

GRAMMAR CHANGE – A CASE OF DARWINIAN COGNITIVE EVOLUTION

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

Abstract

The grammar of a language is a complex neuro-cognitively represented program. Such programs are results of an ongoing evolution, that is, results of the interplay between variation and selection, on the level of cognitive processes. In cognitive evolution, the same principles apply as in biological evolution, but the domains of biology and cognition are different, of course. Evolution of grammar is not a facet of biological evolution. It is a domain of evolution on its own. Since the same general abstract principles are at work, the structure of evolutionary processes is parallel in each domain. As a consequence, many theoretical insights of a century of research in population genetics can be cautiously adduced and applied to the explanation of grammar change. This paper claims that grammar change is a case of Darwinian evolution that targets the cognitively encapsulated, procedural parts of grammar. The cognitively accessible, declarative content of grammars is open for social changes.

1. Evolution – From metaphor to reality

Nature.com introduces the topic as follows: "*Evolution of language is the gradual change in human language over time [...] and can be considered analogous to biological evolution, although it does not necessarily occur through the same mechanisms.*" The analogy to biological evolution is deemed to be loose. Language change is filed under "*cultural evolution*", which "*is the change over time of non-biological aspects of human society. The process is loosely analogous to biological evolution [...] and includes changes in language, art and social behaviour and norms.*" This paper claims that language as a neuro-cognitive phenomenon is subject to Darwinian evolution.

Languages are defined by their grammars. Grammar, as understood here, is an essential part of a cognitive "program package" underlying and constituting the activities in language use and acquisition. Changes in grammar are changes in a cognitively encapsulated neuro-cognitively represented program, that is, changes in a cognitive 'app' for language processing, in perception, production and crucially also in acquisition. It will be argued that *Darwinian evolution* of the neuro-cognitively encapsulated aspects of grammars is essentially involved in the origin and the changes of these programs. Darwinian evolution is the result of constant selection operating on a pool of random variations. In the case of cognition, the target of selection is a neuro-cognitively represented, self-replicating program, namely the cognitive 'grammar app'.

Not all changes in languages are Darwinian. Consciously accessible aspects of languages, as for instance the continuous changes of preferences in the domain of lexical items and their morphological shapes, are open for conscious interventions since they are consciously accessible. Their characteristics of change are different, and an adequate theory of language change has to differentiate them. To put it briefly, changes in the program are Darwinian, changes of the accessible lexical content are not. In a declarative/procedural model of grammar, whose neuro-cognitive basis has been described in Ullman (2001), Darwinian evolution shapes the procedural component. The declarative components are open for socially motivated changes.

The following three subsections review current usages of the concept "evolution" in language change, ranging from loose analogies to approaches that elevate the principles of Darwinian biological evolution to principles of a metatheory of evolutionary processes in fields other than biology. In biology, Darwin's theory has prevailed over Lamarck's theory, but in other fields, notably in socio-cultural domains, Lamarckian concepts have been revived, with mixed reactions, ranging from acceptance to disapproval.

1.1 Evolution as metaphor

"Loosely analogous" approaches employ biological terminology but interpret it differently. Haspelmath (1999), for instance, regards diachronic changes as evolutionary processes since changes in biology and in linguistics are often adaptive, which, in biology, is explained as a result of Darwinian evolution: "*I argue in this paper that linguistic adaptation is in many ways analogous to biological adaptation.*" "*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 1999:186, 180).

The characterization as "*cumulative outcome of numerous individual intentional actions*" makes clear that "evolution" in this view is not Darwinian. In Darwinian evolution, adaptive patterns result from *constant* but *blind* selection operating on a pool of *random* variants. There is no teleology or intentionality involved; see Hanke (2004) for details. Darwinian evolutionary explanation does not appeal to teleology because the theory does not involve any goal-orientation. "*Darwinian evolution is completely myopic*" (Simon 1996: 47).¹ Adaptedness is a result (a posteriori) rather than an (a priori) target of goal-seeking activities (Mayr 1992).

The same is true for grammars. Individual users are unable to intentionally change their acquired grammar since they have no access to its parts and pieces. These are cognitively encapsulated and the users are not aware of the specific grammatical conditions that govern their own verbal behavior. They are aware of the output of the system but not of the intricate details of the system that accounts for the output. They may (intentionally) choose between alternatively available output variants, but this, by itself, is not a change of grammar. It is a source of variation, though, which is a prerequisite for Darwinian selection.

What seems to be taken for granted under Haspelmath's perspective is the synchronization and qualitative transformation of the "individual intentional actions". "Functional selection"² as a consequence of usefulness or functional need³ is claimed to be responsible for the proliferation of specific variants. "*High-frequency structures may become obligatory, and low-frequency*

¹ "*Evolution depends on two processes: a generator and a test. The generator produces variety, new forms that have not existed previously, whereas the test culls out the generated forms so that only those that are well fitted to the environment will survive.*" Simon (1996: 45). "*At each incremental step the evolving organism becomes fitter relative to its current environment, but there is no reason for the progress to lead to a global maximum of fitness of the individuals, separately or severally.*" Simon (1996: 47).

² *Functional selection* (Nettle 1999: 30-35) rests on the notion that some variables may confer fitness because they make language easier to learn or use. Note that in Nettle's view, the gain in fitness goes to the language user, and crucially not to the grammar. In this paper, selection in terms of 'fitness' is selection between grammars as cognitively-based systems that compete for replication, that is, for learning brains. The bonus goes to the grammar.

³ Comrie & Kuteva (2005) object to functional-need explanations on detailed empirical grounds.

items may be lost as a result of their (high or low) frequencies." "*Entrenchment due to frequency thus corresponds to selection in biology*" (Haspelmath 1999: 190).

In Darwinian theory, frequency is not a cause but a potential effect. Higher frequency is an effect of selection if a variant expands in a population because phenotypes with this particular genetic variant happen to contribute to a higher reproduction success. If "frequency selection" were part of Darwinian evolution, novel variants would never have a chance of spreading since they would disappear in the population already at the beginning as a result of their necessarily low initial frequency. Frequency selection is a principle of a debunked theory, namely Lamarckian evolution (see sect. 1.2).

A crucial aspect of Darwin's theory tends to be neglected in these analogies, namely the mechanism of *retention* of traits, independent of their frequency in the population. This is an essential ingredient of the theory of Darwinian evolution. This issue has already been raised forcefully in the early days of evolution theory by Jenkin (1867), known as the "swamping argument". Jenkin objected to Darwin's theory by pointing out that an accidentally appearing profitable variety could not be preserved by natural selection. It would be 'swamped' by the ordinary traits in the course of backcrossing in the population. Darwin (1872: 71-72) accepted the objection as the most valuable one he had ever read (Morris 1994: 313). It took half a century for the theory to remove this difficulty. The weak point was the by-now replaced theory of inheritance as a *blend* of traits of the parents. Genetics provided the necessary insights for replacing *blending inheritance* by *Mendelian inheritance*. What is inherited is not a phenotypic trait but an information package in terms of genetic information passed on to progeny. When expressed, it accounts for the inherited phenotypic characteristics. A variant (i.e. dominant allele) is coded in the cells of its carrier and therefore this allele is immune against being swamped. The offspring will inherit it and will be able to pass it on, even in a community in which no other individual than mother or father originally carried this variant. What is described by "*entrenchment due to frequency*" is not an ingredient of Darwinian evolution but rather a characteristics of the rapid diffusion of innovations and artefacts in society.⁴

Under primarily usage-based perspectives, language design tends to be (mis-)perceived as a human artefact. The grammar of a language is regarded much like a kind of open-source software that is open for collaborative public efforts of improving on it. However, this analogy fails. The source code of our language processing cognitive software is cognitively encapsulated and inaccessible. It is cognitively as inaccessible as the genetic code. Even if I desired to change it, I couldn't do this by intensive wishful thinking. Analogously, we cannot intentionally change the grammar as a mental software. What can be intentionally changed is the behaviour, that is, the choice between items a grammar makes available when we use language. If grammars were open to constant intentional interventions aiming at enhancing their compatibility with all kinds of user demands, grammar change should closely resemble changes in traditions, life-styles, and fashion, which it does not.

⁴ It accounts, for instance, for the spreading of smart phones and frozen pizza worldwide, but not for the dynamics of self-reproducing systems with variants in a population. This is social behaviour rather than natural selection.

What is missing in such a picture is an essential ingredient of the neo-Darwinian theory of evolution, namely the mechanism for the *retention* of a novel and therefore necessarily infrequent variant in the population. This mechanism is essential since it guarantees that there is a pool of variants exposed to constant selection. If the population could not retain infrequent traits, favourable traits within this pool could not gradually spread as a consequence of sieving out less favourable variants, independently of their frequency, by natural selection. Increase in frequency is not the cause of selection but its gradual effect.

Darwinian evolution is a mechanism with two essential components in a feed-back relation, namely the genotype and the phenotype. Natural selection has no direct access to the genotype. Unlike genetic engineering, it does not manipulate a particular trait of the genotype. Selection acts on phenotypic variation in a *population*. The phenotype is the vehicle of the genotype. The spreading of a particular genotype in a population is the effect of environmental conditions that amount to a bonus program in the reproduction of the carriers of a particular genotype. In the case of language, the genotype is the given grammar as a cognitively represented system and the phenotype is speech, that is the output of and input to the system in action. The selective environment for natural selection in the case of grammar is *not* the social environment (see below). It is the neuro-cognitive environment in which the grammar system is cognitively embedded. Finally, the crucial reproduction rate is of course not the sexual reproduction rate of the speakers who have acquired a particular grammar variant, but the *cognitive reproduction rate* of a grammar variant in terms of the number of brains which acquire and use a given variant rather than a competing one.

1.2 Lamarckian evolution and goal-seeking behaviour

Entrenchment due to *frequency* of usage is one of two central theses of Lamarck's theory of evolution (Lamarck 1809). This thesis is known as "*the law of use and disuse*", whereby individual life forms gradually change their traits in proportion to the extent that they are used or not. The second thesis is the "*law of the inheritance of acquired characteristics*". Lamarckism assumes the inheritance of acquired characters as an inherent property of all living organisms. In other words, experiences during lifetime are transformed into heritable features of an individual. Changes are triggered and directed by environmental needs and traits can be acquired and passed on through the efforts of individual organisms, with the effect that organisms gain or lose functions that they need or have no use for, respectively (Johansson 2005: 16).

In today's evolutionary biology, epigenetic effects sometimes are equated with acquired traits that become inheritable, much like what Lamarckian theory assumes.⁵ Haig (2007: 415) objects and argues that epigenetic inheritance expands the range of options available to genes but evolutionary adaptation remains the product of natural selection of random variation. Loison (2018) emphasizes that molecular mechanisms of transgenerational epigenetic inheritance themselves are evolved products of natural selection. "*This means that the kind of inheritance of acquired characters they might be responsible for is an obligatory emergent feature of evolution, whereas traditional Lamarckism conceived the inheritance of acquired characters as a property inherent in living matter itself.*"

⁵ Szyf (2014) reports an inheritable effect of fear-conditioning of mice up to two generations under the title "Lamarck revisited" in Nature Neuroscience.

A third ingredient of Lamarckian evolution is a complexifying, goal-seeking force which is the source of the emergence of new organs in the course of time. If new requirements make themselves felt, they lead to the development of new structures and their upkeep and maintenance.

In today's biology, Lamarck's concepts are of historic interest only, but they are adduced for interpreting cultural changes, not always in a homologous way, though. For qualifying as Lamarckian evolution, properties acquired at the phenotypic level must be encoded in the genotype that is passed on to the next generation. It is not sufficient to appeal to "use and disuse" and a goal-oriented behavior. The appropriate term for this would simply be "social learning".

"Social learning would be literally Lamarckian if the knowledge that an organism acquired about its environment somehow came to be encoded in its genetic material and thereafter was inherited by its progeny. As far as I know, none of the advocates of an evolutionary analysis of conceptual change view social learning in such a literal fashion. The whole point of social learning is that information is transmitted independently of genes. [...] Inheritance is not involved (just transmission)". (Hull 1988: 37).

The crucial point of Lamarckism is that it is a doctrine that admits the possibility of (genotypic) *inheritance* of *acquired* (phenotypic) characters by individual organisms in evolutionary processes. Consequently, a social replicator (as the analogue of genotype) has to be identified, with a view on exploring the distinction between genotype and phenotype at the social level. If in the course of cultural changes, practices are invented, used, accumulated and *directly* passed on to others by social means of communicating such knowledge, this is not Lamarckian evolution.

In linguistic terms, the concept of acquired 'inheritable' traits could be interpreted as follows. If the 'genotype' of a given language (variant) corresponds to 'grammar of a language (variant)', a phenotypic character is the reflex of a genotypic character. Hence, an *acquired* genotypic character in the Lamarckian sense is a character which is triggered by a change in the environment and subsequently coded in the genotype during lifetime. This means that a language user would have to be able to change an already acquired grammar into a new grammar anytime in life and then pass it on. So the crucial question is this: Is a speaker able to (intentionally) change the grammar of the language (s)he has acquired in first-language acquisition in a significant way?

It is the main claim of this paper that the cognitively encapsulated parts of grammar are inaccessible for changes after first-language acquisition. They can only be patched up locally but not changed. Local changes are possible only in the accessible domains of the *inventory* (i.e. lexicon and morphology), but not in the *procedural* domain (see section 3.1 for details).

Haspelmath (1999: 192) admits that linguistic evolution might be "entirely Lamarckian" and concludes that "*structural adaptation in language must be due [to] the effect of constraints on performance combined with a mechanism that turns preferred options of language use into structural patterns of grammar.*"

If "*preferred options of language use*" were the trigger for structural patterns of grammar, then the past millennia of usage would have been sufficient for streamlining any language. Consequently, languages whose users share preferred options of use due to their shared culture and life form will have become structurally highly similar. Conversely, if grammars widely differ in their structural patterns, such as for instance Chinese, English, Japanese, and Russian, this

ought to reflect differences in "the preferred options of language use". This is neither empirically evident nor theoretically plausible.

Let me pick just a single structural difference for illustration: Why would English (and many other languages, such as every Romance null-subject language), but not Japanese (and many other languages, such as German), rule out passivization of intransitive verbs? Isn't there any demand for it in Britain or Spain, but a strong need in Japan and Germany? Why could there be such sharp differences in the "preferred options of language use" persisting over centuries? The availability of such constructions is not a question of language *use* but of the particular system of *grammar* that is compatible (or at odds) with the results of passivizing intransitive verbs.⁶ For analogous considerations with respect to relative clause structures and tense systems see Comrie & Kuteva (2005).

1.3 Generalized Darwinism

Generalized Darwinism or universal Darwinism is the collective term for a variety of approaches that extend the Darwinian theory beyond its original domain, with linguists as early adopters (Schleicher 1873). In social sciences, evolutionary economics has become a productive field⁷ (see Hodgson and Knudsen 2006, Nelson 2007, Hodgson 2011). The unifying bond for Generalized Darwinism is the perspective of a generalized version of the mechanisms of variation, selection and retention so that it lends itself to a common general framework for studying evolutionary processes not only in the biological domain but also in other fields, including even the cultural domain. Hodgson and Knudsen (2008: 51) summarize the essential point as follows: "*It is proposed here that Darwinism provides a general, meta-theoretical framework for dealing with complex evolving systems, consisting of populations of varied and replicating entities, which are found in both nature and human society. There is no alternative to the core Darwinian principles of variation, selection and inheritance to explain the evolution of such systems. Darwinism includes a broad theoretical framework for the analysis of the evolution of all open, complex systems, including socio-economic systems.*"

The shared theoretical background contains a replicating system with sources of variation plus a way of retaining variants, and an environment that exerts a selection effect. The ultimate aim is a general theory of evolution with domain-specific implementations.

2. Mechanisms of evolution in biology

In the popular understanding, evolutionary change is equated with the result of natural selection and the "*survival of the fittest*". However, selection is not the only and in some instances not even the most frequent mechanism of generating changes. Genetic *drift* and *gene flow* are essential factors as well. The homologues of drift and gene flow are non-negligible in language change as well.

2.1. Darwinian evolution by natural selection

⁶ If the standard passive is applied to an intransitive verb in an [S[VO]] language, the subject position must be filled with an expletive subject. If the language does not provide one, passive cannot be applied to intransitives. In no Romance null-subject standard language, intransitive verbs can be passivized.

⁷ In 2006, the *Journal of Evolutionary Economics* devoted an issue, namely 16(5), to evolutionary concepts in economics and biology; see the editor's statement (Witt 2006); for a follow-up literature review see Breslin (2010).

Darwin's theory of evolution by natural selection, as summarized by Mayr (1991) and Gould (2002), consists of five independent sub-theories. Here is a brief summary, with asides on the linguistic parallels of each sub-theory (Haider 2015).

i. *Evolution as such*: Organisms transform over time. They are neither constant nor perpetually fluctuating. Grammars of languages are cognitive systems that are known to transform over time if not impeded by normative efforts (schooling, norms of script culture, etc.). Changes are typically not fluctuating but in many cases directional and irreversible for a given language.

ii. *Common descent*: Each group of organisms descended from a common ancestor, and all groups of organisms ultimately can be traced back to a single origin of life on earth. As for languages, Indo-European studies are a success story of this point. Languages that descended from a single proto-language have spread as far as to Iceland in the West and to the Chinese province in the East (Tocharic) within several millennia.

iii. *Multiplication of species*: This theory postulates that species multiply, either by splitting into daughter species or by budding, that is, the establishment of founder populations that evolve into species, if geographically isolated. „Species“ and „subspecies“ translate into „language“ and „varieties of a language“, such as dialects. An example of budding under geographic isolation is the split of language due to migration (cf. Afrikaans, in comparison with Dutch).

iv. *Gradualism*: Evolutionary change proceeds gradually and not by sudden (saltational) appearances of new complex systems. Sudden changes are not excluded, though.⁸ In linguistics, this is commonplace. Grammar change is gradual, usually spanning many generations. Changes typically develop out of areas with dialectal variants co-existing for a long time.

v. *Natural selection*: Evolutionary change comes about through the proliferation of (genetic) variation in every generation. If individuals have more offspring thanks to a well-adapted combination of *inheritable* characters, these characters spread in the population. Selection sieves out less well-adapted characters.

This sub-theory is the crucial one. Linguists who would subscribe to i.-iv. would not simultaneously endorse natural selection as the mechanism of language change and the emergence of new species (= separate languages). What would it mean that "individuals" with a specific grammar variant thrive and that this variant thereby spreads?

Sceptics focus on an inappropriate concept of "*individuals who thrive*". This concept must not be interpreted as "individuals who use a given language", even if some linguists have interpreted it that way, for instance Pinker & Bloom (1990), Nettle (1999), Solan et al. (2005). Nobody has been able to produce reliable evidence for changes in grammar with a significant effect on the reproductive success of the *users*, conferred by the preference of a particular language variant.⁹ The relevant "individuals" in the context of grammar evolution are the individually embodied *cognitive representations* of a grammar variant in the brains of language users.

⁸ There are phases of less gradual changes (cf. Eldredge & Gould 1972), and "*organisms with a profoundly mutant phenotype that have the potential to establish a new evolutionary lineage have been termed 'hopeful monsters'.*" (Theißen 2009: 43).

⁹ The crucial issue is not having a language vs. no language, but the selection between grammar variants. Furthermore, sharing an individual language may have social benefits, but this reduces variation. It does not explain the actual choices.

These are the "individuals who thrive" and spread in the population of language using brains, and this is the relevant effect of cognitive evolution.

Darwinian evolution is change by selection operating on a pool of variants. For grammars as cognitive systems, the primary selector is a child's processing brain when it acquires the grammar of a given language by being exposed to utterances shaped by this grammar. Like in biological evolution, the winner is the grammar variant of the given language that 'invades' more brains than competing variants. This way, the winning variant multiplies itself more often. And just like in biological evolution, the emergent result is often an accumulation of adaptive qualities relative to the selecting environment, that is, the neuro-cognitive environment.

What are adaptive qualities in such an environment? The environment of grammar is language acquisition and language processing. As in biology, selection becomes a crucial issue once there is competition for limited resources. Limited resources are storage, time, and effort. A limited resource, for instance, is the amount of processing expenditure needed for a given structure. There is a *bottleneck effect* in reception. "*Sequential information, at many levels of analysis, must rapidly be recoded to avoid being interfered with or overwritten by the deluge of subsequent material. To cope with the Now-or-Never bottleneck, the language system chunks new material as rapidly as possible at a range of increasingly abstract levels of representation.*" Christiansen & Chater (2016, sect. 7). If a competing variant can be processed more easily, the responsible grammar variant is likely to gain an advantage by being selected in the course of language acquisition.

However, just as in biology, the need by itself does not create the grammar variant it would prefer. An organism may retain the same form in its environment for any amount of time if the environment does not change radically and no rival variant emerges that competes for the available resources (cf. the species of coelacanths). There is no adaptation without variation. If there is variation, the higher the rate of variation, the higher the potential for adaptive changes, as Fisher's theorem tells (Fisher 1930: 35).¹⁰

2.2. Flow and drift

Gene flow, drift, and natural selection are distinct mechanisms, but do not exclude each other. They may operate separately or simultaneously and influence each other. Drift may prepare the ground for selection, for instance, and gene flow may reduce variation and thereby hamper selection.

Gene flow (also known as *gene migration* or *allele flow*), is the exchange of genetic material and consequently of genetic variation between different populations.¹¹ The linguistic counterpart is the dynamics of grammar changes in multilingual populations, between dialects of a language or between different languages in language contact areas. Extensive bilingualism, see Heath (1984), Thomason & Kaufman (1988), Winford (2003), is a well-recognized variation-prone situation. Take for instance modern Persian. Numerous grammar changes which separate

¹⁰ "The rate of increase in fitness of any organism at any time is equal to its genetic variance in fitness at that time."

¹¹ "Gene flow tends to oppose the effects of local selection. However, it can also replenish the local population and local genetic variation, which are both pre-requisites for evolution by natural selection." (Lenormand 2002: 189).

modern Persian from Eastern Iranian kin-languages like Pashto took place after the Arab conquest with ensuing Persian-Arabic bilingualism, accompanied by the introduction of Arabic script. *"It was not only lexical elements that entered Persian: Arabic morphological and even syntactic features also found their way into the language."* (Gazsi 2011: 1015).

Genetic drift – also known as *allelic drift* or *Sewall Wright Effect* – is a mechanism of evolution by which allele frequencies in a population change by chance (e.g. random births, deaths, and Mendelian segregations in reproduction). Newberry et al. (2017) characterize the linguistic analogue as follows: *"Unlike selective forces, which bias a language learner towards adopting forms that are intrinsically easier to learn or more effective for communication, drift arises purely by chance: the learner chooses randomly among the sample of forms that she happens to encounter."*

Genetic drift is the result of random sampling among variants. The variants in one generation do not reproduce in equal ratios in the next generation. Drift is an evolutionary process, but it is evolution due to chance, not selection. Drift can lead to significant changes over a short period of time in small populations. The linguistic *parallel* to genetic drift is the random nascent of grammar variants which is enhanced by segregation of subpopulations of the speech community, for instance by migration.

A word of warning is appropriate here: The concept "drift" in population genetics must not be equivocated with Sapir's (1921) notion of "drift". Sapir's notion is the exact opposite of the genetic notion since Sapir's characterization of drift is actually a paraphrase of the effect of natural selection: *"The drift of a language is constituted by the unconscious selection on the part of its speakers of those individual variations that are cumulative in some special direction."* (Sapir 1921: 166).

What Sapir describes is the interplay between random *variation* and constant and therefore *directional* selection and *retention*. This is the essence of Darwinian natural selection operating on grammatical systems: *"It by no means follows that the general drift of language can be understood from an exhaustive descriptive study of these variations alone. They themselves are random phenomena. The linguistic drift has direction. In other words, only those individual variations embody it or carry it which move in a certain direction."* (Sapir 1921: 165)

2.3. Viral replication

The replication of the cognitive representation of a grammar resembles the replication of a virus. It is totally dependent on a host. Only a virus that successfully recruits a host is able to induce it to produce copies.¹² So, the host is at the same time the vehicle and the selecting environment. This is where natural selection operates (cf. Rubio et al. 2013). *"Viruses [...] experience strong and diverse selective forces, sometimes acting on timescales that can be directly measured."* Spielman et. al. (2019: 427). Variants that succeed in capturing a host will survive and increase their frequency (= positive or adaptive selection), whereas variants that do

¹² For a briefing see Goulding (2020): Adsorption, entry into the host cell, transcription, and subsequent replication by synthesis are the essential steps. <https://www.immunology.org/public-information/bitesized-immunology/pathogens-and-disease/virus-replication>. [March 15, 2020]

not succeed will decrease in frequency (= negative or purifying selection).¹³ Because of negative selection, natural selection fuels a continuous arms race between the host and the virus. The host wins whenever its immune system prevails and the virus wins whenever it is able to outwit the defence system. Like cell-based life, a virus is subject to evolution by natural selection. With respect to reproduction, however, there is an essential difference between viruses and more complex life forms. Our layman's perspective on viruses is entirely negative because of their pathogenic effects.¹⁴ In this pathogenic perspective, Sobin (1997) has coined the term "grammatical virus" as a device that reads grammatical codes and affects them negatively.¹⁵ Sobin's virus metaphor for external and 'pathogenic' influences on grammar has been carried on by Lasnik & Sobin (2000) to the *who/whom* alternation, by McKenzie (2013) to participle-object agreement in French, or by Sundquist (2012) to Norwegian negation. In all these instances, the virus metaphor is used for characterizing a grammatical phenomenon as alien to the grammatical system and 'pathological'.

However, viruses are not generally pathogenic. Roossinck (2011: 99 and 106): "*Although viruses are most often studied as pathogens, many are beneficial to their hosts, providing essential functions in some cases and conditionally beneficial functions in others.*" Roossinck and Bazàn (2017) call the symbiosis with mutualistic viruses an "intimate partnership". Unlike pathogenic viruses, a mutualistic virus is selected positively.

This paper favours the concept of a grammar as a viral system in a *symbiotic*¹⁶ relation with domain-general cognitive capacities of our brain. A grammar needs the cognitive capacities of our brain as a host and replicator, but it is far from pathogenic. It is at the same time the domain-specific cognitive device for language processing. Such a viral program package is necessarily subject to the selective effects of the host system in the process of replication. Selection has the effect that grammars as cognitive systems adapt to the cognitive capacities they depend on: "*Overall, language appears to have adapted to the human brain more so than the reverse*" (Schoenemann 2012: 443).

3. Elements of a Darwinian cognition-based evolution grammar

3. 1 Grammars as cognitive systems are susceptible to variation and selection.

The expression "Darwinian cognitive evolution" in the title should be read as follows. Darwinian evolution, that is natural selection among variants of a self-reproducing system, applies to variants of a cognitive program, that is, the cognitively represented grammar, with the domain-general computational capacities of the brain as the constant selector environment of the 'viral' grammar. As a viral system, a grammar needs and employs the general cognitive capacities of our brain for reproduction.

¹³ Because more changes are harmful than are beneficial, negative selection plays an important role in maintaining the long-term stability of structures by removing deleterious variants (Loewe 2008:59).

¹⁴ Rhino, influenza, Ebola or Corona viruses are notorious. These lines are written during a virus pandemic curfew.

¹⁵ See Schütze (1999) for a rejoinder. Sobin proposes an analysis of prestige constructions of English, such as plural agreement in expletive constructions, as a result of a "grammatical virus infection" as a process that operates "out of conformity with the principles that govern the proper devices of a grammar" (Sobin 1997: 319).

¹⁶ Symbiosis =_{Def.} two entities living in an intimate association (Roossinck 2011: 99).

Just as in biological evolution, reproductive success is a function of the dissemination of particular variants. In biological evolution, biological success is measured by the number of descendants. In cognitive evolution, the descendants are not the result of sexual reproduction but the result of a viral reproduction process. If a grammar is a (mutualistic) cognitive virus, its reproductive success is measured like the success of any virus namely by the number of hosts it 'infects'. The more brains a grammar variant enters and occupies when these brains acquire a language, the more replicas of this grammar variant will be passed on to learners of the following generation, when they acquire their grammar based on the utterances of the previous generation governed by the particular grammar variant, and so on. An immediate result of this selection-based processes is the adaptation of human grammars to the properties of the processing brains as the selecting environment. This is the familiar effect of adaptation by natural selection in Darwinian evolution.

An essential feature is still missing in this sketch of the evolution of a cognitive grammar program subject to natural selection. It is the mechanism of the *retention* of a variant in the population. In the biological case, a variant is retained because it is coded in the genome of its carrier as an allele.¹⁷

In genetics, an allele is one of the possible forms of a gene. In grammar, an allele is one of the possible 'pieces' of grammatical information that is 'expressed' (in the technical sense of gene expression) in the formation of a particular linguistic expression. That grammatical features can be successfully treated as alleles has been shown in studies that apply algorithms developed in populations genetics to studies of linguistic diversity, such as Reesink et als. (2009) or Greenhill et. al. (2017). For the sake of concreteness, let me adduce a well-studied diachronic phenomenon, namely the fronting of finite verbs in English. In Shakespearean time, two patterns still co-occur as interrogative variants in English, namely (1a,b) and (1c,d):

- | | |
|--|------------------------------------|
| (1) a. <i>Knows</i> he not thy voice? | (All's Well that Ends Well: IV, i) |
| b. <i>Lies</i> he not bed-rid? | (The Winter's Tale: IV, iv) |
| c. <i>Did</i> he not <i>send</i> you twain? | (Love's Labour's Lost: V, ii) |
| d. <i>Did</i> he not <i>moralize</i> this spectacle? | (As You Like It: II, i) |

The clause-initial position in questions could either be occupied by a fronted finite verb (1a,b) or by a dummy auxiliary, namely "*do*" (1c,d). In present-day English, main verbs are not fronted anymore. This rules out (1a,b). According to Ellegård (1953:162), the change covered a time span of roughly four centuries (see Figure 1). A detailed discussion follows in subsection 3.4.

The *do*-variant gained momentum in early Modern English (ca.1500). Then, the frequency of *do* in questions and in negative declaratives rises continuously until *do* is obligatory in these contexts in the early 19th century. In the first period, the grammar variant with *obligatory do*-support is a minority variant, but it is compatible with the majority since it generates a subset of structures that are also generated by the grammar of the majority, in which *do*-support is optional. In the final period of the change, speakers with optional *do*-support are the shrinking minority.

¹⁷ "An allele is a variant form of a gene. Some genes have a variety of different forms, which are located at the same position, or genetic locus, on a chromosome." <https://www.nature.com/scitable/>. Different alleles may result in different observable traits. In the context of grammar, roughly speaking, an allele is a grammatical variant.

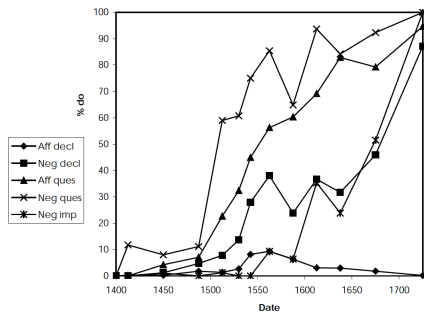


Figure 1: Percentage of dummy *do* by sentence type (Ellegård 1953:162)

If such a development is a result of natural selection operating on grammars as cognitive entities, a grammar variant with *do*-support must be shown to offer some advantage under selection. This selection advantage should become effective in first language acquisition as well as in processing. What is the advantage of a stationary main verb in comparison with variants with alternating verb positions? If the head of the VP is stationary and non-portable, it is easy to identify and to predict. No filler-gap relation needs to be computed. Instead, a dummy auxiliary is placed in positions that formerly have been accessible to any finite verb, that is, main verb or auxiliary. In evolutionary terms, this means that one of two 'grammar alleles', namely a grammar with a specific restriction on the class of verbs that are subject to the rules of fronting has prevailed.

Having illustrated how potential homologues of genetic alleles may look like in the context of grammar in principle, let us return to the crucial issue for Darwinian evolution by selection, namely the *retention* of selected variants. In the above exemplary case, the steady development, starting from nearly zero at about 1400 and reaching the 100% level half a millennium later, is the reflex of a *constant* expansion of the grammar variant with *do*-support. In other words, the grammar variant with *do*-support is retained in the population for generations and gradually spreads, although it once started as a minority variant. What is the mechanism that guarantees retention? It must be a fairly rigid property that cannot be reset or undone easily.

In biology, the mechanism is genetic inheritance. The genome we are born with is the genome we die with. In the case of the cognitive representation of grammars, the rigid property is not genetic inheritance. It is a combination of the inertia of *procedurally* coded systems and their conservation due to the loss of plasticity during brain maturation. The rigid property of grammar is a reflex of its procedural quality: "*Procedural memory is formed more slowly than declarative memory. The other side of the coin is that procedural memory is more robust so that, once formed, it is better preserved, and it is also inflexible, and therefore difficult to change.*" (Lee 2004: 69).

Our language capacities depend on two different neuro-cognitive systems, the procedural and the declarative system (Ullman 2001). The declarative system is the general purpose memory system. Its content is consciously accessible and easily rewriteable. Procedurally organized systems, on the other hand, are encapsulated and therefore consciously inaccessible. They are difficult to rewrite and therefore change resistant to a large extent.

As for grammar acquisition, psycholinguists agree that there is a window of opportunity (Paradis 2009, Morgan-Short & Ullman 2011). It closes in the course of neuro-physiological processes of synaptic pruning and myelinisation on the one hand, and the stabilization of the processing networks by inhibitory regulations on the other hand (Voss et. al 2017). In sum, this

accounts for the significant loss of plasticity of the brain for grammar acquisition. The available evidence points to the conclusion that the *procedural* components of the syntactic knowledge systems are efficiently compiled within a sensitive period that terminates around the age of eight years. It is this ensemble of procedures that is change-resistant to a sufficiently high degree for the rest of the linguistic life of an individual.

Here are two examples for the discrimination of *procedurally* vs. *declaratively* stored information, namely verb morphology and case assignment. Irregular verb forms are stored *declaratively*. *Declaratively* stored information is information about *particular* items. It can be changed easily and a changes may rapidly spread in the population (see the discussion in sect. 3.4). Case *assignment*, on the other hand, is a *procedurally* represented information. A case relation is not an individual property of an item. It is a complex relation between noun phrases, case governing grammatical items and structural relations between them. The two variants of the causative middle construction in (2) are fully synonymous and exchangeable in communication. Nevertheless, the case that has to be assigned to the object of the infinitival verb is different in configuration in (2). Native speakers are not aware of the principles that govern the distribution of accusative and nominative in these and other structures. Their competence is a reflex of procedurally stored information.

- (2) a. Hier ließ sich der/*den Sommer genießen.
 here let REFL the_{Nom}/the_{Acc} the summer enjoy
 'Here it was possible to enjoy the summer'
- b. Hier ließ *es* sich den/*der Sommer genießen.
 here let *it* REFL the_{Acc}/the_{Nom} summer enjoy
 'Here it was possible to enjoy the summer'

A change of a procedurally coded grammatical property is a change in a system that is not open for conscious interventions. If such a system changes, it changes like other complex systems change, namely in the ways channelled by Darwinian evolution. There is variation, there is selection, and if other variants are sieved out by selection, there will be a variant that continuously spreads. Eventually, it may turn out that it is the only remaining variant. In this case, observers will testify that a grammatical change has been completed. The shared properties of such diachronically attested phenomena are directionality, persistence and a sufficient time span. Sapir (1921:180) already identified such "drifts", namely "*the drift toward the invariable word*" and "*the drift toward position as an all-important grammatical method*".

3.2 Darwinian cognitive evolution operates on general *structures*, not on individual *content*.

Since Darwinian (cognitive) evolution operates whenever the following preconditions are met – and cognitively represented grammars meet these conditions – grammars are subject to selection and therefore to Darwinian evolution by selection:

- A self-reproducing (cognitive) system
- generating a pool of (cognitive) variants
- in a selective (cognitive) environment for (cognitive) reproduction
- and a mechanism of (cognitive) retention of selected variants.

There is no denying that grammars are complex cognitively represented systems. The capacity of language processing is to a large extent domain specific. This shows in double dissociations.

Grammar may be spared when general cognition is significantly impaired. On the other hand, grammatical capacities may be impaired while general cognitive capacities are unimpaired.

The cognitively represented program for processing a specific language is self-reproducing, in cycles of output-input-output. The output controlled by a grammar serves as input for the acquisition of this grammar by others, who, when progressing, produce output which is input for others, and so on. The transmission process generates variation, and the pool of variations is fed by external factors, too, for instance by intensive language contacts (cf. "gene flow").

Grammar is first of all a skill-type capacity, that is, it is part of the procedural networks. These systems are inaccessible for conscious interventions and subject to processes of brain maturation that stabilize it. Its effective functioning also involves a network of inhibitory effects. These properties divorce it from declaratively stored grammatical information, such as the lexical and morpho-syntactic inventory of a language. This can be volatile. During the whole life time, it is easy to amend and it is easy to add or replace lexical items by others within a short period by social learning.

The acquisition and cognitive representation of grammar is embedded in, and sustained by, domain-general cognitive capacities. In IT terminology, the domain-general capacities are the operating system. A grammar is an app that runs on this cognitive computation platform. The general capacities determine whether a grammatical variant is 'easier' to acquire or carry out than other variants. In sum, this amounts to a constant selection environment for grammatical variants.

These factors taken together are sufficient for triggering natural selection. However, the selection processes operating on a pool of variants of a cognitively represented and self-reproductive system can result in a grammatical change only if selection finds enough time for becoming effective. So, the variants must be *preserved* in the pool of variants. Natural selection cannot cope with moving targets. Retainable variants are procedurally coded elements of grammar that are entrenched during the sensitive period(s) of first language acquisition. This guarantees retention.

Second, evolution is a slow process and adaptive mutations are selected over periods of many generations before they reach a critical level of sustainability in the population. Bell & Gonzalez (2011) studied biological evolution by selection in real time¹⁸ under lab-conditions and found out that adaptive changes may happen surprisingly fast even for biological system, namely within 50 generations. In biology, "one generation" is one reproduction cycle. For language change, one reproduction cycle is one acquisition-to-production cycle, that is the period from the start of first language acquisition until the production of output that may serve as continuous input for others who acquire the language. Conservatively estimated, this is a period of roughly ten years.¹⁹ Hence 50 generations would mean approximately 500 years. This is a time span for grammar changes that is compatible with time spans found by diachronic studies.

¹⁸ They submitted yeast populations to varying degrees of environmental stress conditions in order to study evolutionary reactions, which is the ability of a population to adapt rapidly through evolution. (McGill University: *Science Daily*, 23 June 2011. <www.sciencedaily.com/releases/2011/06/110622115311.htm>).

¹⁹ The reproduction cycle for language must not be confounded with the concept of "generation" in social sciences (i.e. 25-30 years) since siblings (brothers, sisters, cousins) are effective generators of language input. Bridges &

3.4 Naturally selected vs. drifting.

The linguistic literature on language change focuses on phenomenological properties and does not primarily study them in terms of their *evolutionary* history. Changes may be random (= drift) or channelled by natural selection. In biology, and in particular in population genetics, methods have been developed for identifying the character of changes and for differentiating among them. Linguists typically conjecture potential causalities but do not thoroughly test their hypothesis with such methods because they are still alien to their field.

In the past decade, however, linguists have joined forces with biologists in teams with the mix of competences necessary for tackling such issues. Newberry et al. (2017) apply an analytical technique which tests against random drift as the null hypothesis. The technique²⁰ has been developed in population genetics for discovering the effect of natural selection in microbial populations. It will become clear in the following discussion that a successful transfer of a technique does not automatically generate the appropriate interpretation of the results.

Newberry et. als. (2017) adduce diachronic corpora of English and study two different areas of change, namely the diachronic dynamics of past tense forms (regular – irregular), and the development of *do*-support. For their analysis of morphological changes in past tense formation, they track thirty-six verbs in the 400-million-word corpus of historical American English (CohA), which covers the period of 1810s-2000s. As for *do*-support, they use three parsed corpora.²¹

The interpretation of the result of their investigation of past tense forms of 36 verbs is surprising: "*For six of these verbs we can reject neutral drift*^[22] *for all population sizes N , with nominal $p < 0.05$. Contrary to the standard linguistic expectation, in four of these cases we infer selection towards the irregular variant (dived→ dove, waked→ woke, lighted→ lit, sneaked→ smuck), whereas only two cases exhibit regularization (wove→ weaved, smelt→ smelled).*" Figure (2a) is Newberry's (et al. 2017) figure #2. Figure (2b) is a plot from Masel's (2011: R837) explication of the *drift* of alleles. The graphs look alike, although (2a) plots the subset of verbs that are claimed to testify against drift and for selection.

Figure 2a: Newberry et al. 2017.

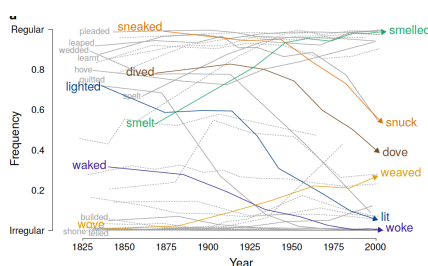
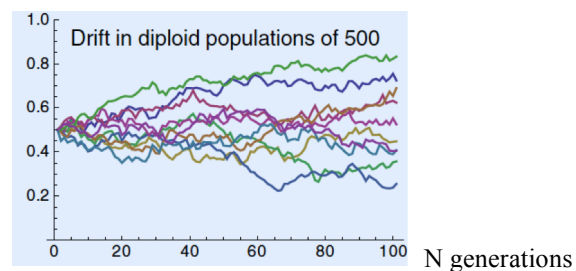


Figure 2b: drift (Masel 2011: R837)



Hoff (2014) found that toddlers with older siblings were more advanced in English language development. Even overheard speech between siblings and parents has a positive effect; cf. Oshima-Takane et al. (1996).

²⁰ *Frequency Increment Test*, as developed by Feder et. als. (2014), based on the standard χ^2 test.

²¹ York-Helsinki Parsed Corpus of Early English Correspondence (1400-1700), the Penn-Helsinki Parsed Corpus of Early Modern English (1500-1700), and the Penn Parsed Corpus of Modern British English (1710-1910).

²² "Neutral drift is the process of change of genotypes by random genetic drift without phenotypic alteration in evolution. It occurs when many genotypes give rise to the same phenotype. In such cases, genotype may change within a given phenotype." <https://www.sciencedirect.com/topics/nursing-and-health-professions/genetic-drift>.

The method of analysis applied by Newberry et. als (2017) tests *against* 'neutral drift', that is, random changes (without phenotypic alterations). "*The frequency increment test (FIT) rejects neutrality if the distribution of normalized allele-frequency increments exhibits a mean that deviates significantly from zero.*" (Feder 2014: 509). The test is not designed for a positive identification of (natural) selection but only for ruling out the null hypothesis, namely random drift. This is sufficient in a situation of a (nearly) dichotomic partitioning of drift and natural selection.

For language change, the situation is not dichotomic. In language use, it cannot be excluded that other factors (e.g. socially-based preferences) influence the choice of a particular past tense form in such a way that the mean of frequency increments deviates from zero. Karjus et al. (in press), who replicate the analysis, point out "*that care should be exercised with interpreting results of tests like the Frequency Increment Test on individual series, given the researchers' degrees of freedom available when applying the test to corpus data, and fundamental differences between genetic and linguistic data.*"

The list of six verbs²³ suspected to be testimonies of (natural) selection, is evidently an inconsistent set. Selection is claimed to have led to opposite outcomes, namely irregular forms as well as regular forms. The 200 years documented by the corpus they use, namely CohA, are not representative since the alternation between 'weak' (= regular) and 'strong' (= irregular) forms can be traced back to the Old English period for many of these verbs.²⁴ Since then (Peters 2004: 293), "*the number of verbs reverting to the regular pattern is much larger than that going the other way. This opposite process can however be seen with hang and sneak (for both past forms).*" "*Dove*" is the exception since it is a singularity. It is a regional innovation, dating back to the 19th century. Outside of North America, the past tense is "*dived*". Figures (3a,b) document search results of the Google nGram-online-viewer [March 23, 2020], restricted to American English books in (3a) and British English books in (3b). It confirms the CocA-based findings.

Figure (3a) shows that the frequency of "*He dove*" steadily increases in the US. It exceeds the frequency of "*He dived*" in the period after 1980. "*It smelled*" is preferred over "*It smelt*" already since a century in the US, while in the UK the lines cross in the sixties. "*Smelled*" is the regular form, "*dove*" is an irregular one. Both directions of change are attested, and in each case, both forms still co-exist. Analogous processes can be documented in any language with regular and irregular verb forms, as for instance in German (Google nGram search), displayed in Figure 4.

Figure (3a): *he dove – it smelled* (USA)

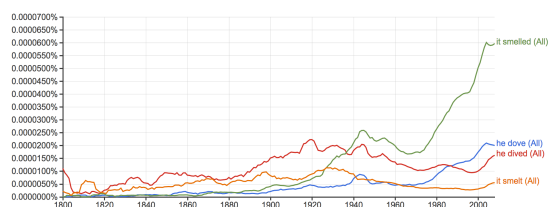


Figure (3b): *he dove – it smelled* (UK)



²³ "Six of these verbs experience selection for either regularization or irregularization."

²⁴ Verbs have individual histories: The alternation between *cleave-cleft-cleft* and *cleave-clove-cloven*, for instance, is a reflex of the time when there existed two homonymous verbs '*cleave*' (Peters 2004:108). The same is true for '*hang*', with '*hanged*' for execution by hanging, and '*hung*' for 'suspended', in the original usage (cf. German *henken – henkte – gehenkt* vs. *hängen – hing – gehangen*).



Figure 4 German: frequency increase of 'gewunken', the irregular past participle of 'winken' (to wave).

Note that the participle form of the particle verb 'zu-winken' (wave to sb.) does not dodge the change of its base form 'winken' since 'zu-gewinkt' remains more frequent than 'zu-gewunken'. Once more, the change is specific to a specific item, namely 'winken'²⁵ and does not generalize.

Fluctuation between alternative tense forms across verbs is an indicator of random variation. It is not a change of a grammatical 'rule' but only a change of a token of the *inventory*. One form is replaced by another for a particular verb. "Dove" is shaped in analogy to a highly frequent and phonetically close verb, namely "drive – drove" (in a subclass with less frequent verbs such as *strive, thrive, or arise*), but crucially, without joining the whole paradigm, that is, *dive – dove – *diven*. The form 'dived' continues to be used as the past participle form. This shows that the switch to "dove" is not a rule-based switch in the paradigm of a verb. The converse change – irregular-to-regular – associates the verb with a morphological rule, as in the case of "smell": *smell – smelled – smelled vs. smell – smelt – smelt*. CohA lists only six tokens of "has smelt" between 1860 and 1960, and nothing thereafter.

Past tense morphology is an example of a persisting change over the last millennium of English that associates forms with rules rather than individually stored information per verb. Standardization and schooling is an antagonistic force. This shows especially in less formal situations, as reported by Gray et al. (2018), who adduced a large corpus of Twitter-data.²⁶ The analysis of the geo-tagged data per county reveals a distinct East-West slope in the US. "In general, western counties show less regularization than average and eastern counties show more, except that the New England area is fairly neutral." (Gray et al. 2018).

A final remark on the *dive-dove* case: It highlights a trade-off relation between storage and rule-application. For highly frequent forms, the cost-benefit-ratio favours storage of individual forms over rule-based formation on demand.²⁷ On the other hand, formation by a rule saves storage costs since per word, only the lemma needs to be stored. For highly frequent forms, the cost of individual storage is less than the cost of frequent activations of a rule. The cost-benefit ratio is the storage cost of the separate forms of an item in comparison with the cost of rule activation multiplied by the number of activations. Associating a form of a verb with a frequent form of related verbs amounts to a free-rider effect on the sub-regularity.

²⁵ Note furthermore, that a paradigm like *winken – winkte – gewunken*, that is, regular past tense but irregular participle does not exist. The model is *trinken* (drink) – *trank* (drank) – *getrunken* (drunk). The analogous form 'wank' is not attested for 'winken', only for 'wanken' (stagger).

²⁶ Their study focuses exclusively on regularization.

²⁷ The 18th century physicist Georg Christoph Lichtenberg noted that the most frequently used verbs are the most irregular ones in all languages: „Diejenigen Verba, welche die Leute täglich im Munde führen, sind in allen Sprachen die irregulärsten; *sum, sono, εἶμι, ich bin, je suis, jag är, I am.*“

Let us turn now to a structure-based change, namely the development of *do*-support in English,²⁸ which took place in the time between the fifteenth and the nineteenth century. In the original caption of this figure, the authors comment that the use of 'do' as an auxiliary verb first arose in the context of interrogative sentences (grey). Drift could be excluded neither for affirmative nor for negative interrogatives. Subsequently, the frequency of *do*-support rose rapidly in negative declarative and negative imperative sentences, what they interpret as a result of selection.

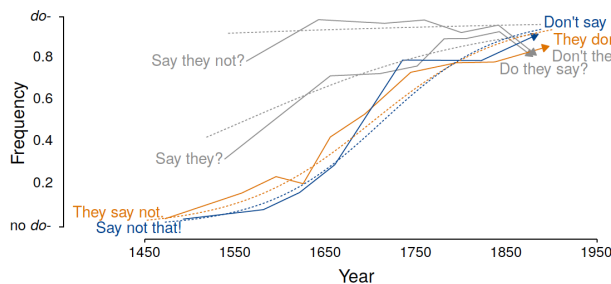


Figure 5: (= figure 3 of Newberry et. als 2017): frequency increase of *do*-support.

In an interview,²⁹ one of the authors, Plotkin, summarizes their findings: “*It seems that, once ‘do’ was introduced in interrogative phrases, it randomly drifted to higher and higher frequency over time. Then, once it became dominant in the question context, it was selected for in other contexts, the imperative and declarative, probably for reasons of grammatical consistency or cognitive ease.*”

What this quote describes is the spread of a grammar variant after a period of free variation. In this phase, a grammar variant gets momentum when the main verb in English is fixed to its base position. It is not fronted anymore to the clause initial position in questions, and subsequently it is not fronted anymore across negation to the position that in present day English, is open only for finite auxiliaries and quasi-auxiliaries. In the middle of the nineteenth century, finally, the grammar variant has successfully 'infected' the whole population.

Within Germanic languages, English *do*-support is an isolated peak in the evolutionary fitness landscape, in Wright's (1932) terminology. Its computational advantage is evident. Main verbs are confined to a unique structural position, namely the head-position of the verb phrase. No filler-gap dependency needs to be computed for verbal heads since the secondary positions for finite verbs are the domain of finite auxiliaries and the dummy auxiliary '*do*'. The 'price' for the advantage of a predictable and invariant structural position of the main verb is the recruiting of a dummy auxiliary for the structures in which a finite verb is required in a fronted position.³⁰

In Generative-Grammar terms, the change started when the use of '*do*' began to be construed as indication that the finite verb is not fronted anymore to the clause initial functional head position. Within a comparatively short period, the grammatical consequences triggered by the consistent implementation of this change become visible: If the finite main verb is not fronted to the clause-initial functional head position, it cannot be fronted to the clause-medial functional

²⁸ It is a textbook example of the grammaticalization of a variation attested in numerous languages; see Jäger (2007), who describes the phenomenon of periphrastic '*do*'-constructions in a sample of two hundred languages.

²⁹ <https://penntoday.upenn.edu/news/luck-plays-role-how-language-evolves-penn-team-finds>

³⁰ If this is an advantage indeed, it should not pose problems during language acquisition, and it does not. In a corpus study, Stromswold (1990) found that on average, the age of the first use of the auxiliary *be* is 2;7, followed by *do*-support at 2;8 while auxiliary *have* is the last. See also Rispoli et al. (2012) for similar findings.

head position (i.e. tense and agreement) either, because the two positions are automatically connected. A finite verb in the lower position is the candidates for the higher position in interrogative constructions. Consequently, *do*-support becomes grammatically mandatory for negated sentences as well. What figure 5 reflects is the initial co-existence of grammar variants, most of which are sieved out within roughly twenty linguistic generations.

4. Natural selection or social change, or both

Language is a neuro-cognitive as well as a social phenomenon. The neuro-cognitive side is the domain of Darwinian selection; the social side is the domain of social change of cultural symbols, rules of behaviour, or value systems. In language change, clear instances of each of these factors occur, and often both kinds of changes are at work simultaneously, see Table (1).

Table 1: D-selected (Darwinian selection) vs. socially motivated

	+ D-selected	– D-selected
+ socially selected	Darwinian & social change	social change
– socially selected	Darwinian change	(other, e.g. flow)

Differentiating between the respective domains requires at least two combined criteria, namely the *time course* and the *declarative/procedural* basis of a change. Social changes are rapid, and they concern accessible properties, that is, items of the inventory. Changes based on selection concern the procedural organization and are slower. Declaratively stored content comprises not only the lexicon but also the morpho-syntactic inventory, including affixes, pronouns, particles, etc. The latter items are tightly governed by rules of grammar and therefore more change resistant. The general vocabulary, can be (ex-)changed or augmented easily during the life-time of a speaker. However, a *specific* subset of the vocabulary – the so-called basic vocabulary³¹ – is fairly stable across a community and over time. This has recently been confirmed in the study of 81 languages of the Pacific area, by Greenhill et. al. (2017: E8822), who use 210 items of basic vocabulary and 157 grammatical features: "*We show that, on average, most grammatical features actually change faster than items of basic vocabulary.*" Such finding replicate results known from Indo-European studies. Present day Indo-European languages differ in their structures and morpho-syntax, but still share a basic vocabulary.

What follows are three illustrations of differences between social changes and changes by selection. The plots are Google NGram search results of book corpora (March 2020). The changes depicted in Figure (6a,b) are social changes. In German, "*am Ort*" (at-the location) gets recently replaced by "*vor Ort*"³² (lit. before location), and "*Fakt*" ('fact') is preferred over "*Faktum*" (Latin loan), especially in the frequent expression "*Faktum ist*" (*fact is ... = 'It is a fact, ...'*). These are rapid, item-based changes that have become effective within approximately two decades. This is characteristic of the socially motivated expansion of 'trendy' expressions in a language community.

³¹ "Basic vocabulary", as defined by Collins on-line dictionary: "*The set of lexical items in a language that are most resistant to replacement.*"

³² Originally, this is a technical term of mining, meaning that the tunnelling has arrived at a spot "immediately before the ore deposition", but not yet "at the deposition". Vernacular usage treats the terms as synonymous.

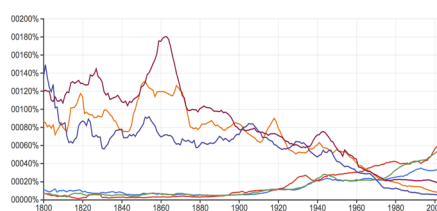
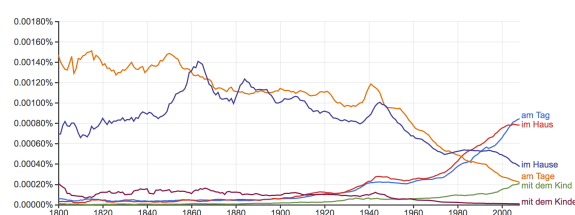
Figure (6): a. "am Ort" → "vor Ort"³³

b. "Faktum ist" → "Fakt ist"



The second example is a change in case marking morphology of datives. The ending "-e", which dates back to Middle-High-German, is cancelled. The change becomes manifest only after the middle of the 20th century, a reason being that until then, schooling enforced a norm with the Dative ending.³⁴ This is a rule-based change antagonised by a social factor, namely the normative force of schooling.

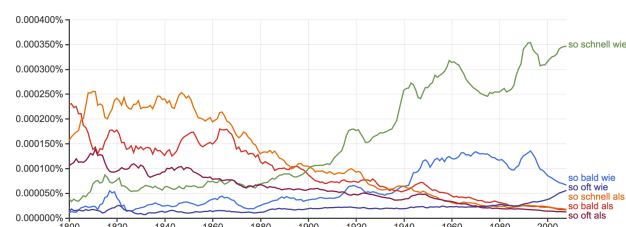
Figure (7): Dative with "-e" → without "-e"



(glosses: *am Tag*: at-the day; *im Haus*: in-the house, *mit dem Kind*: with the child; *dem*: Dative def. article)

The third example is the "as A as Y" construction (e.g.: *as fast as Y*) in German (Figure 8). In present day German, the preferred version is "so A wie Y" (lit.: so A how Y). "Wie" (*how*) replaces "als" (*as, than*), that is, "so schnell als" (*as fast as*) is supplanted by "so schnell wie" (*lit.: as fast how*). What looks like a simple replacement of "als" (*as*) by "wie" (*how*) is in fact a more complex change consisting of structure reduction plus subsequent integration in the wider grammatical fabric.³⁵ The constructions "so A wie" and "so A als" are continuations of "so A als wie"³⁶ (*so A as how*), which is reduced by dropping either "als" or "wie", resulting in two competing variants. The "als wie"-variant is still attested in Southern German vernacular, since it is a regular variant in Bavarian dialects (s. Merkle 1975: 171). Although such a change is a plausible candidate for a selection effect, at least for the first step, that is the reduction of structure, it is still open for social influence, such as normative regulations in schooling: "Combine *so* with *wie*, and comparatives with *als*!"

Figure 8: German: "so A als" → "so A wie"



so schnell **wie** – so fast as
 so bald **wie** – so soon as
 so oft **wie** – so often as
 schnell **als** – so fast than
 so bald **als** – so soon as
 so oft **als** – so often as

³³ "An Ort und Stelle" (lit.: at place and spot) is a common idiomatic expression.

³⁴ e.g.: *Schulgrammatik des Deutschen* (= school grammar of German), by Karl F. Becker, 1845, p. 126, §136, Frankfurt: Verlag von G.F. Kettenteil.

³⁵ Direct replacement of "als" by "wie" would be a change between incongruent categories, since "als" is a preposition and "wie" is a wh-pronoun.

³⁶ Goethe (Faust): "und bin so klug *als wie* zuvor" – and (I) am so wise *as how* before)

Let us turn now to clear cases on the other side of the spectrum, namely clear cases of natural selection operating on the cognitively represented grammar systems. A century ago, Sapir (1921:174, 177, 180) has identified three grammatical 'megatrends', namely the "drift" to fixed position, to the levelling³⁷ between the subject and the object, and toward the invariant word. These three processes are entangled.³⁸ These are changes that on the one hand, *shift the working load* from the declarative to the *procedural* network in production and reception and on the other hand enhance *predictability* in parsing (Levy 2008).

The "drift" to fixed position is a "drift" towards head-initial phrases and an [S[VO]] clause structure. In this structural constellation, each item is uniquely defined by its structural position. As a consequence, morphological markings for grammatical functions such as subject, object or indirect objects become redundant and can be reduced. The result is a language with fixed word order, levelled subject-object relations (= "neutral alignment" in typological terms) and invariant rather than inflected nouns and verbs.

This scenario is well-documented in the history of languages and in particular in the history of Germanic languages. Siewierska (1996), in a study based on a sample of 237 languages, notes that "*in line with common assumptions, neutral alignment in V2 languages is more frequent than in any other word order type; it occurs in just over two-thirds of the V2 languages in the sample*". "Neutral alignment" is the term for "no morphological case distinction" between subject and object, and "V2" is her idiosyncratic term for grouping together SVO and OVS.³⁹ In reality, the correlation between SVO and case-less is much higher since typological surveys classify languages merely by linearization patterns, irrespective of their clause-structure. As a consequence, many languages are (mis-)classified as SVO which are structurally not [S[VO]]; see for instance the supplement⁴⁰ to Gell-Mann & Ruhlen (2011).

What is the cognitive and therefore selectively effective bonus of structural coding? In an [S[VO]] language, grammatical relations are structurally defined. The (obligatorily lexicalized) preverbal argument position is structurally identified as the subject position, followed by the verb, followed by the indirect object and the direct object, without any intervening material between the verb and its objects (3a). The grammatical relation of a noun phrase is identified by a rule that maps linear order on a structure. In a relationally coding language (3c), grammatical relations of noun phrases are identified by their morphological shape and they are freely ordered, as for instance in German (3d) or Polish (3e). In languages that identifies noun phrases

³⁷ "levelling between subject and object" = abolition of morphological case distinctions (s. Sapir 1921:714)

³⁸ "*The drift toward the abolition of most case distinctions and the correlative drift toward position as an all-important grammatical method are accompanied, in a sense dominated, by the last of the three major drifts that I have referred to. This is the drift toward the invariable word.*" (Sapir 1921:180).

³⁹ "OVS" is a very infrequent type. In fact, it is "SVO" with an Ergative-Abs case-system (e.g. Pari, Jur Luo, Hishkaryana, Asurini, Oiampi). Since typologists tend to identify "O" semantically, that is, by the "patient" role, they fail to properly appreciate the fact, that in an Erg-Abs system, the "patient" is the structural *subject*. As the structural subject, it precedes the verb in an SVO clause structure. An ergative "OVS"-language like Hurrian or Kuikuro is in fact an SOV-language, modulo ergative alignment. Erg-V-Abs languages do not exist.

⁴⁰ <http://www.pnas.org/content/suppl/2011/10/04/1113716108.DCSupplemental/sapp.pdf>. Afrikaans, Frisian, German, and Luxemburgish are listed as SVO, although all these languages as [SOV & V2]-languages, just like Dutch, which is filed as "SOV/SVO". Except for Sorbian, all Slavic languages are listed as SVO, although no Slavic language shares the syntactic properties of structural SVO languages, except for the S-V-O linearization as one of many grammatical serialization variants in these languages (see Haider & Szusich 2019). See also fn. 41.

by their morphological case, the acquisition of case morphology and the actual identification is an arduous task, as everyone knows who has acquired such a language after childhood. It usually requires a lot of cross-linking declaratively stored information that needs to be memorized before.

- | | |
|--|--------------------|
| (3) a. [subject [verb _{fin.} object ₁ [object ₂]]] | structurally coded |
| b. that someone envies someone something | |
| c. [... subject _{Nom} ... object _{Dat} ... object _{Acc} ... verb _{fin.}] | relationally coded |
| d. dass dem Mann _{Dat} das Problem _{Akk} jemand _{Nom} erklären muss | (German) |
| dass the man the problem someone explain seized | |
| e. że mężczyźnie _{Dat} dom _{Acc} pokazał Jacek _{Nom} | (Polish) |
| that (to) man house showed Jacek | |

The order in (3e) is one of 24 *grammatically* admissible orders⁴¹ of four elements in Polish and in other Slavic languages. In addition to the free relative order of the arguments, the verb may appear in any of the four possible positions after the complementizer. Hence, linear order does not provide reliable cues for alignment.

The crucial gain of structural coding is the purely structural implementation of alignment. If a grammatical rule changes a relation, as in the case of passive, the structural position changes as well. Structural coding not only reduces the amount of information involved in the identification of grammatical relations. The strict linear order restrictions greatly enhances the accuracy of look-ahead predictions in on-line parsing and thereby enhances its efficiency.

The gain is high enough for lifting an [S[VO]] grammar into a very stable state. The inverse route is not attested, as Gell-Mann and Ruhlen (2011: 17291, fig.1) claim, based on their broad survey. When a language has attained an [S[VO]] structure, it stays [S[VO]]. No language is known that has started out from a system like English and ended up with a grammar like Sanskrit. Many present-day [S[VO]] languages are known to have been SOV, VSO or free-word-order languages in their past, however.

If the [S[VO]]-architecture is a strong attractor, this means that it is strongly selected. If this is the case, then – in terms of evolution theory – any change that gives up the [S[VO]]-architecture would be a change that reduces the "fitness" of a grammar. This state of affairs is in line with Fisher's (1930) Fundamental Theorem of Natural Selection, which implies that the mean relative fitness of a population cannot decrease during evolution, cf. Orr (2009: 531), Koonin (2011:8). This is a necessary consequence if selection constantly sieves out the less adapted variants. Selection does not eliminate a better adapted variant and therefore the fitness peak reached with an [S[VO]] clausal architecture remains stable.

Typologists acknowledge a trade-off between strict word order and rich morphology. Some languages express argument structure, sentence type and other categories primarily by morphological marking. Other languages use word structure for the same functions. Nichols (1992) produced pioneering work, systematically quantifying morphological complexity in a sample of 200 languages, based on grammatical descriptions. Koplenig et al. (2017) present an analysis based on 1.559 Bible translations in 1.196 languages, in a quantitative information-theoretic

⁴¹ "In each of the Slavic languages, all twenty-four possible combinations of a subject, direct object, indirect object and verb occur as grammatical declarative orders." Siewierska & Uhliřová (2010:109).

approach: "*Languages that rely more strongly on word order information tend to rely less on word structure information and vice versa*".

In a diachronic perspective, the crucial point is that the trade-off is *unidirectional* rather than *bi-directional*. Grammars do not fluctuate between the two poles (= word order only vs. morphology only). The unidirectionality of change from morphology to structure is a consequence of selection. It is the unidirectional change toward [S[VO]] that invites the reduction of morphological marking, and not the loss of morphology that invites the structural change, as Siewierska (1996)⁴² or Gibson et al. (2019)⁴³ seem to presuppose.

In fact, there is no trade-off. The change to structural coding makes morphological coding redundant but does not replace it. Icelandic is an [S[VO]] language whose word order restrictions on objects are as strict as in English or Danish (see Dehé 2004:94), in spite of its rich nominal and verbal morphology. Bulgarian and Macedonian, on the other hand, are languages with the typical free word order of Slavic languages but without nominal case morphology or case particles (Avgustinova 1997, Wahlström 2015). The first Icelanders were Norwegian settlers. Norwegian has lost most of its nominal and verbal morphology since this time. Icelandic has preserved its old-to-middle Norse morphology but nevertheless has developed into an SVO language. Bulgarian and Macedonian are linguistically closely related to their neighbouring Slavic languages, all of which have preserved their morphology. Nevertheless, they share their clause and phrase structure and concomitant word order freedom.

What Icelandic confirms, however, is a law of population genetics, namely Fisher's (1930) fundamental theorem: "*The rate of increase in fitness of any organism at any time is equal to its genetic variance in fitness at that time.*"⁴⁴ In other words, the intensity of selection and hence, the rate of evolutionary change due to selection is proportional to the magnitude of variation in an evolving population, which, in turn, is proportional to the effective population size (Koonin 2011: 7). This accounts for the fact that languages in small populations, such as Icelandic, Faroese, or Logudorese Sardinian, change less than, for instance, Swedish, Danish (compared to Icelandic), or Italian (compared to Sardinian), although they are offspring of a common ancestor language in each case. A small population confined to a small region produces less variation and therefore less chance for change. The alleged pressure for effectivity would hold for any language, regardless of the size of its population. In fact, it could much faster reach every member of the smaller speech community. Social changes spread fast, changes by natural selection are gradual and relatively slow.

5. Consequences

The focus on particulars of changes must not divert the attention from the general consequences of Darwinian evolution for language change. If grammars are cognitive systems with a Darwinian evolutionary history, that is, a history governed by the interplay of variation & selection,

⁴² "*The loss of case marking is typically considered to underlie the change from SOV to SVO order in English, the Scandinavian and the Romance languages.*" (Siewierska 1996, fn.1). Jespersen (1894: §75, 96-97) has explicitly rejected such an idea: "*A fixed word order was the prius, or cause, and grammatical simplification, the posterus, or effect*".

⁴³ "*If a language has morphology, then it does not need fixed word order.*" (Gibson et al. 2019: 390).

⁴⁴ Later, Fisher restated it: "*The rate of increase in the average fitness of a population is equal to the genetic variance of fitness of that population.*" Li (1964: 505).

the general characteristics of Darwinian evolution will inevitably have to show as well. These properties separate evolutionary changes from changes with a different aetiology, as for instance socio-cultural alterations (which may nevertheless prepare the ground for Darwinian changes).

- i. Evolution eventually leads to speciation.
- ii. Evolution results in a 'hilly' *adaptive landscape*.
- iii. Evolution may be convergent (because of constant selection).
- iv. Evolution does not trigger changes but constrains them.
- v. Evolution is often punctuated by long periods of little or no change.

Speciation: Darwin's title phrase "*on the origin of species*" points to the long-term outcome of evolution, namely speciation as a consequence of natural selection. If "individual language" is the parallel of "individual species", then dialect split plus the subsequent development into separate languages is a process of speciation, and it happens under the same conditions, namely allopatric (due to geographic isolation), parapatric, or even *sympatric*, that is, *within* a speech community. A particular example is a "complete linguistic system whose use is determined by the gender affiliation of the speech participants" (Dunn 2014: 40). Like in biology, a particular variant is specialized for a particular group of hosts.

Adaptive landscape: The general picture is familiar from biology (Skipper & Dietrich 2012). There are universal properties and there are type-specific ones. Types are clusters of properties that surround a strong attractor. Evolution of grammar accounts for the emergence of different types rather than uniform members of a universal format of grammars.

Evolution operates on a pool of randomly originating variants. It is not goal-directed. As a consequence, system changes that result from constant selection lead to different types of systems. There is no single universal grammar, but there are universal within-group properties. If grammars evolve in a Darwinian way, there is room for linguistic universals. First, there is negative selection, that is, certain properties are always sieved out. The remaining properties are not imposed on languages by a pre-configured "universal grammar"; they unavoidably emerge by negative selection. Such universals are core characteristics any grammar must meet in order to be learnable and processable, given the specific cognitive "system software" a human brain uses for language. Second, there are positively selected properties. These are reflected as typological generalizations. Typological generalization means a type-specific clustering of properties. In evolutionary terms this means that changes are not deterministic since there are alternative attractors, each one with consequential effects of its own. These different sets of consequential effects of each attractor we perceive as types.

In linguistics, two opposing positions dominate the discussion. On the one hand, diachronic change is explained as switching between options provided by a "Universal grammar" (UG) in the course of imperfect acquisition of grammar. This is the perspective of the Generative school of thought. On the other hand, cross-linguistically recurrent structural patterns in grammar are explained as results of recurrent patterns of change fuelled by communicative needs. This is the position of the functional-typological school of thought. In its consequent form, the program seeks to replace UG and explain cross-linguistic invariants as by-products of language change.

In the absence of a theory that explains why certain changes are recurrent while other potential changes are not, this account remains circular,⁴⁵ of course.

Already in the title of his paper, Kiparsky (2008) suggests a compromising position, namely "*universals constrain change; change results in typological generalizations.*" In his view, "*any structural feature that is caused by change is inherently unstable (vulnerable, as Saussure put it). Therefore, recurrent structural features that are caused by recurrent patterns of change are typological generalizations but not true universals.*"

If it is true that changes lead to "*inherently unstable*" states by disturbing the balance of a system, the natural consequence would be that such an unstable system returns to its stable state after a while, or to a new stable state. In evolution, any self-reproductive system is potentially dynamic. In the absence of a positively selected 'competing' variant, a system is stable and protected by negative selection. Since variation is random and not triggered, changes are not predictable and not organized. Once a beneficial variant appears and is selected, a change is initiated if the variant spreads.

The difference between universal features and typological generalizations is one between necessary and possible properties. Typological generalizations are correlations that characterize positively selectable constellations of features that result in stable systems. Since types differ, features of one type may be excluded by other types. Universals – by definition – are features that are never excluded. Any variant that excludes one of these features will qualify as a less fit variant and fall prey to negative selection. How could this be? Features are not selected in isolation, just like genes are not selected in isolation. The subject of selection is the phenotype as the system in action, not the genotype. In the case of grammar, the phenotype is the grammar as it is put to use. The genotype is its neuro-cognitive representation as a computational program which is compiled during language acquisition. A single feature could not be put to use in isolation. It is always embedded in a complex program. Universal features are features of components of these programs without which the program would not work. These are embedded features that cannot be removed or changed in isolation because they are in essential interaction with other features. Any change of such features would be a 'bad' or deleterious mutation; cf. Loewe, and Hill (2010: 1153).

Convergent evolution: It is an open issue whether human languages have a monogenetic or polygenetic history. The time depth for reliable empirical studies on grammars is less than 5k years. The history of homo sapiens dates back at least 300k years (Stringer & Galway-Witham 2017). Since languages do not leave fossils, there is no scientific way of deciding the question of the origin. A unique origin followed by long periods of diversification is not more plausible than a long-period of pre-grammatical pidgin-like communication followed by long periods of grammaticalization with *convergent evolution* of initially different grammar systems. It is highly unlikely that the linguistic capacity of homo sapiens is the result of a unique, huge genetic saltation that spontaneously endowed the new species with the fundamentals of Universal grammar. There is no evidence for Chomsky's evolutionary conjecture, namely a single-mutant theory of human language (De Boer et al. 2020).

⁴⁵ A particular pattern is recurrent if it is the result of a recurrent change. A change is recurrent, if it recurrently produces a particular pattern.

On the other hand, there is ample biological evidence for convergent evolution. For instance, most palaeontologists trace whales, porpoises and dolphins back to Pakicetus, a land-dwelling hoofed mammal. Today, the descendants superficially look like prototypical fish, with streamlined bodies, dorsal fins and flippers. This is a result of convergent evolution (Foote et al. 2015). Adaptive changes lead to structures that have similar forms and functions but were not present in the last common ancestor of those groups. Grammars, just like sea-dwelling mammals, adapt to their respective habitat. In the course of evolution, common superficial characteristics will emerge that are not inherited from any of the ancestors. Since evolution is not goal-driven, languages may even differ in fundamental⁴⁶ grammatical respects.

Untriggered changes: Conjectures that conceive of language changes as products of 'social engineering', driven by "a pervasive pressure for efficiency" that "guides the forms of natural language" (Gibson et al. 2019: 389) in a permanent process of enhancing the efficiency of the social tools of communication neglect the fact that grammars are stable. Shifting complex phrases to the end of a clause (i.e. "extraposition"), for instance, is a cross-linguistically widely implemented way of enhancing efficiency of parsing and production (cf. Hawkins 1994). Nevertheless, thousands of years of an alleged "pervasive pressure" were not enough for turning strict SOV languages into languages that permit extraposition.

Evolution does not work like a steadily grinding mill, as Eldredge & Gould (1972) emphasized. Systems with a shorter evolution evolve mostly by *punctuated equilibrium*, and those with a longer evolution evolved mostly by *gradualism*. Punctuated equilibrium means that after a period of very little alteration one or a few major changes occur. Such major changes close many previously available options forever but open new ways for future possibilities to unfold within the limits of the inherited design. Dixon (1997) proposed such an account, evaluated critically by Joseph (2001).

In sum, a theory of grammar change is a theory of the (neuro-)cognitive evolution of grammar systems. In a procedural/declarative model of grammar, the cognitive encapsulated and therefore inaccessible procedural components are the domain of Darwinian evolution. Declaratively stored content, on the other hand, is consciously accessible and therefore open for various kinds of interventions. This is the domain for social factors, although Darwinian evolution is not principally excluded from these domains since retrieval of declaratively stored information is also a procedural capacity.

Bibliography

- Avgustinova, Tania 1997. Word order and clitics in Bulgarian. PhD dissertation. Saarbrücken: Universität des Saarlandes.
- Bell, Graham & Andrew Gonzalez 2011. Adaptation and evolutionary rescue in metapopulations experiencing environmental deterioration. *Science* 332 (6035): 1327-1329.
- Breslin, D. (2010). Generalising Darwinism to study socio-cultural change. *International Journal of Sociology and Social Policy*. 30(7-8): 427-439.

⁴⁶ Salish languages (Jelinek & Demers 1994) seem to have conserved a design that has been changed in most other languages. Salish languages lack lexical categories. Although this impedes phrase-structuring, the grammar has not changed. Needs do not create grammars.

- Bridges, Kelly & Hoff, Erika 2014. Older sibling influences on the language environment and language development of toddlers in bilingual homes. *Applied Psycholinguistics* 35(2): 225–241. doi:10.1017/S0142716412000379.
- Christiansen, M., & Chater, N. 2016. The Now-or-Never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences* 39, E62. doi:10.1017/S0140525X1500031X.
- Comrie, Bernard & Tania Kuteva 2005. The evolution of grammatical structures and ‘functional need’ explanations. In: Tallerman, Maggie (ed.) *Language origins: perspective on evolution*. 185-207. Oxford University Press.
- De Boer, Bart, Thompson, Bill, Ravignani, Andrea, Boeckx, Cedric 2020. Evolutionary dynamics do not motivate a single-mutant theory of human language. *Nature Science Reports* 10: 451. <https://doi.org/10.1038/s41598-019-57235-8>
- Darwin, Charles 1872.⁶ *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. London: John Murray (6th edition).
- Dehé, Nicole 2004. On the order of objects in Icelandic double object constructions. *UCL Working Papers in Linguistics* 16: 85-108.
- Dixon, Robert 1997. *The rise and fall of languages*. Cambridge: Cambridge University Press.
- Dunn, Michael 2014. Gender determined dialect variation. In: Corbett, Greville G. (ed.) *The expression of gender*. 39-67. Berlin: Walter de Gruyter.
- Ellegård, Alvar. 1953. *The auxiliary Do: The establishment and regulation of its use in English*. (Gothenburg Studies in English vol. 2). Stockholm: Almqvist, Wiksell.
- Eldredge, Niles and S. J. Gould 1972. Punctuated equilibria: an alternative to phyletic gradualism. In: Schopf, Tom J.M. (ed.) *Models in Paleobiology*. 82-115. San Francisco: Freeman Cooper.
- Feder, Alison F., Sergey Kryazhimskiy, and Joshua B Plotkin 2014. Identifying signatures of selection in genetic time series. *Genetics* 196(2): 509-522.
- Fisher, Ronald A. 1930. *The genetical theory of natural selection*. Oxford, UK: Clarendon Press.
- Footo Andrew, Liu, Yue, Thomas, Gregg W. et al. 2015. Convergent evolution of the genomes of marine mammals. *Nature Genetics* 47(3): 272–275. doi:10.1038/ng.3198
- Gazsi, Dénes 2011. Arabic-Persian Language Contact. In: Stefan Weninger (ed.) *The Semitic Languages: An International Handbook*. (Handbook of linguistics and communication science 36). 1015-1021. Berlin: De Gruyter.
- Gibson, Edward, Richard Futrell, Steven T. Piandadosi, Isabelle Dautriche, Kyle Mahowald, Leon Bergen, and Roger Levy 2019. How efficiency shapes human language. *Trends in Cognitive Sciences*. 23(5): 389-407.
- Gell-Mann, Murray and Merritt Ruhlen 2011. The origin and evolution of word order. *PNAS* 108(42): 17290-17295.
- Goulding, John 2020. *Virus replication*. British Society for Immunology. <https://www.immunology.org/public-information/bitesized-immunology/pathogens-and-disease/virus-replication> [March 15, 2020].
- Gray Tyler J., Reagan A.J., Dodds P.S., Danforth C.M. 2018. English verb regularization in books and tweets. *PLoS ONE* 13(12): e0209651. <https://doi.org/10.1371/journal.pone.0209651>
- Greenhill, Simon J., Chieh-Hsi Wu, Xia Hua, Michael Dunn, Stephen C. Levinson, and Russell D. Gray 2017. Evolutionary dynamics of language systems. *PNAS* 114(42): E8822-E8829.

- Haider, Hubert 2015. “Intelligent design” of grammars – a result of cognitive evolution. In: Aria Adli & Marco García García & Göz Kaufmann (eds.): *Variation in language: System- and usage-based approaches*. 205-240. Berlin/New York: de Gruyter.
- Haider, Hubert & Szusich, Luka 2019. *Slavic languages – "SVO" languages without SVO qualities?* lingbuzz/004973. (submitted and under review).
- Haig, David 2007. Weismann Rules! OK? Epigenetics and the Lamarckian temptation. *Biology and Philosophy*. 22:415–428. DOI 10.1007/s10539-006-9033-y.
- Hanke, David 2004. Teleology: the explanation that bedevils biology. In: John Cornwell (ed.) *Explanations: Styles of explanation in science*. 143-155. Oxford 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.
- Heath, Jeffrey 1984. Language contact and language change. *Annual Review of Anthropology* 13: 367-384. Palo Alto: Annual Reviews, Inc.
- Hodgson, Geoffrey M. 2011. Organizational Adaptation and Evolution: Darwinism versus Lamarckism? In: Anna Grandori (ed.), *The Elgar Handbook of economic organization*. 157–171. Cheltenham, UK: Edward Elgar.
- Hodgson Geoffrey M. & Knudsen, Thorbjørn 2008. In search of general evolutionary principles: Why Darwinism is too important to be left to the biologists. *Journal of Bioeconomics* 10: 51–69. <https://doi.org/10.1007/s10818-008-9030-0>.
- Hodgson, Geoffrey M. and Knudsen, Thorbjørn 2006. Why we need a generalized Darwinism and why a generalized Darwinism is not enough. *Journal of Economic Behavior and Organization* 61:1-19.
- Hull, 1988. Interactors versus vehicles. In: Plotkin Henry C. (ed.) *The role of behavior in evolution*. 19-50. Cambridge, Mass.: MIT Press.
- Jäger, Andreas 2007. Grammaticalization paths of periphrastic 'do'- constructions. *Studies van de Belgische Kring voor Linguïstiek* 2. <https://sites.uclouvain.be/bkl-cbl/nl/publicaties/studies-van-de-bkl/volume-2-2007/>.
- Jelinek, Eloise, and Richard Demers 1994. Predicates and pronominal arguments in Straits Salish. *Language* 70: 697–736.
- Jenkin, Fleeming 1867. The origin of species. *North British Review*, 46:277–318.
- Jespersen, Otto 1894. *Progress in language*. New York: McMillan & Co.
- Johansson, Sverker 2005. *Origins of Language: Constraints on hypotheses*. Amsterdam: John Benjamins.
- Joseph, Brian 2001. Review of R. M. W. Dixon, *The rise and fall of languages*. *Journal of Linguistics* 37: 180–6.
- Karjus, Andres, Richard A. Blythe, Simon Kirby, Kenny Smith (in press). Challenges in detecting evolutionary forces in language change using diachronic corpora. *Glossa*.
- Kiparsky, Paul 2008. Universals constrain change; change results in typological generalizations. In: Jeff Good (ed.) *Linguistic universals and language change*. 23-53. Oxford: Oxford University Press.
- Koonin, Eugene V. 2011. *The logic of chance: The nature and origin of biological evolution*. Upper Saddle River, New Jersey: FT Press Science.

- Koplenig Alexander, Meyer P., Wolfer S., Müller-Spitzer C. 2017. The statistical trade-off between word order and word structure – Large-scale evidence for the principle of least effort. *PLOS ONE* 12(3): e0173614. <https://doi.org/10.1371/journal.pone.0173614>.
- Lamarck, Jean-Baptiste Pierre Antoine de Monet de 1809. *Philosophie zoologique, ou, exposition des considérations relative à l'histoire naturelle des animaux*. Paris: F. Savy.
- Lasnik, Howard, Sobin, Nicholas 2000. The who/whom puzzle: On the preservation of an archaic feature. *Natural Language & Linguistic Theory*. 18: 343–371. <https://doi.org/10.1023/A:1006322600501>
- Lee, Namhee 2004. The neurobiology of procedural memory. In: John H. Schumann, Sheila E. Crowell, Nancy E. Jones, Namhee Lee, Sara Ann Schuchert, and Lee Alexandra Wood (ed.) *The neurobiology of learning: Perspectives from second language acquisition*. 43- 73 Mahwah, NJ: Erlbaum.
- Lenormand, Thomas 2002. Gene flow and the limits to natural selection. *Trends in Ecology & Evolution* Vol.17(4):183-189.
- Levy, Roger 2008. Expectation-based syntactic comprehension. *Cognition* 106(3): 1126-1177.
- Loison, Laurent 2018. Lamarckism and epigenetic inheritance: a clarification. *Biology & Philosophy* 33(3-4), Art. 29.
- Loewe, Laurence and William G. Hill 2010. The population genetics of mutations: good, bad and indifferent. *Philosophical Transactions of the Royal Society. B* 365: 1153-1167. doi:10.1098/rstb.2009.0317
- Loewe, Laurence 2008. Negative selection. *Nature Education* 1(1): 59.
- Masel, Joanna 2011. Genetic drift. *Current Biology*. 21(20): PR837-R838.
- Mayr, Ernst 1992. The idea of teleology. *Journal of the history of ideas*. 53(1): 117-135.
- Merkle, Ludwig. 1975. *Bairische Grammatik*. München: Hugendubel.
- Morgan-Short, Kara & Michael T. Ullman 2011. The neuro-cognition of second language. In Gass, Susan M. and Alison Mackey (ed.) *The Routledge handbook of second language acquisition*. 282-299. New York: Routledge.
- Morris, Susan 1994. Fleeming Jenkin and the origin of species: A reassessment. *The British Journal for the History of Science*, 27(3), 313-343. doi:10.1017/S0007087400032209
- Nelson, Richard R. 2007. Universal Darwinism and evolutionary social science. *Biology and Philosophy* 22: 73–94.
- Nettle, Daniel 1999. *Linguistic diversity*. Oxford: Oxford University Press.
- Newberry, Mitchell G., Christopher A. Ahern, Robin Clark, Joshua B. Plotkin. 2017. Detecting evolutionary forces in language change. *Nature* 551(7679): 223-226. DOI: 10.1038/nature24455.
- Nichols, Johanna 1992. *Linguistic diversity in space and time*. Chicago: University of Chicago Press.
- Orr, H. Allen 2009. Fitness and its role in evolutionary genetics. *Nature Reviews Genetics*. 10: 531–539.
- Oshima-Takane, Yuriko, Elizabeth Goodz and Jeffrey L. Derevensky 1996. Birth order effects on early language development: Do second born children learn from overheard speech? *Child Development* 67(2): 621-634.
- Paradis, Michael 2009. *Declarative and procedural determinants of second languages*. Amsterdam: John Benjamins.
- Peters, Pam 2004. *The Cambridge guide to English usage*. Cambridge: Cambridge University Press.

- Pinker, Steven & Paul Bloom 1990. Natural language and natural selection. *Behavioral & Brain Sciences* 13: 707-784.
- Reesink, Ger, Singer Ruth, Dunn Michael 2009. Explaining the linguistic diversity of Sahul using population models. *PLoS Biology* 7(11): e1000241. doi:10.1371/journal.pbio.1000241.
- Rispoli, Matthew, Hadley, Pamela, and Holt, Janet 2012. Sequence and system in the acquisition of tense and agreement. *Journal of Speech, Language, and Hearing Research* 55: 1007–1021. doi: 10.1044/1092-4388(2011/10-0272)
- Roossinck Marilyn J. 2011. The good viruses: viral mutualistic symbioses. *Nature Review of Microbiology* 9: 99–108.
- Roossinck, Marilyn J. and Edelio R. Bazán 2017. Symbiosis: viruses as intimate partners. *Annual Review of Virology*. 4(1): 123-139.
- Rubio Luis, Guerri J., and Moreno P. 2013. Genetic variability and evolutionary dynamics of viruses of the family Closteroviridae. *Frontiers in Microbiology* 4:151.
- Sapir, Edward 1921. *Language an Introduction to the Study of Speech*. New York: Harcourt, Brace & Co.
- Schleicher, August 1873. *Die Darwinsche Theorie und die Sprachwissenschaft*. Weimar: H. Böhlau.
- Schoenemann, P. Thomas 2012. Evolution of brain and language. In Michel A. Hofman and Dean Falk (eds.) *Progress in Brain Research* 195: 443-459. Amsterdam: Elsevier.
- Schütze, Carson 1999. English expletive constructions are not infected. *Linguistic Inquiry* 30(3): 467–484.
- Siewierska, Anna. 1996. Word order type and alignment type. *STUF - Language Typology and Universals* 49(2): 149-176.
- Siewierska, Anna / Uhlířová, Ludmila 2010. An overview of word order in Slavic languages. In: Siewierska, Anna (ed.) *Constituent order in the languages of Europe*. 105-150. Berlin: Mouton DeGruyter.
- Simon, Herbert 1996. *The sciences of the artificial*. Cambridge, Mass.: MIT Press.
- Skipper, Robert A. Jr. and Dietrich, Michael R. 2012. Sewall Wright's adaptive landscape: philosophical reflections on heuristic value. In: Svensson, Erik & Ryan Calsbeek (ed.) *The adaptive landscape in evolutionary biology*. 17-25. Oxford: Oxford University Press.
- Sobin, Nicholas. 1997. Agreement, default rules, and grammatical viruses. *Linguistic Inquiry* 28(2): 318–343.
- Solan, Zach, Eytan Ruppín, David Horn and Shimon Edelman. 2005 Evolution of language diversity: why fitness counts. In: Tallerman, Maggie (ed.) *Language origins: perspective on evolution*. 357-371. Oxford University Press.
- Spielman, Stephanie J., Steven Weaver, Stephen D. Shank, Brittany Rife Magalis, Michael Li, and Sergei L. Kosakovsky Pond 2019. Evolution of viral genomes: Interplay between selection, recombination, and other forces. 427-468. In Maria Anisimova (ed.) *Evolutionary Genomics: Statistical and Computational Methods* (Methods in Molecular Biology, vol. 1910). New York: Humana.
- Stringer, Chris, Galway-Witham, Julia 2017. On the origin of our species. *Nature* 546: 212–214. <https://doi.org/10.1038/546212a>
- Stromswold, K. 1990. *Learnability and the acquisition of auxiliaries*. Unpubl. PhD dissertation. Cambridge, MA: Massachusetts Institute of Technology.

- Sundquist, John D. 2012. Negative movement in the history of Norwegian: the evolution of a grammatical virus. In Dianne Jonas and John Whitman (eds.) *Grammatical change: origins, nature, outcomes*. 293-312. Oxford: Oxford University Press.
- Szyf, Moshe. 2014. Lamarck revisited: epigenetic inheritance of ancestral odor fear conditioning. *Nature Neuroscience* 17, 2–4. <https://doi.org/10.1038/nn.3603>
- Theißen, Günter 2009. Saltational evolution: hopeful monsters are here to stay. *Theory in Biosciences*. 128: 43–51.)
- Thomason, Sarah G. and Terrence Kaufman 1988. *Language contact, creolization and genetic linguistics*. Berkeley: University of California Press.
- Ullman, Michael T. 2001. A neurocognitive perspective on language: The declarative/procedural model. *Nature Reviews Neuroscience* 2: 717–726.
- Voss, Patrice, Thomas M.E., Cisneros-Franco J.M. and de Villers-Sidani É. 2017. Dynamic brains and the changing rules of neuroplasticity: Implications for learning and recovery. *Frontiers in Psychology* 8: 1657. doi: 10.3389/fpsyg.2017.01657.
- Wahlström, Max 2015. *Loss of case inflection in Bulgarian and Macedonian*. Doctoral Diss. Helsinki: University of Helsinki, Department of Modern Languages.
- Winford, Donald 2003. *An Introduction to Contact Linguistics*. Oxford: Blackwell.
- Witt, Ulrich 2006. Evolutionary concepts in economics and biology. *Journal of Evolutionary Economics* 16(5): 473-476.
- Wright, Sewall 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proceedings of the 6th Int. Congress on Genetics* 1: 356–366.