricted resources (viz. the number of brains that can be „infected“ by a given grammar variant, whose „ease“ of processing during acquisition and use is the major selection factor). Adaptation is the product of constant but „blind“ selection. Biological and cognitive evolution are identical in terms of the abstract processes (self-replication, variation plus selection), but they clearly differ in terms of their domains of application.

In sum, the descent of species and the descent of languages encompass the same abstract mechanism (self-replication, variation, selection) in two different domains. Darwin has opened our eyes for the domain of biology, but he was also aware of the fact that the mechanism of evolution is applicable to many other domains as well. One of these domains is the domain of grammars as cognitive organisms residing in our brains.

The basic issue of the dissent between linguistic functionalists and („transcendental“) structuralists today is the same as in nineteenth century biology, namely the explanation of adaptive design. The path leading to the correct answer is unnecessarily blocked by the same fallacies that can be found in the corresponding dispute in biology.

1. Introduction

We always have been functionalists, from ancient times[2] until today: Why do sea-dwelling mammals have fins for limbs? – In water, fins are more useful than legs. Why do languages employ acoustic signs? – In order to be independent of sight contact. Why are we fond of functional explanations? – Because we are social animals whose minds apparently have a predominant disposition for analyzing complex situations in terms of actors, intentions and purposes.

Functionalism is a deeply entrenched and instinctively attractive common sense perspective on complex design. The appropriate scientific perspective is intuitively less accessible, as revealed in the history of science.[3] A fairly recently lost bastion of functionalism is life science. An initially entirely functionalist standpoint was given up for a less anthropocentric but more explanatory account, namely adaptation by evolution. The functionalist predisposition is characteristic of our explanation-seeking mind but we must not project this on the object of our scientific enquiry. Our understanding is functionalist; the ontology of the objects of enquiry is not.

Linguistics is faced with the very same problem that Darwin solved. The basic question was and is this: What explains functional design in the absence of a designer? In Dawkins' (1996) words, in evolution, the watchmaker is blind, but his products are working aptly. There is „intelligent“ design, but there is no intelligence that designed it. This reads like a paradox and the anti-Darwinian camps regarded this as a fatal defect of Darwin's idea of evolution as fed by random variation. Intuitively, order out of random variation seems to contra-

[2] Why do celestial bodies move? In Aristotle's view, the source of movement of the outer sphere that triggers the movements of all inner spheres is not the causal "unmoved movent" as defined in Physics (VIII). It is "aspiration" [sic] as a final cause (causa finalis); see Metaphysics 1074b, 34.

[3] During the Stalinist period in the Soviet Union, opposing functionalism and defending the scientific position costed dissidents their life, when they refused to follow the Lamarckian-functionalist beliefs of Trofim Lysenko, then the director of the Academy of Agricultural Sciences (Graham 1987).

dict experience and the second law of thermodynamics. How could random processes produce order rather than chaos? What this intuition overlooks is this: Variation („mutation“) indeed enhances entropy, but selection is the antagonistic feature. It eliminates most of the variation. “Natural selection acts as a sieve; it does not single out the best variations, but it simply destroys the larger number of those which are, from some cause or another, unfit for their present environment.” (De Vries 1909:70).

If order and complexity emerge without an organizing ordering force, this is a result of evolution. The order parameters that happen to emerge are a non-random result of selection processes (see Heylighen (1999) on emerging complexity). What we perceive as problem solving is an emergent property and not a collectively engineered solution, as functionalism would have it.

The nineteenth century teleologists – i.e. functionalists – regarded adaptation and the fit of the organisms to their environment as the primary issue of research. Their opponents, the morphologists – i.e. structuralists – regarded commonalities of structure as the primary issue, thereby rejecting the centrality of adaptive properties. (Amundson 1998:154). “The most important and widespread biological debate around the time of Darwin was not evolution versus creation, but biological functionalism versus structuralism.” (p. 153). “The primary theoretical goal […] was the explanation of biological form. […] The continental biologists favoured structural explanations, the British favoured functional explanations. Functional facts seemed concrete and empirical to the British, and in comparison the continental structuralist theories (positing hypothetically-inferred unities) seemed transcendental”[4] (Amundson 1998:171).[5]

In our linguistic context, the functionalist conviction reads as follows. Language structures (i.e. forms determined by grammar) are the way they are because this is a function of requirements of language use. A structuralist, on the other hand, would deny that language structure could be explained solely by reference to the functions for which the structures are employed. For a structuralist, the particular properties of grammars cannot be reduced to conventionalized routines of usage.

Newmeyer (1998, 2001, 2005 and elsewhere) distinguishes two kinds of functionalist approaches, namely atomistic vs. holistic functionalism. According to the stricter conviction, namely atomistic functionalism, there is always a direct causal link between functional motivations and particular properties of grammars (2005:174). Holistic functionalism, on the other hand, is not committed to any such direct linkage but locates the subtle influence of the former on the latter in language use and acquisition. It’s shaping force becomes visible in language change. (2005:175). For Newmeyer, atomistic functionalism does not pass thorough

[4] “Major transcendentalist figures include Johann Wolfgang von Goethe, Etienne Geoffroy St.-Hilaire, Louis Agassiz, and Richard Owen. Each advocated the *primacy of structure or form* over function, of the *unity of type* over the conditions of existence.” (Amundson 1998:155; emphasis mine).

[5] “Many functionalists see rejection of generative theory as a fundamental component of functionalism.” (Dryer 2007:245). Mainstream generative grammar with its numerous hypothetical, hidden unities employed (e.g. overt & covert movement, covert lowering, remnant movement, roll-up movement, strong or weak features, overt or covert checking of features, etc.) is a good examples of (transcendental) structural explanation attempts. With or without these transcendental (i.e. empirically untestable) amendments, an account in terms of usage-independent principles that determine structure and form would not be accepted as explanation by functionalists.

empirical testing. He opts for the explanatorily weaker theoretical position, namely holistic functionalism.

Dryer (2007: 247) objects to Newmeyer's compromising withdrawal position as it "*seems to exclude an intermediate position, that in at least certain cases, a property of a particular grammar is directly motivated by some functional consideration, but that the locus of this functional explanation was at the level of historical change. Such a position seems to be a coherent one and is likely to be widely held by functionalists. […] In other words, Newmeyer's characterization obscures the distinction between two different issues; that is, it conflates the questions whether there is a direct link between functional explanation and grammatical properties and whether the locus of functional explanation is at the level of historical change or somewhere else (such as at the level of language usage).*"

What is meant by a direct link is a causal relationship between a functional aspect and a grammatical property in a functional explication of the grammatical property. Obviously, Dryer is willing to partially accept the stronger version, i.e. atomistic functionalism, and considers it scientifically and empirically appropriate.

Givón (undated: 7) is equally categorical on this issue, favoring atomistic functionalism when he refers to „roughly isomorphic matching*"*: "*The process of change itself, the invisible teleological hand that guides the ever-shifting but still roughly isomorphic matching of structures and functions, is driven by adaptive selection, i.e. by functional-adaptive pressures.*"

Givón (undated: 1) presents Aristotle as the founding father of functionalism: "*Aristotle outlined the governing principle of functionalism, the isomorphic mapping between form and function*." He explicitly refers to, and emphasizes, the teleological hand and functional-adaptive pressures. But, in reality, there is no pressure and no pressure generating device, and there is no teleological hand, as will be argued below.

The essential drawbacks of this kind of functionalism is exactly these two leading ideas, namely the „invisible hand" that designs more adaptive forms and the „pressure" on improving functionality by adopting these forms. An explanation based on the notion that future functionality drives present changes is invalid, however. This insight was established in biology more than a century ago. Functionalist „explanations" are appealing narratives that please our function-addicted mind, but these narratives are what they are, namely narratives rather than scientific explanations.

The scientifically correct core of these narratives is the purpose free process of evolution. There is no teleology, no final causes (i.e. in the sense of *causa finalis*), and no pressure; there is merely variation and selection from constantly being exposed to a given environment. Darwin's break through is a scientific theory of adaptation without any functionalist narratives, a theory of evolution with purpose-free random variation and purpose-free non-random selection as major components. What appears to be functionality driven is but the emergent effect of adaptation by selection. Final causes are not part of this system; they are merely part of the perceptual filter through which many researchers perceive their simplified world.

Of course, when biologists tell popular short-cut versions of examples of evolution, they often talk as if they are telling a functionalist narrative, but this is merely a façon de parler. They implicitly understand that a functionalist rendering is easier to grasp. The basic story, however, is a causal explanation in terms of variation and selection. In a profound study on Darwin and adaptive design, Ruse (2003) analyzes the pitfalls of our commonsensical desire for functional understanding of complex adaptive design. Our common sense is creationist; it prefers an engineering perspective on design and our favorite approach for understanding it is (invalid) functionalist reasoning.

Darwin expelled the „argument from design" from biology. His theory does not have adaptive design built in as a premise; it emerges when evolution does its work. Evolution presupposes an independently structured system that is replicating. This is the structural side.[6] Adaptation by selection covers the apparent functionalist side of evolutionary developments.

This paper contends that in linguistics, we foster the same fallacies as biologists did more than a century ago. The point is not structuralism vs. functionalisms. It is „form meets function by means of evolution". For grammar theory, this mode of explanation is novel. But of course it is not novel at all, since it is the standard mode of explanation in biology. The orthogonal viewpoints – structure-geared vs. function-driven – are wrong if maintained in isolation or as opponent positions. It is the synthesis in the concept of cognitive evolution that does justice to the correct insights of each of the competing standpoints, without their respective drawbacks.

Here is a non-linguistic example to start with: The rhinovirus that successfully recruited me as a host organism follows no teleology and is not pushed by any functional-adaptive pressure. It just happens to be a virus variant that my immune system is not prepared for. It might have successfully blocked other variants, but not this one. The success of the rhinovirus family is its variability. This feature guarantees its ability to regularly successfully infect a host and spread. If one describes this as an armament race between the „attacking" virus and the 'defending" immune system, its description will be a functionalist narrative. This narrative is easy to grasp but misleading in a crucial respect.

It suggests a causal functional relation that is not causal at all: Indeed, the virus changes its appearance quickly and regularly, but not in order to be able to outfox the immune defense and it is under no functional pressure to do so. The functionality is not programmed in, it is emergent. If it did not have the property of changing frequently it would not be a successful rhinovirus. In other words, what we see is adaptation, and what we misapprehend by overinterpreting it is a conjectured functional causality, which, according to Darwin is the exact opposite of a functionalist narrative. There is adaptivity without teleology. The virus does not change in order to spread; it spreads, because it changes.

The rhinovirus is successful and its recipe for success is rapid change. It is this property which proves successful in the selection process that we interpret as an armament race between virus propagation and virus elimination by the potentials hosts' immune systems. The

[6] Here is a simple example. Flying has been 'invented' several times in several distinct forms, e.g. bats, birds, bumblebees, etc. In each case, the predecessor structures originally had served different functions (e.g. thermoregulation).

immune system, on the other hand, is the cause for the particular property of the virus, since the virus is adapting to it due to selection. Our immune system spurs the rhinovirus into becoming the kind of virus it is now. Neither the virus nor my immune system anticipate each other.

Although this is an accurate account, it does not fit into our functionalist narrative scheme. In this scheme, there is always an agent that is coerced into doing this or that in order to gain this or that. In these narratives there is no role for an immune system that is as it is (due to an independent line of evolution) and a virus that changes frequently. Neither the immune system nor the virus would qualify as teleological agents, but their interaction can often be described as if they were. If a virus is said to be „under functional pressure" to change, our commonsensical understanding of functional connections is happy with it, but it is merely an appealing metaphor. If the virus did not change, it would not have the chance to spread, and this would be the end of the story. What we see is a virus population that spreads. Hence it happens to have this property, but not „in order to" spread. Crucially, our functionalist tunnel vision does not perceive the numerous other virus populations that lost their chance to spread because they were sieved out. The „in order to"-supposition is the superfluous and misleading ingredient we add. If the virus did not have the property of being able to outwit human immune reactions by constantly changing, it would not spread. Hence, our supposition is that it changes in order to be able to spread. This is the misattribution of a teleological component to a complex situation, simply because it is easiest for our common-sense understanding of adaptivity.

Our functional narrative is one of progress-seeking problem-solving changes. It claims that A changed into, or was replaced by B because B is a functionally superior structure and a successful step on the stair case leading to better solutions. This is a misperception, however. If we have two types, A and B, and B is selectively favored over some time period, this does not have to do with how well B is doing in absolute terms, but with how well it is doing in relation to A (Millstein 2006:648). The change may in fact be a step leading to an impasse. Both A and B may be inferior to C, and A may be closer to C than B, but C happens to not belong to the set of variants at this time in this population. The diachrony of grammar changes is full of suboptimal solutions that are local maxima, that is, the particular solution implemented for a subsystem of grammar is suboptimal for the complete ensemble of grammar. This is what evolution predicts. It is "bricolage" (i.e. tinkering) in Jacques Monod's (1971) words.

Our willingness to accept functional narratives as scientific accounts reveals our profound misunderstanding of the true causality of adaptivity, namely blind variation, as in the family of Picornaviridae (i.e. the family of viruses of which the rhinoviruses are a subbranch), and blind selection as in the case of viruses that are killed by our immune systems. The emergent property is a virus with a complex adaptive „behavior". It did not adapt by functional pressure in a teleological conspiration; it is simply the variant that is not eliminated. It has not been sieved out, in De Vries' wording above. In retrospect, it is the best-adapted variant, and only in retrospect we can tell our functional narratives: Had the virus not changed in this or that sequence, it would not have prevailed. And linguistics is full of functionalist

narratives, too. You can easily identify them by this recurrent shibbolet: If a given language did not have property A, it could not function in the way B.

In the following section, I shall argue that functionalist explanations cannot be regarded as satisfactory scientific explanations. A functionalist perspective on grammar theory may be a useful heuristics, as it is in biology, but it does not qualify as a full-fledged scientific explanation. A functionalist explanation is basically a flawed explanation since it presupposes backward causation. It must be replaced by the empirically adequate notion of adaption by evolution. This domain of evolution will be argued to be cognitive evolution.

This paper advances and elaborates on ideas that were published first in Haider (1998) and Haider (2001) and adduced for explaining adaptive design of grammars in Haider (2013). It is organized as follows: Section 2 reviews the fact that for functional explanations there is no logically valid deduction scheme. It has been acknowledged in theory of science since the 1950s that functional explanations do not qualify as full-fledged scientific explanations and are not on a par with causal explanations. Section 3 explicates the parallels between cognitive evolution and Darwinian biological evolution and illustrates cognitive evolution, with linguistic examples in section 4. Section 5 briefly examines attempts at compromising between a Darwinian and a functionalist Lamarckian approach to language and grammar change. A summary and outlook section reviews the presented assumptions and some of their implications.

2. Functionalist explanations miss the benchmark

Haider (1998) addressed the issue with the intentionally ambiguous title: „Form follows function fails“. This should be read as „form follows function“ fails as opposed to „form follows, function fails“. The combined message is this: The idea that form follows from function is going to fail since form follows, but functions may fail. Purpose or potential for future use does not explain a design, except for intentionally designed tools. For self-replicating systems with „intelligent design“, functions may be used to describe them but not to explain the causality of their design. This has become a commonplace in the theory of science since the classical work of Hempel and Nagel (Cummins 1975:742). Functional explanations are not causal explanations.

Either a property P of a system S is taken to be necessarily present to guarantee a function F, then this premise is empirically incorrect. F could be guaranteed by alternative means (cf. Nagel 1961:403). Or, the presence of P is taken to be a sufficient condition for F, then the inference from the function F to the necessary presence of P in S is not valid. All we may infer is that the presence of F contributes to a function (Hempel 1959: 310).

The grammars of natural languages are good sources for examples of alternative means of implementing identical functions. For instance, „parts of speech“ may be identified by morphological means (affixes), particles, or word order. Information structure properties, like focusing, may be coded by particles or word order or both, or merely by intonation. In each case, an identification function is implemented in different ways by different forms.

Why is a functional explanation not causal? In a causal explanation, we hypothesize that a cause C produces an effect E. Whenever C applies, we expect E (ceteris paribus). The infer-

ence from C to E is valid.[7] On the other hand, from the absence of E we infer the absence of C. Finally, if C applies and E is absent, the particular causal explanation is wrong. In sum, we regard C as an explanation for E.

In a functional explanation we hypothesize that a system contains a functionally characterized item F in order to produce a result R. Here, we regard R as the cause for F. But in this case, R cannot be a causal explanation for F (because this would amount to backward causation). The inference from R to F, given R, is not valid. And there are indeed cases in which R is given, but F is not (because R may be the effect of F′≠ F). We cannot readily falsify a functional explanation, either. Modus tollens[8] cannot be applied: If we correctly describe a given result R and hypothesize a function F as the prerequisite for R, and it turns out that our predicted F is empirically wrong, this has no consequence for R.[9]

Of course, the inference in the other direction, viz. from F to R, would be the familiar causal explanation of R as the effect of F, but its functionalist inverse is not. For functional reasoning there is no logically valid mode of inferencing. Cummins (1975:765), who tries to defend some version of functional analysis, offers this as his conclusion:

> When a capacity of a containing system is appropriately explained by analyzing it into a number of other capacities whose programmed exercise yields a manifestation of the analyzed capacity, the analyzing capacities emerge as functions. Since the appropriateness of this sort of explanatory strategy is a matter of degree, so is the appropriateness of function-ascribing statements.

Let us be more concrete and analyze a linguistic example. A functionalist explanation of „extraposition“ usually refers to the increased ease of processing for structures with material extraposed. In numerous languages, phrases may be optionally placed at the very end of the containing phrase or clause.[10] Evidently, this enhances efficient parsing of otherwise center-embedded structures by reducing the working memory load.

A functional explanation postulates that „extraposition“ is the active adaptation of the grammar by the users for the benefit of parsing; cf. Hawkins' (1994) EIC = „early immediate constituents“ measure as a functionalist trigger scenario. The functionalist explanation of extraposition is this functionality in usage. Grammars are said to provide extraposition „in order to“ avoid center embedding since this is well known to impede parsing. The language users are said to have shaped this property of their grammar for the indicated functionality.

A formal account of extraposition does not deny its functionality, but the functionality is an epiphenomenon. The primary issue is explaining how a grammar with extraposition differs from a grammar without extraposition. The explanation of the extraposition phenomenon is the grammar device that accounts for it. This is a causal explanation. Grammars that provide this device will produce extraposition structures as a result. The grammar property is the

---

[7] It is reasoning by modus ponens: [C → E] and C, hence E.

[8] [[R → F] & ¬ F], hence ¬ R.

[9] [[[F → R] & R], hence F] is an invalid inference, and so is [[[F → R] & ¬F], hence ¬R]

[10] Example: [The city [that the flood [that Hurricane Katrina had triggered] destroyed]] has been partly rebuilt.
⇒ The city that the flood destroyed *that Hurricane Katrina had triggered has* been partly rebuilt.
⇒ The city has been partly rebuilt that the flood destroyed that Hurricane Katrina had triggered.

cause for extraposition as a language phenomenon. „Functionality“ does not play a role in the explanation. The issue of grammar-parser fit is a higher-order question. It is relevant only when we compare grammars that provide extraposition with grammars that don't. The adaptivity of grammars is best accounted for in a theory that does not assume a direct causal relationship between a processing effect and a formal detail of a grammar. However, ease of processing plays an important role in language change, and language change is part of the cognitive evolution of grammars.

As described above, a functional explanation is not cogent. First, there are „strict“ OV languages that do not admit „extraposition“ and have existed for millennia (e.g. Japanese). This is a flat contradiction for a direct functional grip on grammar. An even greater embarrassment for a functional explanation is the fact that there are languages with extraposition (e.g. German, Dutch) that do not allow extraposing a class of items in spite of their functional similarity to extraposable items:

(1) a. Ist [die Erklärung, [die uns hier von ihm angeboten wurde]] wirklich richtig?
'is the explanation [which (to)us here by him offered was] indeed correct'
b. Ist [die [uns hier von ihm angebotene] Erklärung] wirklich richtig?
'is the [(to)us here by him offered$_{Agr}$] explanation indeed correct'
c. Ist [die Erklärung] wirklich richtig [die uns hier von ihm angeboten wurde]?
'is the explanation indeed correct [which (to)us here by him offered was]'
d.*Ist [die Erklärung] wirklich richtig [uns hier von ihm angebotene]?
'is the explanation indeed correct [(to)us here by him offered$_{Agr}$]'
e.*Ist [die Erklärung [uns hier von ihm angebotene]] wirklich richtig?

The relative clause in (1a) may be extraposed as in (1c), but the complex participial modifier in (1b), which is functionally (i.e. discourse functionally) equivalent to a relative clause, must not be extraposed, neither to the end of the clause (1d) nor to the end of the NP (1e). Obviously, it is the grammar that regulates extraposition and reducing center embedding does not dictate the particular grammar design. Note that the participial attribute in (1b) is on a left branch in the nominal phrase, in between the determiner and the head noun. Therefore it is a much greater obstacle for the parser than the post-nominal relative clause in (1a). However, extraposition is ruled out for (1b).

Second, in extraposing languages, even items may be extraposed that would not pose any problem for parsing.[11] On the other hand, center-embedding may be avoided[12] by „alternative means“ (cf. Nagel 1961:403), that is, paraphrases.

So, ease of parsing is neither a necessary nor a sufficient condition for a causal explanation of extraposition. It is not a necessary condition because there are strict OV languages that do not allow extraposition, and it is not a sufficient condition either since items may be extraposed that do not matter for ease of parsing. Extraposition is a system's potential of a

[11] Example: Er hat nicht *damit* gerechnet ⇒ Er hat nicht gerechnet *damit*. (‚He did not reckon *with-it*‘).
‚*Damit*‘ (‚with-it‘) is a single (compound) word. Both variants are equivalent under complexity measures.
[12] Example: The destroyed city has been partly rebuilt. It was destroyed by a flood triggered by Hurricane Katrina.

subset of human language grammars that is utilized for, but not explained by, enhancing the ease of parsing.

In general, whatever system of grammar we as humans had at our disposal, we would have to use it willy-nilly, irrespective of its user(un)friendliness, simply because there would be no alternative. But it is very plausible that nevertheless there would have soon appeared correlations between forms and contexts of usage. Crucially, these correlations are post hoc. Consequently, for natural languages, we (may) find fairly stable correlations between structure and use. But use is not the causal factor for shaping structure. Tooby and Cosmides (1990: 761-62) are quite explicit in this respect:

"[…] the only scientifically coherent account for the origin of any complexly organized functionality is (ultimately) evolution by natural selection. […] All (non-chance) functionality is ultimately attributable to the operation of adaptations, that is, naturally selected innate aspects of the cognitive architecture. Cognitive science and the adaptationist branches of biology are natural intellectual companions." (Tooby and Cosmides 1990: 761). "It is magical thinking to believe that the „need" to solve a problem automatically endows one with the equipment to solve it. For this reason, the invocation of social and practical „needs", pragmatic factors and acquired heuristics, or „functionalist" hypotheses to explain language acquisition need to be reformulated in explicitly nativist terms" (Tooby and Cosmides 1990:762).

The latter statement is correct, but too specific in one point, namely the reference to „explicitly nativist terms". This is an unnecessary restriction, and not a very plausible one. It is far from clear that there was time enough in the biological evolution of human brains to become endowed with a rich enough innate language faculty as a consequence of biological brain evolution. All our language capacities are parasitic on previously evolved capacities of the brain. The human innovation is not so much one in terms of newly acquired brain mechanism but rather in terms of evolutionarily improved brain capacities that already existed and of more efficiently crosslinking them (see Rauschecker & Scott 2009).

Tooby and Cosmides think in terms of an obvious dichotomy for biologists, namely invalid functionalism vs. valid evolutionary structuralism. In biology, a structure is usually innately determined (by the genome). In our case of grammars as cognitive information structures, the structure is a cognitive „organism" that utilizes organically determined structures (i.e. our brain functions). Therefore, the phenotype (i.e. language) is not exclusively dependent on innate qualities. Cognitive evolution is an evolution that is principally independent of its biological implementation and substrate.

Biological evolution on the genetic level is not the exclusive source of adaptive functionality.[13] In the case of cognitive capacities, natural selection gets a chance to operate on cognitive programs and their representations. The theory of evolution as developed by Darwin is principally substance-neutral, although it was developed and explicated as a theory of explaining the „origin of species by means of natural selection" (Darwin 1859). Basically, all it

[13] "As a causal theory natural selection locates the causally relevant differences that lead to differential reproduction. These differences are differences in organisms' fitness to their environment. Or, more fully, they are differences in various organismic capacities to survive and reproduce in their environment." (Stanford Encyclopaedia of Philosophy [http://plato.stanford.edu/entries/natural-selection/]).

requires is a replicating system that produces enough variants that are constantly exposed to selection.

Biological selection (which is based on the reproductive success of the phenotype) could not explain the intricate and biologically irrelevant grammar-internal details of languages. Nevertheless, the idea of a piece-by-piece biological evolution of grammar has been ventured by Pinker and Bloom (1990).[14] However, evolutionary success in biological evolution must be translated into reproductive success. It is hard to see what an accidental change in a cognitively encapsulated system of formal operations for symbol recombination could contribute to the reproductive success of those whose brain supports the change compared to those whose brain does not[15]:

For everyday life purposes of language use (including cognitive operations on propositionally structured knowledge representations) a much more primitive system of grammar seems to be flexible enough a language tool for all the purposes of the hard life of (the predecessors of) stone-age humans. The luxury of grammar systems of natural languages is by far underdetermined by the functionality of use. Replace „evolutionary" by „functionalist" in the first line of the following quote and it becomes an anti-functionalist argument, too. In fact, it is strong counter evidence that has been underestimated. The overshooting complexities of grammars are not captured by functionalist approaches and of course they are not captured by biological evolution, as Premack (1985:30) stresses:

> "Human language is an embarrassment for evolutionary theory because it is vastly more powerful than one can account for in terms of selective fitness. A semantic language with simple mapping rules, of a kind one might suppose that the chimpanzee would have, appears to confer all the advantages one normally associates with discussions of mastodon hunting or the like. For discussions of that kind, syntactic classes, structure dependent rules, recursion and the rest, are overly powerful devices, absurdly so."

Why are grammars luxurious[16] and diverse? They are luxurious because the neural substrate freely provides processing capacities for this luxury. What appears to be a superfluous complexity is but the costless exploitation of the system's potential of the human brain that happens to be available for free. The „programmer" of this potential is not an „invisible

[14] In the early days, (Friedrich) Max Müller tried to make the strongest possible point against an all-encompassing concept of evolution. He emphasized the impossibility of the biological evolution of language as a strong argument against Darwin's theory of evolution: "Language is the Rubicon which divides man from beast, and no animal will ever cross it [...] The science of language will yet enable us to withstand the extreme theories of the Darwinians, and to draw a hard and fast line between man and brute." (Müller 1862: lecture IX).

[15] See Bierwisch (2000) for a detailed discussion of the conundra and paradoxa of attempts to explain the emergence of language as a direct product of biological evolution.

[16] See Haider (2001) for an *arbitrary* and merely illustrative list of communicatively immaterial details:

i) a language with(out) fronting of finite verbs (cf. Germanic vs. Romance languages vs. strict OV languages)
ii) a language with(out) a case system consisting of subsystems that correlate with the inflection class of the verb (e.g. Georgian Nom-Acc & Nom-Ergative system vs. languages with a plain case system)
iii) a language with(out) clitic pronouns (e.g. Romance vs. Germanic languages)
iv) a language with(out) gender agreement (in the article system: English vs. Dutch vs. German)
v) a language with(out) negative concord (cf. standard German vs. Russian)
vi) a language with(out) multiple fronting of wh-expressions (e.g. Slavic vs. Romance and Germanic).

hand“ and it is not a society-based net of communicative needs.[17] It is an ongoing process of cognitive evolution. Just like biological evolution produces luxurious organisms – fantastically colored butterflies or fish populations in coral reefs, to name just two instances – cognitive evolution produces luxurious systems of grammar. Their luxury may sometimes even hamper acquisition or usage and gets cut back in the course of diachronic changes (e.g. the loss of the system of rich inflectional morphology of Latin, both in terms of the inventory and the categories, as a result of the changes that lead to present-day Romance successor languages).

Functional linguists who insist on a form-function isomorphism grounded in social interaction must feel at a loss when confronted with grammar systems that provide a much more elaborate system. “The nature of language follows from its role in social interaction” (s. Beckner et. als. 2009:3) is a bewildering dogma, given the fact that the „role in social interaction“ could be perfectly fulfilled by much simpler systems. Take for instance relatively new languages, viz. creole languages. In terms of their morphological inventories they are much simpler than many present-day Indo-European languages whose diachronic development is reconstructed for a period of roughly up to three millennia, such as Icelandic.[18] It differs greatly from present day Norwegian varieties in morphology. In terms of morpho-syntactic inventories, Norwegian is much closer to the „impoverished“ system of creole languages while Icelandic has conserved the grammatical morphology of Old Norse. It is difficult to see how one could prove that this follows from their different „roles in social interaction“ on the Norwegian or Icelandic coasts.

The merely rhetorical reference to the „role in social interaction“ is a regular ingredient of functional narratives. It is less illuminating than a claim that the nature of the shape of dogs follows from their role in social interaction. There are big Blood-hounds and small Toy Poodles and this reflects their different functions in society. This is true, but the reason is that there were breeders that intentionally selected certain properties for these breeds. For languages, there are no breeders. Any human language can serve in social interactions, and the purposes of social interaction do not cover the overall complexities in the grammar systems that determine these languages.

In sum, a functional analysis of the inventory and processes of the grammars of human languages may describe aspects of their functionality, but it does not explain it. You may correctly describe the functionality of the human eye as a component of visual perception, but you cannot explain it in terms of this functionality. An instructive example is the structural anatomy of the human eye, and in fact the vertebrate eye. It suffers from an evident „constructional“ defect. Unlike the octopus (cephalopod) eye, its wiring design is the result of “tinkering” design in Monod’s (1971) diction. The nerves approach the retinal cells from the side at which the light arrives. The smarter „engineer“ of the octopus eye correctly placed the nerves

[17] Here is an example of this claim: “Since communities are defined by shared practice, and human beings engage in a great variety of joint actions with different groups of people, the community structure of human society is extremely complex.[…] As a consequence, a language as a population is equally complex.” (Croft 2013:107)
[18] The great-grand-mother of Icelandic and the modern Norwegian varieties is Old Norse. The Icelandic settlers were Norwegians from North-West Norway.

on the dark sides of the cells. As a consequence of this design, the human eye has a blind spot (scotoma).[19]

Functional reasoning may account for the advantage of (some) vision, but it cannot explain the structures that enable vision and how they developed. Analogously, functional analysis may classify linguistic structures in terms of their contexts of use, but this does not explain how they developed and why exactly these structures are used and not others that would serve the same function even better.

3. Uniform theory of evolution – different fields of application: biological or cognitive

3.1 The background

Already in 1871, Darwin pointed out that the theory of evolution is not substance-dependent and consequently, the developments in languages appear to be parallel to biological evolution in terms of adaptation and „struggle for life“ as a consequence of variation and selection. In this publication on human physical and cultural characteristics, evolution of culture and differences between sexes, to name but a few topics, Darwin (1871: 59) made it clear that his theory of evolution is substance-neutral: “The formation of different languages and of distinct species, and the proofs that both have been developed through a gradual process, are curiously parallel. […] The survival or preservation of certain favored words in the struggle for existence is natural selection.”

Intriguingly, linguists in those days (and in fact today, if we do not count metaphorical allusions), did not take this eye-opener seriously.[20] Instead, some linguists attacked Darwin precisely on linguistic grounds, in complete misjudgment of the nature of the problem (see Müller’s fierce attack, quoted in footnote 14). In hindsight, this is understandable. In the second half of the nineteenth century, linguistics was not in command of a concept of grammar as a cognitively real knowledge system, nor of an understanding of how this knowledge system is structured, acquired and put to use. Linguistics was mainly concerned with developing a method for comparing ancient languages, particularly the (lexical as well as inflectional) morphology and the sound systems of Indo-European languages.[21]

Darwin had to convince his audience that the mainstream opinions of that time were wrong. The leading idea was the theory of Lamarck, a functionalist theory.[22] Organisms would actively adapt to their environments in order to prevail and pass on these adaptions to their successors (i.e. by heredity of acquired traits). According to Gould (2002), the first one of his ideas was „l'influence des circonstances“, an adaptive force by which the use or disuse of led organisms to become more adapted to their environment. This would make organisms adapted to their environment, taking them sideways off the path from simple to complex. The

[19] There is a blind spot at the back of each eye at the place where the optic nerve passes through the eyeball since in this region there is nor room for receptor cells. The brain computationally eliminates the blind spot and we are not aware of it.

[20] But biologists of today do. See Fitch (2007) for quantitative relationships between how frequently a word is used and how rapidly it changes over time.

[21] It was a thoroughly history-minded and text-focused science that was gradually growing out of philological disciplines. Ironically, a corner stone of Indo-European studies tuned out to be identical with one of Darwin’s sub-theories, namely the *theory of common descent*.

[22] See de Vogelaer (2007) for a confrontation in his explication of a particular grammar change.

second idea was „le pouvoir de la vie' as a complexifying force. This was thought to drive organisms from simple to complex forms. Movements of fluids would allegedly "etch out" organs from tissues and lead to more and more complex constructions regardless of the organ's use or disuse. The third and crucial ingredient was the equally wrong idea that an organism can impart on its offspring characteristics that it acquired during its lifetime.

Darwin's theory of evolution by natural selection, as summarized by Mayr (1991 ch.4) and Gould (2002), consists of five independent sub-theories. Here is a brief summary, with an appended outline of some linguistic implications for each sub-theory.

i. Evolution as such: The objects of the theory are not seen as constant or recently created nor perpetually cycling, but rather as steadily changing. Organisms transform over time.

   Linguistics: Grammars of languages are steadily changing, if not impeded by normative efforts (schooling, script culture, etc.). Changes are not cycling but follow drifts. Acquisition and language contact are the primary sources of grammar change. Another source of variation is the drive for linguistic in-group differentiation.

ii. Common descent: This theory states that each group of organisms descended from a common ancestor, and that all groups of organisms, including animals, plants, and microorganisms, ultimately can be traced back to a single origin of life on earth.

   Linguistics: Indo-European studies are a success story that illustrates this point. Languages that descended from a single proto-language have spread as far as to Iceland in the North-West and to the province of Xinjiang in China's North-West (Tocharic). Today these languages are different beyond superficial recognition, but they all are descendants of a single „mother language", with a research depth of about 3500 years.

iii. Multiplication of species: This theory explains the origin of the enormous organic diversity. It postulates that species multiply, either by splitting into daughter species or by „budding", that is, by the establishment of founder populations that evolve into new species, if geographically isolated.

   Linguistics: „Species" and „subspecies" translate as „language" and „varieties of a language". Latin, or more precisely its regional varieties, is the ancestor of a number of languages (= species).[23] What biologists call „budding", is dialect split in language change. Geographic isolation and „cross-fertilization" (i.e. language contact in bilingual brains) are catalysts for „budding".

iv. Gradualism: According to this theory, evolutionary change takes place through the gradual change of populations and not by the sudden (saltational) production of new individuals that represent a new type.

   In linguistics, again, this is commonplace. Languages change over generations. Language change is gradual, usually taking several generations and time spans of centuries. Changes typically develop out of communities with dialectal variants co-existing for a long time.

[23] As linguists we know that there are many more descendant languages of Latin than merely the 'official' Romance languages and the already extinct ones (like Dalmatian), from Sicilian, Neapolitan, Istriot, to Friulian and Piemontese, to name just a few languages on the Apennine peninsula.

v. Natural selection: According to this theory, evolutionary change comes about through the proliferation of genetic variation in every generation. The individuals who thrive thanks to a particularly well-adapted combination of inheritable characters give rise to the next generation.

This last sub-theory is the crucial point. Linguists who would subscribe to i.-iv. would not simultaneously assume natural selection to be the mechanism of language change and the emergence of new species (= languages). What would it mean that „individuals" survive and become the founding individuals of a new „species" of language? All we have to do here is to step back and rethink the analogies carefully. Of course it is not a question of survival and reproductive success on the level of the human phenotype. However, there is an exact parallel to biological evolution on a different and relevant level which has been hitherto overlooked. It is the level of the cognitive evolution of cognitive representations of replicating cognitive algorithms, namely grammars.

Evolution is adaptation by selection operating on a pool of variants. For grammars as cognitive systems, the selector is the (child's) processing brain that must acquire the grammar of a given language merely by being exposed to the language. Like in biological evolution, the winner is the grammar variant of the given language that „infests" more brains than the other competing variants. The winning variant multiplies itself more often. And just like in biological evolution, the emergent result is an accumulation of adaptive qualities.

What are adaptive qualities? The adaptive quality, as in biological evolution, is a quality that becomes effective in the environment of the system/organism. The environment of grammars is language acquisition and processing. Grammar is the „computational program" of a procedural knowledge system that the brain employs for language processing, both in reception and in production. As in biology, selection becomes a crucial issue once there is „competition" for limited resources.

A limited resource is, for instance, the amount of processing time needed for a given structure (Haider 1997). This is a limited resource for reception. The production process may take as much time as it needs, but a listener's brain must be finished with processing when the next utterance arrives. If „extraposition" saves processing time, the competing grammar that allows „extraposition" is likely to gain a selective advantage. However, just as in biology, the need by itself does not create the grammar variant it would prefer. An organism may remain in the same form in its environment for any amount of time if the environment does not change too much and no rival variant arrives that competes for the available resources. In other words, ease of processing cannot turn Japanese into an „extraposing" language as long as Japanese does not face a gradual grammar „mutation" that allows extraposition and is able to spread. Adaptation is a question of scope. The higher the rate of mutation, the higher the potential for adaptive changes (see the rhino-virus example).

A well-known mutation-prone situation for grammar changes[24] is language contact with extensive bilingualism (cf. Heath (1984); Winford (2003)). Take for instance modern Persian. Many of the grammar changes that separate modern Persian from its kin languages

[24] A parallel in biology would be the direct influencing of a genome, for instance by ionizing radiation.

like Pashto occurred after the Arab conquest with ensuing Persian-Arabic bilingualism, accompanied by the introduction of Arabic script for Persian.

Finally, the selecting filter of the processing brain is responsible for the complete absence of many simple, easy to understand potential rules of grammar across languages. For instance, there is no language that employs word-by-word mirroring, i.e. mirror-image inversion, as a grammatical rule.[25] The reason for this is clear. Our brain does not support this type of symbol manipulation. Hence, no language employs this operation for coding a grammatical rule. It is easy to program on a computer but very hard to perform with our own „wet"-ware software.

3.2. Grammars as a result of cognitive evolution

Let us now try to answer the relevant questions: Which entities are selected and how does selection work in the case of language and grammar? We have to make clear the „what-is-what" in terms of a theory of evolution, namely what the vehicle is, what the replicators are, what the interactors are and who is benefitting. The minimally necessary background for applying these notions is easily accessible thanks to Lloyd's (2012) online contribution on units and levels of selection that the following exposition takes advantage of:

Dawkins (1978) introduced „replicator" and „vehicle" to stand for different roles in the evolutionary process. „Vehicle", is defined as "any relatively discrete entity which houses replicators, and which can be regarded as a machine programmed to preserve and propagate the replicators that ride inside it" (Dawkins 1982b: 295). According to Dawkins (1982a: 62), most replicators' phenotypic effects are represented in vehicles, which are themselves the proximate targets of natural selection. The term replicator, modified by Hull (1980), is used to refer to any entity of which copies are made.

An „interactor" (Hull 1980: 318) denotes an entity which interacts, as a cohesive whole, directly with its environment in such a way that replication is differential - in other words, an entity on which selection acts directly. The process of evolution by natural selection is "a process in which the differential extinction and proliferation of interactors cause the differential perpetuation of the replicators that produced them" (Brandon 1982: 317–318).

Next, let us clarify what the corresponding referents are in the domain of cognitive evolution with respect to language. Let us start with „grammar", regarded as a cognitively represented program for processing a given language. It is the program that our language processing capacity for the given language is based on. The grammar is the „replicator". It is the entity that is replicated by language acquisition based on productions of the grammar (utterances in the given language).

The replicator is the grammar of a language understood as an information structure. The parallel to biology is very close. The genome is the information structure that governs the

[25] For example: i. Auch sie lachte (declarative) 'also she laughed' ii. Lachte sie auch? (interrogative) 'laughed she also?'

In this simple German example (which can be replicated in any of the Germanic V2-languages), ii. as a yes-no question is the word-by-word mirror image of i., but no child would wrongly jump to the conclusion that questioning means mirroring the declarative order. No processing routine of the brain would support this operation required by a rule of grammar.

make-up of an organism. Grammar is the information structure that governs the operations of the language usage system (in acquisition, production and reception). The grammar of language is the system that determines most properties of the utterances of a speaker of that language.

Next we have to identify the „interactor". The interactor is the language as a population of utterances. More precisely, it is the set of utterances the language users produce and the set of utterances that are the input for language acquisition by an individual. In other words, the grammatical properties of utterances that serve as input for language acquisition are the basis for acquiring the grammar that is responsible for the make-up of these utterances.

We now turn to the „vehicle". It is the cognitive representation of the grammar in the individual speaker's brain. It is the cognitive "software" system that enables us to produce and understand the language we have acquired. Importantly, we have to distinguish between grammar as an information structure and its cognitive representation in the brain. Grammar as an information structure is a cognitive virus, and the brain is the host. Like any virus, it needs the host for reproduction: The cognitive grammar guides the brain in language processing and acquisition. The produced language is the input for language acquisition which carries the cognitive virus into the next language-acquiring brain.

Finally, we have to identify the selection environment. Remember that natural selection acts directly on interactors and is the "process in which the differential extinction and proliferation of interactors cause the differential perpetuation of the replicators that produced them" (Hull 1980:318).

The replicators are the grammars of a language in the population of users of the particular language. It is not „the" grammar since the speech community of a given language typically is not completely homogenous. There is always variation and the variation corresponds with a set of grammar variants that differ minimally. Selection has an effect on the different perpetuation of the replicators that produced them. This is cognitive selection. Some grammar variants win, while some loose and become extinct.

What is responsible for this selection? It is the system of brain functions that are recruited for language processing. Let me clarify this with an analogy of processing functions on a computer: Grammar is the language processing software. It is implemented as a software package on a (bio-)computer with general processing properties that are largely independent of the needs of the particular software package. The software package recruits and employs the general system architecture for its functioning and it houses the grammar. The grammar is the replicator; the software package is the vehicle.

Selection becomes operative when the „software package" for the given language gradually is implemented during language acquisition. Language acquisition is not an unconstrained trial-and-error geared learning. It is guided by narrow constraints that follow from the particular combination of recruited brain functions for language processing. The language software design is constrained by the room the general system provides for such a system.[26]

[26] Note that at this level of abstraction it is not essential to decide whether there exists a domain-specific restriction on possible grammars (i.e. UG. It may exist or it may merely be the reflection of just those restrictions

It is this interaction that constitutes the selection environment. If in first language acquisition there are competing grammar variants for a given set of utterances, the winning grammar will most likely be the variant that is more easily accessible for the learning system. It will win the „competition“ for becoming implemented. Since the set of utterances that acquisition is operating on is fed by speakers whose grammars are not completely identical, there will always be variation. This pool of variants (pool of interactors) is the biotope for natural selection during language acquisition.

Finally, who is benefitting? The benefit goes to the replicator that wins. It is the grammar that becomes implemented in a brain and thereby gets the chance to replicate again. Crucially, it is not the language user who benefits. For me as a language user it does not matter whether the grammar that is implemented in my brain is computationally more or less demanding than an alternative grammar I could have acquired. It is there and I am using it. The „competition“ for becoming my grammar is not influenced by my intentions. The competition is a process I am never aware of. It is encapsulated in the operations of my learning capacity that organizes the neuro-cognitive build-up of the grammar system based on the language sample it is confronted with.

And how does the implemented grammar obtain the benefit? It has passed through the sieve of selection while other grammar variants did not. It passed because of a property that turned out to be less resource consuming or easier to acquire for the containing system than a corresponding property of the alternative grammar variant. The selecting system reacts passively; it merely is the sieve. It does not actively influence the grammar package. Crucially it does not restructure the grammar system – during or after acquisition – by improving its compatibility with the general system properties. Things that work better gain admittance, while things that do not work that easily are likely to be rejected during acquisition.

4. Functional efficiency without functionalism

The degree of functional efficiency is a function of available variation. Only if selection can get hold of a variant that enhances efficiency, it is possible for the language to change. No variation, no change. Moreover, change is local. Particular languages are functionally more efficient in some respect and less in others. These are the finger prints of evolution by mutation and selection, not a reflection of alleged needs of a speech community which keeps an eye on improving its tool of communication. Evolution is always local optimization, but what is locally advantageous may be a disadvantage on a higher level of the system since the local improvement may hamper other functions. Take for instance the restrictions on expletive subjects in English.

Modern English ‘there’ as an expletive is bound to co-occur with a postverbal nominal subject while it as expletive is an antecedent of a postverbal subject clause. As a consequence, English is left without an expletive for the subject position in clauses without a subject argument. As a direct consequence, English has become the only Germanic language that has no

that the neuro-cognitive architecture recruited for language processing imposes on the kind of ‘cognitive programs’ it is able to support.

passive for intransitive verbs (2a), since there is no expletive. VO languages (like Norwegian 2b,c), however, require a lexicalized subject position ('det'), unlike OV languages (2d).[27]

(2) a. *Often it/there was phoned
b. Ofte vart det telefonert — Norwegian
'often was it telephoned'
c. Ofte telefoneres det
'often telephons-Passive it'
d. Oft wurde (*es) telephoniert — German
'often was (it) telephoned'

The alleged discourse-communicative need of being able to leave the reference to the actor open and the actor argument unmentioned is patently ignored by the grammar of English in contrast to many other languages. English is obviously dysfunctional in this respect.

Another apparently dysfunctional trait of VO languages, compared to OV languages, is the exclusion of „why" or „how" in combination with a wh-subject. Why is German (and any other OV language) allowed to ask a question that is forbidden to the English speaker? The explanation is the causality of a structural constraint that happens not to meet „communicative needs" (see Haider 2010b, ch. 3).

(3) a.*It is unclear who left why
b.*It is unclear why who left
c. Es ist unklar wer weshalb weggelaufen ist — German
‚it is unclear who why left has'
d. Es ist unklar weshalb wer weggelaufen ist
‚it is unclear why who left has'

Even if grammars are highly efficient, they nevertheless contain quite a few dysfunctional traits. The search for an optimal grammar would be in vain, however, just like the search for the optimal animal. Efficiency is a matter of degree because the selectors in the environments correspond to independent and hence sometimes conflicting demands. What is optimal for production may be suboptimal for perception, and vice versa. What is optimal on the phonological level (e.g. cluster reduction), may be suboptimal on the morphological level (e.g. cluster reduction that produces non-distinct forms). This is a well-known and typical situation for adaptation by selection. It is localistic and may create globally dysfunctional local maxima.

A strong case for adaptation by evolution and against society-driven functionalism is the irreversibility of change. Interestingly, the irreversibility is acknowledged by functionalists and declared a consequence of functionalism (Givón undated: 8, on the „unidirectionality of change"). Functionalism does not provide a demonstrative causality, however. The needs of a society are not coherent and they may come and go, like trends in fashion. Language

[27] This requirement may be masked by the pro-drop property: VO languages with pro-drop may drop the unstressed pronominal subject, but they do not tolerate subjectless clauses, that is, clauses without a subject argument. In VO languages without pro-drop, clauses without an subject argument require an expletive subject (see Haider 2010: 35-38; 2013: 221-22). Null-subject languages are languages in which the pronominal subject argument may be phonetically null, but it is syntactically present. VO languages permit null subject clauses, but they do not permit subjectless clauses.

change, however, is generally irreversible. When case morphology is gone it is not re-introduced by the next-but-one generation. When a language becomes verb-second, like all Germanic languages (except English), the V2 grammar is not given up for the previous grammar again. It is this strong drift that is characteristic of evolution by natural selection, and not at all characteristic of fluctuating changes in a society.

Linguistic evolution is fed by two sources of grammar „mutations“, like in genetic evolution. One source is internal. It is the imperfect transmission of grammar by the language structure it determines. This is parallel to the imperfect transmission of genetic information. The other major source is external. The transmission of a grammar is disordered by external influences on the interactor (i.e. the given language). For the language learning brain, a disturbing factor is externally triggered variation. This is typically the case in multidialectal or multilingual situations. The multilingual adult brain is happy to playfully mix the languages, but the learner is confronted with enhanced variation in the given language variety. Let us briefly recapitulate:

- Evolution proceeds by the random process of variation, and an environmentally-based non-random process of selection. For grammar, the environment is the language processing brain.
- Individual intentions do not play a role. Organisms do not fabricate what they „need“ through „inner drives“ or intentional „use and disuse.“
- Mutations are not directed at the overall benefit of the interactor (see cancer-triggering mutations in biogenetics that kill the carrier of the replicator).
- Evolution is neither goal-directed nor random. It is fed by variability and driven by the non-random but non-directed process of selection.

This is true of evolution on the level of the biological genotype as well as evolution on the level of a cognitive representation (viz. the cognitive representation of grammar, with grammar as the „genotype“ of the language it determines). In each case, a reproductive system produces variation and this pool of variation is exposed to blind selection. Selection is an environmental property. In biology, it is the environment where the phenotype finds its resources, e.g. food. Analogously, in cognitive selection the environment provides the resources for the phenotype. The environment for cognitive evolution is the ensemble of brain resources for language acquisition, production and perception. The brain resources constitute the „biotope“ in which the grammar „virus“ resides after it has won the „struggle for life“ in the course of language acquisition.

Here is a concrete case for the sake of providing a more vivid impression, given the abstract points raised above. It is the split of the Germanic language family into a VO and an OV group during the time of the development of the Germanic V2 property. This is a sketch of the crucial points only. For a detailed discussion please consult Haider (2010a) and Haider (2013; chapters 1 and 5).

In the Old Germanic languages, verb positioning was variable. Its base position could be VP-final, VP-initial or VP-medial. Old English (Fischer et als. (2000:51)) is representative here (Haider 2010a; 2013, chapter 5).

(4) a. Se mæssepreost sceal [mannum [bodian þone soþan geleafan]]$_{VP}$
'the priest must [people [preach the true faith]]' (Ælet 2 (Wulfstan1) 175)
b. þæt hi [urum godum [geoffrian magon ðancwurðe onsægednysse]]
'that they our god offer may thankful offering' (ÆCHom I, 38.592.31)
c. Ac he sceal [þa sacfullan gesibbian]
'but he must the contenders reconsile' (Ælet2(Wulfstan1)188.256
d. Se wolde [gelytlian þone lyfigendan hælend]
,he wanted humiliate the living saviour' (Ælet 2 (Wulfstan1)55.98

When the V2-pattern[28] became grammaticalized, this introduced an additional, structurally distinct, verb position outside the VP, accessible only for the finite verb. As a result, the structures became highly ambiguous since the finite verb in the V2 position now had to be related to its base position. Given the alternative V positions within the VP, there always were several alternatives for the required linkage to a base position:

(5) V2 + variable V-positioning inside the VP (pre-change situation)

a. XP V_{fin} YP ZP	3	alternative base positions (see 6)
b. XP Aux_{fin} V YP	3 + 2	alternative base positions
c. XP Aux_{fin} YP V	3	alternative base positions
d. XP Aux_{fin} YP V ZP	3 + 2	alternative base positions

The three alternative base positions for (5a) are indicated in (6). What this amounts to is a high degree of indeterminacy for identifying the filler-gap relation of the fronted finite verb.

(6) XP V_{fin} YP ZP
a. XP V_{fin} [YP ZP **-i**]
b. XP V_{fin} [YP **-i** ZP]
c. XP V_{fin} [**-i** YP ZP]

The present-day situation is as follows. Every Germanic language with a single exception[29] has changed into a language type with a fixed head position for the verb in the VP. The Northern group is head-initial (VO); the continental Western group is head-final (OV). This is a unique split within a language family, and it is parallel with the grammaticalization of V2. In fact, the latter change invited the former.

(7) a. V2 + fixed V-position in the VP

b. XP V_{fin} YP ZP	1	(VO) or (OV)
c. XP Aux_{fin} V YP	1(+1)	(in VO)
d. XP Aux_{fin} YP V	1	(OV)

The advantage of the change is obvious. It replaces a grammar with a high degree of indeterminacy with a grammar with an easy to determine filler-gap relation. The simpler grammar variant wins, and since there are two possible implementations (namely OV and VO, masked by V2), it comes as no surprise that both found their way into the brains of language learners and users. The simpler grammar wins because it suffices for processing the given language

[28] [XP [$Vfin_i$ [… **-i** ….]$_{VP}$]]
[29] Yiddish has conserved a grammar that all other Germanic languages changed into one with rigid head-positioning. Arguably, this is due to the ad-strate effect of being embedded in Slavic speaking communities. Slavic languages are languages with flexible head-positioning. The all show the variation illustrated in (3).

structure and there is no necessity (in terms of a large residue of patterns that cannot be analyzed with this grammar) for a more complex analysis that would be imposed by the previous grammar. This is selection on the level of cognitive representation of alternative grammar variants.

5. No need for metaphors

5.1 The path of evolution is not paved with intentions

Haspelmath (1999) suggests a compromise: a functionalist characterization of grammar change as a consequence of evolution. He describes linguistic „evolution“ (without any reference to selection) as a diachronic process driven by cumulative intentional actions. He is more precise than others who describe it vaguely as „cultural evolution“ with “replication, and variation” occurring “when we use language in the service of joint actions between human beings in a community” (Beckner et als. 2009:9). “As in biology, observed adaptive patterns in language can be explained through diachronic evolutionary processes as the unintended cumulative outcome of numerous individual *intentional* actions” (Haspelmath 1999:180, emphasis mine).

The reference to intentional actions makes this position completely incompatible with Darwinian evolution. Of course, one may use the term evolution in a non-technical, metaphorical sense, but the alleged parallel to biology becomes a mere equivocation then. What is taken for granted in this particular case is an invisible synchronization of the „individual intentional actions“. The typical behavior of individual entities in a closed system is governed by the statistical law of growing entropy (i.e. the second law of thermodynamics), however. Not synchronization but dissipative development is the normal process when a system is continuously changing locally.

What is crucially missing in this conception is a precise notion of selection. Evolution is adaptation by natural selection. There is no adaptive change without selection. Adaptability is relative to the selection mechanisms and these are diverse and unstable if conceived of as „intentional actions“. Without a precise characterization of selection, evolution is as insignificant a notion as percentage figures without a baseline.

Evolution of grammar cannot be based on intentions. Intentions can be served by many different means, and intentions are not constant. They themselves may change or be replaced continuously, and they cannot be assumed to be uniform over a large group of individuals. The selector for the selection process behind linguistic evolution is not a homogenous grammar council of users who formulate their annual motions. It is as blind as the selectors in natural selection in biology.

The only constant selector is the uniformity of our language processing brains. Every child processes language by means of the same brain structures and resources. This is the uniform selection environment. When this selector is exposed to language variation, grammar variation is the pool for selection. To put it in a simple way, the grammar variants are „competing“ for a host brain as a vehicle for grammar replication. The brain provides the vehicle and the grammar determines the language structures this human being will use. This language

is the interactor that reflects the grammar that will be picked up by the next language-acquiring brain.

The selector is blind. Any feature of a grammar that makes grammar acquisition, reception and production easier than a competing grammar will win because brains will acquire this grammar more easily than the less efficient competing grammars, and in the end the winner takes them all. There is no intention at issue, nor could it be. Ease of processing is of course not the only selection filter. Storage and retrieval is selective, too, and many other factors at the neuro-cognitive interface. The structure of grammar is always cognitively encapsulated. As a speaker I have no idea how the grammar in my head is structured, and I have no idea how I could change its structure in order to change its usability properties.

5.2 The target of evolution is not the utterance, but the grammar

A position similar to Haspelmath's (1999) but more elaborated in terms of neo-Darwinistic concepts has been worked out by Croft (2000, 2013). According to Croft (2013:98): "The evolutionary framework gives us a theory of structured entities that vary through replication, and a systematic relationship between the products of language use (the replicators) and the knowledge and behavior of language speakers (the interactors)."

Note that utterance types (viz. the products of language use) are seen as the replicators and the speakers as the interactors. This is a fundamental misunderstanding. Utterances do not replicate; they are produced by the vehicle and are the interactors. The speakers are the containers of the vehicle. The misidentification of utterances as replicator is bound to lead to wrong conclusions. The cause for this misunderstanding seems to be the functionalist conviction that variation has to be understood as a function of language use in a speech community (Croft 2013:107): "The population of linguemes (the replicators) – the variants that are propagated or go extinct – is defined by the population of speakers (the interactors), who replicate the linguemes in language use. Thus we must examine the speech community and its role in the linguistic selection process."

This will not work. Selection does not operate on utterance variants; it operates on grammar variants that determine the structure of utterances. For grammar change, the target is certainly not the utterance. The target is the grammar. An utterance is merely a „molecule" of the phenotype of a grammar, that is, the set of utterances determined by a grammar. In cognitive evolution, the language is the phenotype and the genotype is the grammar. The replication of an utterance is clearly not the replication of grammar.

Biological replication means the replication of the genome by expression in the phenotype. A single utterance, either as a token or as the set of tokens (= type) crucially is not the representative of the phenotype of grammar. Genetic variation is accessible for selection only by its effects on the phenotype. Analogously, variation in an utterance types is the reflex of variation of the grammar that determines the type. Selection operates not on the utterance, but on the grammar that is accessible via a set of utterances. The grammar is replicated by the language-acquiring and -using brain.

Croft does not explain how a frequent novel „lingueme" could have a causal effect on the grammar. The „theory of utterance selection" (Croft 2000) merely tells us that a given

utterance is used frequently if it turns out to be useful for any context of usage and that this frequency of usage propagates this lingueme. Crucially, it does not tell us how the frequency of a specific utterance can be the cause of grammar change.

Here is a counter-example: Several years ago, the ungrammatical utterance (8a) became popular because a prominent person uttered it sincerely. In fact, a web search confirms that the ungrammatical utterance (8a) is still very popular. It is seven times more frequent than its grammatical variant (8b),[30] but nobody would assume that the frequency of this specific utterance could make it a catalyst of grammar change that would replace dative by nominative in the German passive from now on.

(8) a.* Hier werden sie geholfen
'here are you$_{Nom}$ helped'
b. Hier wird ihnen geholfen
'here is you$_{Dat}$ helped'

Croft's "theory of utterance selection" may account for the on-going process of lexical changes (since this is by and large a Lamarckian kind of change), but not for changes in the procedural system of language (viz. grammar change). Changes in the lexicon are ubiquitous and continuous. This is not language change, however. Language change is not so much a change in the inventory; it is a change in the computational system. The token frequency of an utterance can only explain the fossilization of an utterance[31] as an idiom or the adoption of a novel lexeme. It is neither the type frequency nor the token frequency of an utterance that matters. What matters is the availability of an alternative structuring of an utterance. This is not a question of token frequency, but rather a question of the size of a type set.

Dryer (2007: 245) emphasizes the importance of frequency: "One of the central ideas of functionalist linguistics, especially over the past fifteen or so years, is that frequency of usage plays a central role in explaining why languages are the way they are." Given the discussion above, it should be clear by now that we are facing a chicken-and-egg problem. Adaptive success means higher utility, and higher utility is likely to have an effect on the frequency of use. It is not frequency that drives functionality. Causality goes in the other direction.

Newmeyer (1998:124) puts it this way: "All linguists would agree that text frequency is a response to a variety of factors, from cognitive complexity to pragmatic usefulness. The question is to what extent frequency itself can legitimately be called upon as an „explanation" for whatever phenomena seem to be sensitive to it."

„Ease" or „naturalness" in processing by the language processing brain is the explanandum and frequency is the effect.[32] In language change, higher frequency may be a cor-

[30] 2.9 million pages for (7a) versus 394.000 pages for (7b) (google-search on April 17th, 2012).
[31] ‚*Vater unser*' (the Lords Prayer, lit. 'father our') is ungrammatical in German, but it is the first verse of the prayer every Christian knows. It is the direct translation of '*pater noster*'. It's extremely high token frequency in German has not had any effect on the grammar of German, though. Obviously, token frequency had not had any effect on the Grammar of German, of corse. Token frequency merely has the effect of fossilizing forms, but not of generalizing and establishing the type.
[32] As in biology, frequency is a multi-faceted phenomenon. Typological frequency is a sign of successful adaptation by selection and shows on the populations' scale, i.e. cross-linguistically. Lightfoot's (2002) "threshold frequency" is a within-language (i.e. within population) frequency criterion. It is the critical frequency at which a variant gets a chance to spread.

relate and it may be part of an explanation, but it is not the cause. The cause is a grammar change that makes a variant available that prevails over „competing“ variants. Prevailing may be reflected in frequency.

Frequency is a topic in first language acquisition, too. Here, an obvious problem with frequency as a causal factor becomes perspicuous. The usual argumentation uses frequency considerations as evidence and as a basis for a suspected functional connection and overlooks that this explanatory move would work only post hoc. If you know which functionality to look at, you can count frequencies. But crucially, observing frequencies would not tell the child what to do with frequency gradients. This is clearly stated by Yang (2004:452): “Although infants seem to keep track of statistical information, any conclusion drawn from such findings must presuppose that children know what kind of statistical information to keep track of. After all, an infinite range of statistical correlations exists” and “statistics requires UG”. Statistical learning “is constrained by what appears to be innate and domain-specific principles of linguistic structures, which ensure that learning operates on specific aspects of the input” (Yang 2004:455).

6. Summary and outlook

Language is both a cultural and a neuro-cognitive phenomenon embedded in a biological substrate. Which kind of language change is a cultural[33] or a „natural“ phenomenon is therefore an empirical question. This paper argues that the „cognitive evolution“ of grammars is an instance of classical evolution that is operative in a different domain. Darwin showed how his theory of evolution operates on the level of biological organisms but was aware that his theory is not substance-bound. Cognitive evolution operates on the level of the structures of cognitive representations of self-replicating cognitive systems. Biology and neuro-cognition are different domains, but the mechanisms of evolution are the same.

As Darwin (1871) realized, the development of languages and the development of species follow the same general laws of evolution but they are implemented in different domains. The laws of grammar change are the laws of evolution that hold for replicating systems, whose variants (mutations) are exposed to selection. The effect is adaptation (to the selection parameters) and diversification, with luxurious side effects not excluded.

Grammar change is a domain of cognitive evolution. The replicating system is the grammar of a language. It is an ontologically real object that resides in human brains as a cognitive virus and makes the brain a servant for the purpose of its replication (by producing language structures that serve as the basis for the acquisition of this grammar by other brains).

Cognitive evolution is nothing more than evolution working on cognitive entities. Hence, the outcome is totally parallel to biological evolution in terms of yielding systemic adaptation. This paper argues for abandoning the customary metaphorical allusions to evolution and for taking the ground-breaking Darwinian insight seriously, on the appropriate level of theoretical generalization. Crucially, this does not mean that the evolution of grammars is to be subsumed under the domain of biological evolution. This would obviously be misguided.

[33] Adaptation as a consequence of cognitive evolution by natural selection is one source of grammar changes. There are of course other sources as well whose best explanation, for instance, may be a socio-linguistic one; see Trudgill (2011).

It means that a hitherto overlooked domain of application of the theory of evolution is the domain of self-replicating cognitive systems. The prominent case are grammars as the neuro-cognitive programs for the ensemble of processing systems for language production and reception, and, most importantly, for language acquisition. Once cognitive evolution is recognized, the adaptive properties of grammars find a scientific explanation and the frequently felt desire for invoking (functionalistic) teleological explanations can be satisfied in an unexpected but logically valid way. It is the very same access road that has been opened and paved by Darwin in the domain of organismic biology.

The evolution of biological species is an exemplary case of evolution, but it is not its exclusive domain. Evolution applies to any replicative system that replicates in a domain of restricted resources. Once grammars are seen as natural, replicative systems that come in variants whose replication depends on limited resources, Darwinian evolution is predicted to apply, with adaptation and diversification as the inevitable outcome. Eventually, cognitive evolution is the solution for „Newmeyer's dilemma": On the one hand, as he has shown in great detail, it is hopeless to advocate „atomistic functionalism". But on the other hand, „holistic functionalism" is hard to prove since this view maintains that the link between grammatical constructs and functional motivations is "extremely indirect" (Newmeyer 2005:225). As he explains "There is no direct linkage between external and grammatical properties. The influence of the former on the latter is played out in language use and acquisition and (therefore) language change and is manifested only typologically." (Newmeyer 2005:175). This is exactly what we expect for system adaptivity under natural selection. There is no direct influence of function on form but nevertheless the system of forms will end up showing an adaptive design.

Cognitive evolution provides a causal relation and predicts this overall picture. The result of cognitive evolution is a family of adaptive changes in grammar systems that may be described, but not explained, from a vantage point of holistic functionalism. Cognitive evolution of grammars explains adaptation without any functionalist backward causation and without any direct linkage between properties of particular grammars and conjectured functional motivations for each of these properties.

It is needless to emphasize that the precise understanding of a theory of cognitive evolution of grammars as cognitive systems replicated by brains is at least as far away as it was for Darwin. He developed his theory without any precise understanding of the real source of variation (i.e. genetic mutations), without any idea of the real target of selection (viz. the genome), and had to link natural selection crucially with the idea of heredity long before the basic concepts of genetics and inheritance were discovered. It took several generations of researchers until the theory had arrived at its standard form as the modern evolutionary synthesis. But he dealt with palpable entities, namely animals and plants.

We linguists deal with grammars. Like in nineteenth century biology, the situation in linguistics will not make fundamental progress before neuro-cognitive research offers us a more precise understanding of the selection environments for various kinds of structurally different cognitive representations. For biological evolution, the selection environment was much easier to estimate. Darwin understood it to be everything that enhances or impedes reproductive success. The environment for cognitive selection is not so easily accessible, alt-

hough it is very close to every theoretician since it is located somewhere between his right and his left ear.

The other disadvantage in comparison to Darwin is equally basic, namely the precise identification of the species and their formal distinctions. In biology, it sufficed to observe, analyze and describe; in grammar theory, observation does not help. Presently, all we have is less than a dozen sufficiently analyzed grammar systems. Typological cross-linguistic descriptions of human language grammars are comparable to poor-quality photographs that barely suffice for telling apart a gnu and a zebra.

For a sufficient understanding of cognitive selection we need a much better understanding of the neuro-cognition of language processing and language learning as an arguably domain-specific capacity and its specifics in the ensemble of domain-general processing functions. At the moment, we hardly know anything. As everyone knows, it is difficult to catch a black cat in a dark room, but it helps at least to know that it is there.

For some time this set of circumstances will leave room enough for linguists who are happy with functional narratives: "Unfortunately, the vast majority of self-designated functionalists, of whatever sect, tend to expostulate about cognition without studying the cognitive literature" (Givón undated: 10) - and without studying the well-documented history on the rise and fall of functionalism in biology, one might feel tempted to add.

References

Amundson, Ron (1998): Typology reconsidered: two doctrines on the history of evolutionary biology. *Biology and Philosophy* 13: 153–177.

Beckner, Clay, R. Blythe, J. Bybee, M.H. Christiansen, W. Croft, N.C. Ellis, J. Holland, J. Ke, D. Larsen-Freeman, T. Schoeneman (2009): Language is a complex adaptive system. *Language Learning* 59: 1–26.

Bierwisch, Manfred (2000): Evolutionstheorie und Linguistik – Drei Probleme. In: Klaus Richter (ed.) *Evolutionstheorie und Geisteswissenschaften.* Univ. Erfurt: Acta Academia Scientiarum 5: 129-189.

Croft, William (2013): Language use and the evolution of languages. In: Kenny Smith and Philippe Binder (ed.) *The Language phenomenon.* Berlin: Springer. p. 93-120.

Croft, William. (2000): *Explaining language change: an evolutionary approach*. London: Longman.

Cummins, Robert (1975): Functional analysis. *The Journal of Philosophy* 72:741-765.

Darwin, Charles (1859): *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. London: John Murray, Albemarle Street.

Darwin, Charles (1871): *The descent of man, and selection in relation with sex.* London: John Murray, Albemarle Street.

Dawkins, Richard (1978): Replicator selection and the extended phenotype. *Zeitschrift für Tierpsychologie*, 47: 61–76.

Dawkins, Richard (1982a): Replicators and vehicles. In: King's College Sociobiology Group, Cambridge (eds.) Current Problems in Sociobiology, Cambridge: Cambridge University Press. p. 45–64.

Dawkins, Richard (1982b): *The Extended Phenotype*. Oxford: Oxford University Press.

Dawkins, Richard (1996) [1986]: *The Blind Watchmaker*. New York: W.W. Norton & Company, Inc.

De Vogelaer, Gunther (2007): Innovative 2pl.-pronouns in English and Dutch - 'Darwinian' or 'Lamarckian' change? In: *Studies van de BKL 2007 / Travaux du CBL 2007 / Papers of the LSB (Linguistic Circle of Belgium)* 2:1-14.

De Vries, Hugo (1909): Variation. In: Albert C. Seward (ed.) Darwinism and Modern Science. Cambridge: Cambridge University Press. p.66-84.

Dryer, Matthew S. (2007): Review of Newmeyer's *Possible and probable languages*. *Journal of Linguistics* 43:244-252.

Fischer, Olga & Ans van Kemenade & Willem Koopman & Wim van der Wurff (2000): *The syntax of Early English*. Cambridge: Cambridge Univ. Press.

Fitch, W. Tecumseh (2007): An invisible hand. *Nature* 449: 665-666.

Givón, Talmy (undated): *On the intellectual roots of functionalism in linguistics*. Ms. Univ. Oregon. URL http://linguistics.uoregon.edu/wp-content/uploads/2013/03/FuncLing11Rus.pdf (18.4.2013)

Gould, Stephen Jay (2002): *The structure of evolutionary theory*. Cambridge MA: Harvard Univ. Press.

Graham, Loren (1987): *Science, philosophy, and human behavior in the Soviet Union.* New York: Columbia University Press

Haider, Hubert (1997): Economy in Syntax is Projective Economy. In: Chris Wilder/Hans-Martin Gaertner/ Manfred Bierwisch eds.: *The Role of Economy Principles in Linguistic Theory.* Studia Grammatica 40: 205-226. Berlin: Akademie Verlag

Haider, Hubert (1998): Form follows function fails - as a direct explanation for properties of grammars. In: Paul Weingartner & Gerhard Schurz & Georg Dorn eds. *The Role of Pragmatics in Contemporary Philosophy.* Vienna: Hölder-Pichler-Tempsky. p. 92-108.

Haider, Hubert (2001): Not every why has a wherefore – Notes on the relation between form and function. In: Walter Bisang (ed.) *Aspects of Typology and Universals*. Berlin: Akademie-Verlag. p. 37-52.

Haider, Hubert (2010): *The Syntax of German.* Cambridge: Cambridge University Press.

Haider, Hubert (2010a): Wie wurde Deutsch OV? Zur diachronen Dynamik eines Strukturparameters der germanischen Sprachen. Ziegler, Arne (ed). *Historische Textgrammatik und Historische Syntax des Deutschen - Traditionen, Innovationen, Perspektiven*. Berlin: De Gruyter (p. 11-32).

Haider, Hubert (2013): *Symmetry breaking in Syntax.* Cambridge: Cambridge University Press.

Haspelmath, Martin (1999): Optimality and diachronic adaptation. *Zeitschrift für Sprachwissenschaft* 18:180-205.

Hawkins, John A. (1994): *A Performance Theory of Order and Constituency*, Cambridge: Cambridge University Press.

Hawkins, John A. (2004): *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

Heath, Jeffrey G. (1984): Language Contact and Language Change. *Annual Review of Anthropology* 13:367-384. Palo Alto: Annual Reviews, Inc.

Hempel, Carl (1959): The logic of functional analysis. In: Llewellyn Gross (ed.) *Symposium on Sociological Theory*. New York: Harper & Row.

Heylighen Francis, (1999): The growth of structural and functional complexity during evolution. In: F. Heylighen, J. Bollen & A. Riegler (ed.). *The Evolution of Complexity.* Dordrecht: Kluwer Academic. p. 17-44.

Hull, David L. (2001): *Science and selection*. *Essays on biological evolution and the theory of science*. Cambridge: Cambridge University Press.

Lightfoot, David W. (2002): *The development of language: Acquisition, change and evolution*. Oxford: Blackwell.

Lloyd, Elisabeth (2012): Units and levels of selection. Edward N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*.

URL (18.4.2013) http://plato.stanford.edu/archives/win2012/entries/selection-units/

Mayr, Ernst (1991): *One long argument*. Cambridge, Mass.: Harvard University Press.

Millstein, Roberta L. (2006): Discussion of "four case studies on chance in evolution": Philosophical themes and questions. *Philosophy of Science* 73 (5):678-687.

Monod, Jacques (1971):. *Chance and necessity: An essay on the natural philosophy of modern biology*. New York: Alfred A. Knopf.

Müller, Friedrich Max (1862): *Lectures on the Science of Language*. London: Longman, Green, Longman, and Roberts.

Nagel, Ernest (1961): *The structure of science: Problems in the logic of scientific explanation*. New York and Burlingame: Harcourt, Brace and World Inc.

Newmeyer, Frederick J. (1998): *Language form and language function*. Cambridge, Mass.: MIT Press.

Newmeyer, Frederick J. (2001): 'Where is functional explanation?' In: Mary Andronis, Christopher Ball, Heidi Elston, and Sylvain Neuvel (eds.) *Papers from the thirty-seventh meeting of the Chicago Linguistic Society. Part 2: The Panels*, Chicago: Chicago Linguistic Society, pp. 99-122.

Newmeyer, Frederick J. (2005): *Possible and probable languages*. Oxford: Oxford University Press.

Pinker, Steven & Paul Bloom (1990): Natural language and natural Selection. *Behavioral & Brain Sciences* 13: 707-784.

Premack, David (1985): *"Gavagai!" or the future history of the animal language controversy.* Cambridge, Mass.: MIT Press

Rauschecker, Josef P. & Sophie K. Scott. (2009): Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature Neuroscience* 12: 718-724.

Ruse, Michael (2003): *Darwin and design: Does evolution have a purpose*? Cambridge, Mass.: Harvard University Press.

Tooby, John & Leda Cosmides (1990): Toward an adaptionist psycholinguistics. *Behavioral and Brain Sciences* 13 (4): 760-763.

Trudgill, Peter. (2011): *Sociolinguistic Typology: Social determinants of linguistic complexity.* Oxford: Oxford University Press.

Winford, Donald (2003): Contact-induced changes – classification and processes. In: *OSU WPL* 57: 129–50.

Yang, Charles D. (2004): Universal Grammar, statistics, or both? *Trends in Cognitive Sciences* 8: 451–456.