Syntactic Change: The Word-order Cycle*

Paul Kiparsky

*I am grateful to the audiences at ICHL and several earlier venues for their questions and criticisms. There is no space here to address all the points they have raised, but at least I can offer this first instalment. Thanks also to Cleo Condoravdi for discussing the ideas and presentation with me, to Jadranka Gvozdanović for her patience and support, and to Joe Salmons for his reading of the final MS.

Abstract

Long-term changes, and their independent recurrence in different languages and families, are an unsolved problem for historical linguistics. They are hard to reconcile with the standard view that morphological and syntactic changes are due to error-driven reanalysis by language learners. I defend instead a *bias-driven* model based on Optimality Theory, which posits that the course of acquisition is shaped by learners' construction of successively more accurate fully ranked constraint systems.

Suppose the probability of a language is measured by the proportion of fully ranked constraint systems that generate it, its *r-volume* (Riggle 2010), and that learners continuously revise their evolving grammars to accommodate new data they encounter. The likelihood of a language being projected by a learner will then be proportional to its *r*-volume. The learning process is thus intrinsically biased in favor of the most probable language that is consistent with the learner's experience. The hypothesis is that this is the mechanism behind drift. Here we investigate word-order change from this perspective.

Languages with consistently head-initial phrase structure tend to have a richly branching spine of functional projections, whereas languages with head-final phrase structure tend to have relatively flat syntax. Over time, languages can change from either of these types to the other. I present a simple OT constraint system that derives the core categorical universals of clausal word order, including the preference for Subject-Object order over Object-Subject order, and an improved version of the Final-Over-Final-Constraint, and then show that this constraint system also predicts the observed correlation between basic word-order type and the repertoire of syntactic functional heads. Coupled with the assumption of bias-driven language acquisition, this yields an explanation of word order drift. The well-established processes by which function words are morphologized and lost, and then replaced again with new syntactic heads by grammaticalization, can be seen to drive a syntactic word order cycle.

1 Drift, convergence, and unidirectionality

1.1 The problem of drift

A prominent challenge to historical linguistics is the phenomenon of drift – sustained unidirectional change that takes place in small steps over a long period, sometimes even millennia. Adding to the puzzle is the fact that the same long-term developments occur independently in different languages and even across families, sometimes involving large-scale typological shifts, in what Meillet (1918) called *convergence*, the counterpart of *homoplasy* in evolutionary biology.

Drift sits uncomfortably in the theoretical landscape. It is inconsistent with the widespread view that language change is a random walk in the space of possible grammars (Battye and Roberts 1995:111, Jäger 2019, 2025). Its gradual character is also at odds with the idea of abrupt catastrophic change proposed by Lightfoot and others,[1] and with cue-based and parameter-resetting models. Though drift is well-documented, there are few efforts to theorize it. Even Kroch and his collaborators, who did much to document the long-term course of change and to establish its characteristic trajectory, did not really tackle the question of what actually drives such change.

When drift is countenanced at all, it tends to be presented as a somewhat mysterious mechanism with a teleological or even occult vibe. Sapir (1921: 155) imagined "the unconscious selection on the part of its speakers of those individual variations that are cumulative in some special direction". In what is probably the first discussion of drift from a generative perspective, R. Lakoff (1972) characterized it as change governed by a "metacondition" — a constraint on change which is neither part of a grammar nor a universal condition on grammars, "some principle, outside our present concept of grammar, governing the forms the rules are to take." Hale 1998 speaks of the Slavic open syllable shift as a language's realization of its "desires": "How can a "language" (if we even wanted to admit the relevance of such a concept into our considerations) that does not have a restriction against closed syllables (e.g., pre-Proto-Slavic) acquire a compulsion to develop such a restriction, a compulsion that achieves its desires hundreds of years later only to almost immediately surrender it (and develop closed syllables again)? Where would such a conspiracy reside and how would it exercise its influence on grammar construction by the relevant generations?" And Lightfoot (1999: 209; 2003: 104) even invokes "racial memory", asserting that the protracted SOV > SVO change in English "raises the question of how a child attains a language which is exactly half-way between the subject-verb-object and the subject-object-verb types; how does she know whether this is a subject-verb-object language changing into a subject-object-verb, or vice versa? How does she know that her generation must push the language a little towards, say, a subject-verb-object type? It seems that the only conceivable answer is to posit a racial memory of some kind, such that the child knows that this is a subject-verb-object language changing towards subject-object-verb." These authors imply that the phenomenon itself is dubious because it seemingly requires a teleological approach to change that posits vitalistic powers and impulses residing in language itself.

But Hale's allusion to "conspiracies" is unintentionally right on target. That term was coined to highlight the problematic character of another phenomenon that the linguistic theory of the time could not explain: that certain phonological patterns emerge from the interaction of disparate and apparently unrelated processes. In the classic article that first drew attention to the problem, Kisseberth (1970) showed that Yowlumne Yokuts phonology has several deletion, epenthesis, and vowel shortening rules that collaborate to enforce a simple rigid syllable canon. He noted that the rule-based phonology of the time was unable to bring out the functional unity of these processes, and

[1] For example, that modal auxiliaries originated by a sudden reanalysis of verbs in about 1575.

introduced the new device of output constraints, which restrict the operation of rules and thereby allow some of the rules that participate in the conspiracy to be simplified. Output constraints were a precursor of Optimality Theory (OT), which abandoned rules altogether in favor of ranked violable constraints. An OT constraint can be satisfied in multiple ways due to its interaction with other ranked constraints within the phonological system. This resolved the problem of conspiracies.

I propose that drift, which seems just as intractable in traditional frameworks, can likewise be demystified by the shift from rules to constraints. A OT-based theory of change can make sense of convergent long-term change, in particular of drift in basic word order and its relation to functional structure, which are among the most basic, widespread, and best-documented instantiations of the phenomenon.

1.2 Indo-European

On the evidence of its early daughter languages, proto-Indo-European had a uniformly head-final basic word order, with displacement of topics and focused constituents to left-peripheral positions, and "scrambling" within the clause sensitive to definiteness and givenness. It had a rich inflectional morphology and a small repertoire of function words: no auxiliaries, no articles, no *that*-type complementizers, a fairly flat syntactic constituent structure, head-final syntax (SOV), an enclitic coordinating conjunction *=k^we (Goldstein 2019), and pronominal and adverbial clitics at the left periphery (Hale 1987). There is no evidence of an InflP or TnsP. Finite subordinate clauses were arguably adjoined rather than embedded, and limited to relative clauses of the relative-correlative type, and perhaps also quotatives. Otherwise propositional subordination was done by a rich system of nominalizations, which quite early began to become grammaticalized into infinitives and participles.

Each branch acquired new function words, including entire new word classes such as articles, prepositions, propositional (*that*-type) complementizers, and auxiliaries, beginning with TAM auxiliaries for periphrastic constructions. Almost all these new function words governed to the right, to form left-headed Determiner Phrases (DPs), Prepositional Phrases (PPs), Complementizer Phrases (CPs), and Tense Phrases (TPs). In most branches of Indo-European, Verb Phrases (VPs) then switched from OV to VO to establish a consistent head-initial syntax. Continental West Germanic (German, Dutch, Flemish, Frisian, and Afrikaans) did not fully participate in this final phase, and retain basic OV, with a residual layer of postpositions. Indo-Aryan languages remain a holdout of head-final syntax, except for the initial phase of the drift, the acquisition of initial *that*-type complementizers.[2]

The shift from head-finality to head-initiality is the cumulative result of many distinct steps, each of which itself unfolded over centuries. The steps are ordered from top to bottom in the constituent structure. Most of the trajectory can be tracked in the abundant and chronologically deep historical record of Romance, Germanic, and Greek.

1.3 The OV > VO drift

Nearly half the world's languages are SOV. Almost as many are SVO. About 8% are VSO. The SOV order broadly correlates with head-finality in other categories, and VO (SVO, VSO) broadly correlates with head-initiality in other categories (Dryer 2013), where a head is defined as an obligatory single constituent (X°) that selects a complement and determines the category of the construction that it forms with it. For example, the complementizer *that* is of category C, selects

[2]Bayer (1999) shows, conclusively to my mind, that final complementizers in Indo-Aryan languages are generated in final position, and the clauses they head are not fronted across them, contra Kayne 1994.

a finite clause, and combines with it into a CP (Complementizer Phrase), turning propositions into arguments.

OV order tends to go with other HEAD-FINAL properties: postpositions (rather than prepositions), nouns follow genitive and relative clause modifiers, auxiliaries follow the main verb, copulas follow their predicates, question particles follow their clauses, and in some languages even conjunctions (complementizers) follow their clauses, which is never the case in VO languages.

It is commonly assumed that the change from OV to VO is more common than the change of VO to OV, and more "natural" (Maurits & Griffiths 2014). It has even been proposed that OV > VO is universally unidirectional (Newmeyer 2000, Gell-Mann & Ruhlen 2011, Haider 2021, 2023; I mistakenly did so myself in Kiparsky 1996). From this it would follow that all languages were originally OV. And Gell-Mann & Ruhlen indeed go so far as to claim that all languages are descended from a single OV "proto-world" language.

But *why* would the primordial order have been SOV, and *why* would all subsequent changes have gone unidirectionally towards SVO? Neither of these orders has been shown to have any *overall* processing advantage over the other (Hawkins 1994, Engelhardt, Filipović & Hawkins 2024). And even if one order was superior to the other, that could not explain *both* a putative original SOV word order *and* a putative unidirectional contrary shift to SVO. They may well have *different* functional advantages, say learnability versus parsing efficiency, or dependency locality versus information locality, or perception versus production (Hall, Mayberry & Ferreira 2013, Hahn & Xu 2022). But why would one of these advantages have been all-important when language originated, whereas the other became dominant later?

Another difficulty is that unidirectional OV $\rightarrow$ VO predicts eventual universal convergence on VO word order. Even the weaker version of the unidirectionality claim, which restricts it to endogenous changes, and concedes that VO to OV changes can result of borrowing (Li 1977: xii–xiii) entails this prediction. But the prediction is certainly false. Historical comparative linguistics shows that word order change is common, and that the fundamental typological features of a language, including basic word order, can change within a few thousand years. Language is thought to have originated at least 135,000 years ago. So, unidirectional OV $\rightarrow$ VO change at anything like the known rate of change would have made most of the world's languages VO by now. But in reality half of them are OV.

The hypothesis that OV > VO is universally unidirectional is in any case untenable for the simple empirical reason that shifts in the reverse direction VO > OV do happen. Some of the known cases are reviewed below. I will argue that in the long term, word order shifts are equally frequent in both directions, due to a perennial ... OV $\rightarrow$ VO $\rightarrow$ OV... cycle that causally interacts with shifts in the syntactic spine of functional heads.

1.4 Is drift a mirage?

Since drift is theoretically problematic, we have to either explain it away or revise our understanding of language change. One way to explain away drift is to claim that the textual record is misleading. Pointing to the apparently rapid shift from primarily SOV word order to fixed SVO in late 12th century English, Lightfoot (1979) conjectured that drift is merely apparent, and that all change is actually abrupt. But on closer examination it turns out to be the discontinuity that is a mirage (Allen 1995, Ch. 5). The apparent stability of Old English in the monastic texts from the eighth to the twelfth century reflects a standard written form of Old English that was maintained throughout the literature of that period, and reinforced by the copying and recopying of the

manuscripts. The spoken vernacular language was changing substantially throughout this time, as revealed by errors in texts from the final stages of that scriptorial tradition (most dramatically towards the end of the Peterborough Chronicle). After a near-hiatus in English-language literary activity caused by the Norman Conquest, the thirteenth century saw a reset with new works written in the regional dialects of the time. It may look as though the language suddenly leapt from SOV to SVO word order and concurrently reduced its inflectional morphology, but on the ground, the syntax and morphology evolved steadily in small steps, as always.

1.5 Is drift the result of chance?

Could drift be simply due to chance? Even the proverbial blind ant in a labyrinth can get lucky sometimes. In Lightfoot's (2003) words, "What's wrong with a series of coincidences?" In the same vein, Keiser (2009: 29) conjectures that convergence is "nothing more than ordinary, non-end-driven, internally-induced language change that, through a combination of universals and chance occasionally results in parallel developments in related languages".

But there is no reasonable way the parallel long-term unidirectional word order shifts from Latin to Romance, from Old English to Middle English, and from Homeric to Modern Greek, with no backtracking or false steps,[3] can be due to a chance series of unconnected changes over several millennia, or for that matter explained by error-driven reanalysis or some occult teleology. When all blind ants head the same way, a rational biologist credits them with a sense of direction and looks for the underlying mechanism. The same goes for directionality in linguistic change.

1.6 Is drift supervescence?

Convergence and drift within a language family can sometimes be explained as the emergence of an incipient or latent feature of the proto-language. Sanskrit did not have ergative case, but it developed everywhere in Middle Indic. So it is at least possible that the actual ancestor of the Middle Indic dialects was a vernacular Prakrit that had already developed ergative case, contemporaneous with normative literary Sanskrit, and diglossic with it. In Latin, too, colloquial registers were consistently ahead of formal literary registers, particularly in the case of Verb-Object order (Ledgeway 2012: 202-35). Some apparent drift phenomena, therefore, might just ne the "bubbling up" in the daughter languages of inherited features that were originally submerged for sociolinguistic reasons ("supervescence", Joseph 2006, 2013, Janda and Joseph, in press). However, supervescence is unlikely to be the origin of the major across-the-board typological changes under discussion here. There is no evidence of any inherited *that*-complementizers, articles, auxiliaries, or prepositions that might have been preserved from some vernacular register of Proto-Indo-European down into the daughter languages. These things clearly arise in the separate branches by independent grammaticalization processes, which is why they govern to the right like all new functional categories, even in otherwise still predominantly head-final languages. And inherited head-finality can be seen to give way to head-initiality by a long series of innovations within each individual branch of Indo-European, not by a sudden surfacing of previously stigmatized usage.

1.7 Is drift simplification?

If drift is neither chance, nor a textual mirage, nor supervescence, what is it? Some instances of drift have been portrayed as successive simplifications. According to Bauer (1995) Latin evolved

[3]Proviso: deliberate backtracking can of course result from reimposition of an older prestige standard, as in Greek Atticism of the 1st century BC, or from contact with a prestige language, as in 15th century Swedish, where OV order, after being nearly completely lost, reappears due to intensive cultural, commercial, and political relations with Low German, and then recedes again to complete the drift (Petzell 2011, 2016, Sangfelt 2019).

into the Romance languages because it was "a difficult language to master". Actually SVO word order (or left-headed syntax in general) does not seem to have any overall acquisition or processing advantage (for references, see section 1.3). As for formal complexity, languages certainly differ in word structure, whether in the richness of morphological categories, the number of morphemes in the inventory, the average number of morphemes in a word, the amount of allomorphy, the number of exceptions, or by any other measure. They also differ in their phonological complexity, and probably in their syntactic complexity. But there is no known operationally feasible way to compare the overall complexity of entire languages. Minimum description length in a fixed notation provided by the theory of grammar, which translates simplicity into generality (Chomsky & Halle 1968), is theory-dependent and leaves open the question what to measure. On the (contested) assumption that the lexicon is the only locus of language-specific information (Borer 2005), it is the size of the lexicon. Or it could be the number of formal features that are specified in the grammar (Roberts 2007: 235), or the number of marked parameter settings (Roberts 2007: 21). A theory-independent definition of complexity, unimplementable and untestable as things now stand, is minimum description length, in bits, in any universal notation, known as Kolmogorov Complexity (Goldsmith 2011). An approximation to it would be to define the complexity of a language as the size of its complete explicit grammar and lexicon, including the rules (or constraints) of phonology, morphology, syntax, and semantics, and the exceptions to them, encoded in some sufficiently well-defined format. Unfortunately a complete explicit formal description has never been achieved for any language, and even approached only once, and that was two and a half thousand years ago. On such a global measure of complexity, regardless of how exactly it is implemented, there would probably not be much difference between Latin and French, or between Old English and modern English. In any case, thes languages would most likely all score somewhere in the middle range, presumably simpler than Navajo, and more complex than Tok Pisin.[4]

Any theory of drift must reconcile it with the existence of contrary changes and continuing typological diversity. If we claim that OV > VO changes and their typological concomitants are driven by unidirectional processes, we owe an account of the existence of the opposite changes, and of the fact that OV-type languages are so common. The proposal to be advanced here does not posit a single unidirectional trajectory, but a long-term oscillation, or cycle, and it spells out their opposing phases and what triggers them. It predicts that the word order types and the changes that produce them are equally frequent in the long term.

1.8 Is word order change triggered by morphology?

Two centuries years ago, the first linguistic typologists classified languages into three types according to the complexity of their words and the degree of fusion of their parts: analytic, synthetic, and isolating. A fourth, polysynthetic type was soon added. From there it was a short step to envision a long-term morphological cycle, during which analytic morphology first fuses into synthetic morphology and eventually becomes isolating, and the resulting monomorphemic words then combine with each other to reintroduce analytic morphology (Hodge 1970, Dixon 1994:182–184).

Sapir (1921) later tidied up the typology by distinguishing the complexity of words from the degree of fusion as distinct typological parameters. The complexity of words is more likely to be correlated with syntactic typology than the degree of fusion, for syntax cares about the presence of case inflection and verb agreement morphology but not about how intimately the morphemes bond with their stems.

[4]On complexity, see McWhorter 2007, Andrason et al. 2023, Dahl 2004, Miestamo et al. 2008, Sampson, Gil & Trudgill 2009, Miller 2012, Baerman et al. 2015, Arkadiev & Gardani 2020, Catasso, Bastiani & Coniglio 2022, Engelhardt, Filipović & Hawkins 2024, Polinsky, Putnam & Salmons (forthcoming), and in general Mitchell 2009.

The exploration of morphology/syntax connections has continued in minimalist syntax with efforts to correlate verb movement (V-to-Tense raising) with the richness of verb agreement (Biberauer & Roberts 2009, Koeneman & Zeijlstra 2014), which so far remain inconclusive.

A much older hypothesis posits an organic connection of rich case and/or agreement inflection with free clause-internal word order ("scrambling"), as opposed to topicalization or focalization to clause edges. The intuition is that inflectional morphology and fixed word order of arguments are alternative means of distinguishing subjects from objects, and that one of these means is necessary, but enlisting both violates a preference for economy. There are certainly striking examples of this co-variation. Aleut and Dutch, which have impoverished case systems, distinguish subject and object by rigid SOV word order,[5] while their respective relatives, Eskimo and German, retain much freer word order, with case marking doing the work of distinguishing nominal subjects from objects. The generalization is however too strong in one respect and too weak in another. Rich inflectional morphology does not *entail* free word order: some richly inflected languages have fairly rigid word order, for example, Icelandic, Grisons Swiss German, and modern Greek. And rigidity of word order is not a simple typological parameter: many richly inflected languages are strictly verb-final, yet tolerate "scrambling" of direct arguments within the core clause, in addition to movement to left-peripheral positions. A narrower generalization that I believe is valid is that no language which lacks inflectional morphology allows scrambling, although even this weaker connection between linearization and morphology is not predicted by current syntactic theory, with the possible exception of licensing theory (Galbraith 2023).

For obvious functional reasons, strict SVO languages can do without morphological case marking (Vennemann 1975, Hawkins 2004, Siewierska & Bakker 2009). Indeed, the convergent shift from head-finality to head-initiality in most branches of Indo-European and in the Western Uralic languages is broadly correlated with a reduction of inflectional morphology. But this correlation too is imperfect. Many of these languages, among them Greek in Indo-European, and Finnish in Uralic, have changed from basic SOV to SVO while retaining distinct grammatical cases and subject-verb agreeement. Conversely, Dutch is fairly rigidly SOV, yet inflectionally poor. Similar examples can be cited in other language families. The Ịjọ languages are SOV, yet inflectionally poor (Williamson 1965, Worufah & Obikudo 2023). All dialects of Chinese lack inflectional morphology, but the Southern ones are SOV. The fact is that both fixed SVO word order and fixed SOV word order can occur independently of morphological richness. The attrition of the case system therefore does not suffice to explain the change of SOV word order to SVO word order (*pace* Salvi 2004). Textual evidence also makes a direct connection doubtful: SVO order began to spread in late Latin and Old English *before* the distinction between nominative and accusative case was lost (Ledgeway 2012: 29, sections 3.3.2, 7.3.2.2, Pintzuk & Taylor 2006).

To the extent that there is a genuine causal relation here, it is arguably indirect and mediated by syntax: left-headed functional categories (such as auxiliaries) induce VO order and are associated with meager inflectional morphology. I propose to explain this correlation in the following way. Function words in the syntax are functional counterparts of inflectional morphology, so languages with sparse morphology are rich in function words. These preferentially precede their complements, as explained below. The left-headed order of functional categories propagates downward to lexical categories by a syntactic harmony principle, eventually inducing consistent VO order.

There is a strong cross-linguistic correlation between syntactic left-headed TPs and SVO order (Dryer 1992: 100). Almost all languages in which auxiliaries precede the VP have VO order. A

[5] Apart from movement to designated topic and focus positions; for Aleut, see Bergsland 1997: 343.

rare exception is Canela-Krahô (Popjes & Popjes 1986). It has Tense/Aspect auxiliaries to the left of verb-final VPs:[6]

(1) wa ha pïχó jühkà
1P FUT fruit buy
'I will buy fruit.' (Popjes & Popjes 129)

This combination is not an outright exception to the word order constraints formulated below in (3), but it is an instance of the marked structure that I refer to as benign disharmony.

2 Reanalysis and drift

2.1 Acquisition and change

If unidirectional long-term change and convergence by parallel developments in unrelated languages are indubitably real phenomena in language change, we have to identify the assumptions that make them problematic and revise them. The stumbling block, I think, is the idea that the major cause of change is error-driven reanalysis in the course of language acquisition. The notion of error-driven reanalysis has become part of a larger conception of linguistic change, and specifically of syntactic change. It fits hand in glove with the doctrine of change as a random walk in typological space, and it is often paired with the idea that change does not proceed in a sequence of smaller but discrete increments, but in large abrupt steps ("catastrophes") (Lightfoot 1999). This doctrine probably owes its popularity in part to its resonance with generative grammar's idealized model of instantaneous language acquisition within a homogeneous speech community, where learners acquire the simplest grammars compatible with UG and the ambient language – a useful simplifying abstraction for the formal theory of grammar, but not so innocuous for linguistic change. The actual course of language acquisition is undoubtedly shaped by the realities of language variation and contact, and by learners' substantive preferences of the kind that markedness and OT constraints aim to encode.

Reanalysis is commonly held to be a two-stage process consisting of an initial covert phase, and a subsequent overt phase in which the change is actualized (Anttila 1972, Langacker 1977, Timberlake 1977, Harris & Campbell 1995, Roberts 2007, Trask 2015). In the covert ('abductive') phase, learners parse certain structures in a new way; their I-language thereby ends up differing from that of the speakers from whom they acquire it, although there is no detectable change in the output at this stage. In the overt ('deductive') phase, the new grammar is then extended ('actualized') to contexts where it is manifested as innovative usage. In line with current usage, I refer to both stages together as 'reanalysis'.[7]

[6] Kashmiri has large number of periphrastic auxiliaries, and the verb moves to second place if there is no auxiliary or negation/affirmation marker there (Hook & Koul 1996, n.d.a, n.d.b, Koul & Wali 2006, examples from pp. 78, 127, 153). It looks like Kashmiri has a PolP/ΣP projection (on these, see below), but no separate TP. Karitiana (Tupí-Arikém, Brazil) has obligatory verb fronting (V2, sometimes V1) in main clauses, and V-final subordinate clauses (Storto 1999: 120). Fronted verbs have tense and agreement morphology, while final verbs do not. Aspectual morphology is present in both positions. Storto's analysis posits basic Subject-Aux-Object-Verb with verb movement to the Tense/Agr auxiliary. Shlonsky (1997) also proposes to analyse Modern Hebrew (Afro-Asiatic Semitic) as V2. V2 is noted, but less well documented, in Papago (Pimic group of the Uto-Aztecan, see Bhatt 1999:epilogue and references therein).

[7] A note on terminology: strictly speaking, 'reanalysis' proper is the first, covert stage, "change in the structure of an expression or class of expressions that does not involve any immediate or intrinsic modification of its surface manifestation." The overt changes that it may lead to, "the natural and expected result of functionally prior modifications in

Error-driven reanalysis, the catastrophic nature of change, and its character as a random walk have all been challenged on both empirical and theoretical grounds. In the next sections I summarize and extend these critiques, and argue for an alternative, *bias-driven* learning model, in which learners' expectations play a role in determining the output of the acquisition process, and which predicts the existence of long-term change and convergence (homoplasy). I formalize it in OT, whose mathematical simplicity affords a particularly clear characterization of the biases that drive the direction of change. I support it here by evidence from the empirical typology of word order and word order change, which I argue exhibits a major long-term cycle.

2.2 Error-driven reanalysis

The error-driven reanalysis scenario is puzzling in several respects. Why would learners converge on one grammar before the change, and then, on the basis of the very same Primary Linguistic Data (PLD), converge on *another* grammar in sufficient numbers to initiate the covert stage of the change? What happens at the transition from the covert phase to the overt phase? Do the innovations appear as soon as the new grammar is applied in contexts that reveal them, or does it, as some researchers maintain, involve further changes, such as either analogical extension of the new grammar (Anttila 1972, Timberlake 1977, Hopper & Traugott 2003, Madariaga 2017),[8] or 'adaptive rules' which retrofit the new grammar to the old linguistic givens, and are then themselves eliminated to reveal the effects of the innovation (Andersen 1993)? How and why is the new grammar internalized and replicated by speakers who do not themselves go through this reanalysis process? Why does the language not oscillate back and forth between the two grammars? In short, error-driven reanalysis leaves open two core questions: how a language changes from one state to a succeeding state, and what triggers the change – the *transition problem* and the *actuation problem* (Weinreich, Labov & Herzog 1968). Indeed, it complicates them by positing two empirically indistinguishable phases for each change event (Haspelmath 1998).

Another reason to doubt that major syntactic changes are the result of simple mislearning is the sociolinguistic finding that learners are attentive to any discrepancy between their own usage and that of others because language is a marker of social identity. Moreover, change can occur even when the ambiguous structures that invite multiple analyses are very infrequent (De Smet & Markey 2021), and it can involve salient features which are known to be acquired early, such as basic word order.

A subtler variant of the error-driven change hypothesis posits frequentistic E-language precursors of I-language reanalysis (Lightfoot 1991:67-68, 1999, Clark & Roberts 1993, Hale 1998). In line with Saussure's view that change originates in language use *(parole)*, accidental fluctuations in the frequency of primary linguistic data are assumed to affect the result of language acquisition by changing learners' triggering experience. When the frequency of the data supporting a feature of grammar for whatever reason drops below some critical threshold, learners can no longer take it in, and they acquire a new grammar which generates a different output.

Positing curtailment of evidence in the ambient data as the cause of change would predict that change is a unidirectional simplificatory process that reduces structures or eliminates arbitrary restrictions from grammars. Certainly many changes have that effect, but many others create new

rules and underlying representation," are 'actualizations' (Langacker 1977: 58) or 'extensions'. The covert and overt phases of change correspond to what Andersen (1973) calls abductive and deductive change, see however Deutscher (2002)'s argument that this terminology is based on an incorrect understanding of Pierce's philosophical work.

[8] "What is couched in philosophical terms as 'abductive change' can simply be called 'reanalysis'; what is called 'deductive change' is just '(analogical) extension'." (Deutscher 2002).

structure, in English for example the progressive, the *going to* future, reflexive pronouns, articles, and the auxiliary system, and in general the large class of changes subsumed under the term grammaticalization. it is not clear how new categories could arise just through the mislearning of old ones, at a stage when there is evidence for the existing grammar and no evidence whatsoever for the new grammar. And mere variation in usage could hardly have caused learners to mislearn features of core grammar, such as basic OV word order as VO. Studies of L1 acquisition shows that children have word order well under control quite early, even before they master inflectional morphology.

Moreover, the fluctuations in usage that the frequentistic version of the reanalysis model posits as triggers of reanalysis cannot be verified independently of the changes that they are supposed to be the causes of. In fact, there is evidence that changes in frequency are consequences of grammatical change, rather than causes of them. This is shown by the systematic frequency profile of variation (or "grammar competition") known as the Constant Rate Effect (Kroch 19892), and particularly by the discovery that variation is governed by the same grammar-based constraints that underlie categorical regularities of I-language (Anttila 1997, 2007). So locating the cause of reanalysis in frequentistic E-language variation does not offer a solution to the actuation and transition problems – it only covers them up by reversing the causal connection.

The hypothesis that change is caused by random or extralinguistically motivated fluctuation of usage frequency is also inconsistent with evidence that low-frequency constructions can be acquired and remain stable for long periods.

The hypothesis is in any case not only implausible but unnecessary. If some portion of primary linguistic data becomes so infrequent that learners no longer take it on board, the reduced data that remains accessible is necessarily still compatible with the old grammar. If learners at that point innovate, the new grammar must have some kind of advantage. That is where any explanatory content of the account would lie. The bias-driven model leverages that advantage to explain the innovation without relying on hypothetical extralinguistically triggered fluctuations in usage.

Similar criticisms apply to other versions of the reanalysis account, including cue-based models (Lightfoot 1999), where first-language learners do not evaluate candidate grammars against sets of sentences: they look for bits of diagnostic structure that unambiguously reveal parameter settings. The cues are a kind of universal checklist assumed to be available to the learner prior to language acquisition. Checking all the cues picks out the target language from the set of possible languages provided by UG. Since the acquired structures can be fairly abstract, with empty categories and displacement of constituents, the question arises how a learner can discover whether a given piece of data constitutes an unambiguous cue for a given parameter setting. An attempt to overcome this problem was the decomposition of cues into more specific microcues, each corresponding to a microparameter – for example, separate cues for V2 in each clause type (Lightfoot & Westergaard 2007, Lightfoot 2010). The inventory of microcues is effectively a list of "constructions", which would make this approach a variant of construction grammar or structuralism.

The conclusion is that error-driven reanalysis has no solution to the transition and actuation problems – how and why learners replace a grammar with a new one when the available data favors the existing one, or at least underdetermines the choice between them.

2.3 Empirical objections to error-driven reanalysis

Error-driven reanalysis has been challenged from the empirical side as well. Recent more fine-grained work on syntactic change questions the idea that syntactic reanalysis occurs by reseg-

mentation of constituent structure (Allen 1995, Garrett 2012, Whitman 2012, Weiß 2019). Let us review some familiar standard examples.

The origin of Nominative experiencer subjects has been attributed to reanalysis of the surface dative object as a nominative subject in sentences like *the woman liked those words* (Stockwell 1976, Fischer 1987, Harris & Campbell 1995: 63, 83, among others). Instead, this change appears to be a lexically gradual reduction of lexical case assignment to Experiencer arguments (Allen 1995, Ch. 7).

English *for*-infinitivals were thought to have arisen by the rebracketing of prepositional phrases, for example, "It was [good for him] to go" > "It was good [for him to go]" (Harris & Campbell 1995: 62, following Jespersen and other earlier scholars). This idea was questioned by de Smet 2009 and Garrett 2012, and their objections still stand. The change is probably connected to the late Middle English innovation by which IP (or TP) becomes an obligatory syntactic projection (Kiparsky 1997). At that point, *to* joins the modals as a nonfinite Infl head (van Gelderen 1993). Being nonfinite, an IP whose Infl is *to* must form a subordinate clause. The complementizer that heads such a subordinate clause is *for*. Thus the rise of *for-to* infinitives is part of — in fact, *caused by* — a larger syntactic shift.

Since neogrammarian times, *that*-clauses of Germanic and other language families have been said to arise by reanalysis of paratactic main clauses, e.g. **He said that: "They came"* > *He said that they came*. Their source is almost certainly a correlative construction, by a development of the following kind: "I know [$_{DP}$ it [that$_{Wh}$ he left]]" $\xrightarrow{pro-drop}$ "I know [$_{DP}$ $\emptyset$ [that$_{Wh}$ he left]]" $\xrightarrow{reanalysis}$ "I know [$_{CP}$ that [he left]]". In the light of recent work on grammaticalization it can be recognized as an instance of Spec-to-Head restructuring, motivated by Head Preference (van Gelderen 2009). *That*-type complementizers arose from relative pronouns in each branch of Indo-European, e.g. Germanic *that, dass* (Axel-Tober 2017), Latin *quod*, Greek *ho-ti, pou*, and later within the branches of several of its daughters, e.g. Sanskrit *yad*, Bangla *je*, Hindi *ki*; see further Weiss (2020).

The change of SOV word order to SVO order in the non-continental Germanic languages (English and Scandinavian) and in Bantu has been attributed to the rebracketing of extraposed objects "afterthoughts") from outside the VP as part of the VP (Hyman 1975). The hypothesis is that the output of the synchronic derivation (2a) is reanalyzed as in (2b).

(2) a. Extraposition: [Subject [Object Verb]$_{VP}$]$_{CP}$ → [Subject [Verb *t*]$_{VP}$ Object]$_{CP}$

b. Reanalysis: Subject [Verb *t*]$_{VP}$ Object]$_{CP}$ > [Subject [Verb Object]$_{VP}$]$_{CP}$

This rebracketing account of OV → VO shift has been refuted (Whitman 2012); for more recent perspectives below see section (5.2) on Germanic, and section (5.6) on Niger-Congo.

There is no question that a change can be caused by prior changes which make the old analysis unlearnable, or which make available a better one. Obviously phonological change can erode morphology and ultimately wipe it out or force its drastic renewal. In the same way, syntactic change could be the result of morphological change, particularly of the realization of inflectional categories like case and agreement, and of argument-structure changing morphology such as passive, causative, and antipassive. Such cases are not error-driven reanalysis in sense of the preceding section. A traditional term for them (at least in phonology) is *restructuring*, which usefully distinguishes it from reanalysis. Because restructuring is a direct consequence of the modified triggering experience, no two-phase process of change is required for restructuring, and it would

not even make sense to assume a covert abductive phase. For example, if Allen (1995, Ch. 8, 9, 10.4) is right that the loss of the 'indirect passive' construction and the rise of the recipient passive in Middle English are consequences of the collapse of the accusative/dative case distinction and the fixing of word order, the new grammar is is a consequence of the new givens of the language, and we have a genuinely explanatory account of the change. In the same way, Faarlund (2021) shows how a word order change can force restructuring of a whole series of constructions.

3 Bias-driven change

3.1 Acquisition and change

To explain the initiation, spread, and completion of drift and grammaticalization, I propose a *bias-driven* model that attributes change to learners' prior biases – preferences in a non-pejorative, non-teleological sense – rather than simply to passive learning failure. The biases guide the acquisition process and thereby indirectly propel the language on an evolutionary path. Change occurs when hypotheses embodied in intermediate grammars and the outputs generated by those grammars are retained and become part of the linguistic norm. Languages will tend to change towards the structures privileged by learners' expectations. Drift, then, is a optimizing path through a FITNESS LANDSCAPE (lets 2004) defined by UG.

This bias-based approach is compatible with the postulate that syntactic reanalysis passes through a covert stage, but it does not depend on it. If we assume that prior shared biases can override learners' experience (Garrett 2012, Culbertson, Smolensky & Legendre 2012, Garrett & Johnson 2013), they may cause learners to converge on "wrong" innovative grammars, which may spread when their advantages over the old grammar cause them to be adopted. It would follow that changes can be initiated and persist in spite of accessible counterevidence, and that they can involve not just loss of old features and categories, but the creation by grammaticalization of new ones that have no pre-existing model in the language. The key point is that learners' expectations provide a causal mechanism for the initiation, spread, and completion of change across environments over many generations, which in the long term can result in drift (long-term unidirectional change) and convergence (independent parallel change different languages). In short, instead of Saussurean change-in-*parole*, bias-driven models adopt the Jakobsonian conception of change as an evolutionary process driven by and rooted in the principles that constrain language itself, where "language moves down time in a current of its own making" (Sapir 1912: 160).

A key part of the hypothesis is that the biases have substantive content, which can be modeled by grounded OT constraints, specifically by the subset of OT constraint systems that the learner brings to the acquisition process (leaving open the possibility that learners may augment the set with constraints acquired empirically on the basis of language-specific evidence). The role of biases in this approach is in some respects comparable to that of the "attractor states" posited in recent typological studies (Haig 2018, Seržant & Moroz 2022, Haider 2023, van Gelderen 2024: 200) and of unmarked parameter settings (Dresher 1999, Roberts 2007), but it is conceptually different: biases are neither reified as global configurations (constructions, attractor states) nor reduced to separate parameter settings, but emerge from elementary violable constraints which interact through ranking to define a whole structured landscape of states: thus they are not extraneous to grammar but emerge from grammar itself.

This hypothesis subsumes and reconciles two previously proposed alternatives: that learners select the *simplest* grammar consistent with the data encountered (by some measure of simplicity such as the ones mentioned in 1.7), and that they select the *most restrictive* grammar. The rationale

for the former is that simplicity corresponds to generality, and that the search for the simplest grammar therefore results in the best grammar. The rationale for the latter is that a conservative learning strategy pays off because undergeneration can be corrected simply by positive evidence, but cutting back on overgeneration requires the harder task of discovering the unacceptability of some structures, either by attending to systematic gaps or by receiving explicit corrections or other negative evidence. A formal articulation of this idea is the Subset Principle (Berwick 1985, Manzini & Wexler 1987, Clark and Roberts 1993: 304-5), which claims that the learner seeks the most restricted grammar consistent with what has been acquired at each stage of the learning process. On the other hand, the well-documented class of changes that involve generalization of grammatical rules/processes would seem to show that learners on the contrary prefer the most general grammar, rather than the most restricted one.

From the present perspective, there is no contradiction here. The two apparently incompatible formal principles – simplest grammar versus most restrictive grammar – are both partial truths, and the OT-based approach pursued here subsumes them. In particular, the typology of anaphoric systems justifies a constraint system that formally imposes the narrowest binding domain, and by our assumption learners are endowed with that constraint system (in effect, the Subset Principle) and are therefore biased in favor of the narrowest binding domain for each pronoun. This correctly predicts the pattern of unidirectional historical development towards the narrowest binding domain, where pronominals become anaphors, long-distance anaphors become local anaphors, and so on (Kiparsky 2002, 2012). On the other hand, a learner will expect the context of an arbitrary alternation to be as general and natural as possible, and this is the basis for analogical generalization in historical change,

The general idea, then, is that learners are biased in favor of the most *probable* language consistent with the Primary Linguistic Data (PLD) so far encountered.

An exact, simple, and testable definition of the probability of a language can be formulated within Optimality Theory (OT). An OT grammar is a set of ranked constraints that function as a sequence of filters. The grammar maps base forms into surface forms by eliminating ungrammatical outputs. The grammatical output is obtained from arbitrary inputs by filtering them through successive constraints. Each constraint chooses only between the items that score best on the previous one. Languages differ in how they rank the constraints. Different rankings may converge on the same language, or generate different languages (Prince & Smolensky 2002).

We define the probability of a language as its *r(anking)-volume*. The *r*-volume of language *l* under *k* constraints is the number of constraint rankings (grammars) that generate *l*, divided by *k!* (Riggle 2010). The most probable language is the one with the greatest *r*-volume. *r*-volume is a promising metric because it has already proved to be a good predictor of typological distribution. The same constraints that characterize typological space also define canonical ("unmarked") systems as *r*-volume maxima. Moreover, *r*-volume is a good predictor of the relative frequency of free variants (Kiparsky 1993, Anttila 1997 *et seq.*, Bane & Riggle 2008), including syntax (Bane 2011).

The proposal is consistent with several plausible learning models. Let us assume that a learner starts with a set of unranked constraints, or some fixed ranking that encodes UG preferences, and ends with a fully ranked constraint system, a grammar whose output is the target language. Supposing that the learner at each point in the acquisition process randomly projects a language defined by the constraint ranking acquired on the basis of the Primary Linguistic Data encountered so far, then the higher *r*-volume of the language is, the more likely she is to project that language.

Alternatively, a learner can at each stage of acquisition select, among grammars consistent with the PLD, the one with the largest *r*-volume. Such a learner will make fewer acquisition errors and master the language more quickly, at the price of the substantial overhead cost of computing *r*-volumes at every stage. These are abstract models of acquisition, of course, and unrealistic until concretized to take account of the real-time acquisition. Still, they yield interesting and surprisingly accurate predictions about the direction of language change.

My proposal does not preclude functionally grounded preferences, such as those introduced by Chomsky (1995 et seq.), which favor the shortest derivation (minimizing computation and speaker effort), the shortest move (minimizing long-distance dependencies to reduce burden on short-term memory), and MERGE over MOVE (shorter outputs).[9] These should be implementable as OT constraints. More global functional criteria, defined specifically for language acquisition and with direct implications for language change, include Fitness (Yang 2000), which selects the grammar that generates the highest frequency of unambiguous outputs, the Tolerance Principle (Yang 2016), and Hawkins' (2004) Minimize Forms principle. To the extent that such functional preferences can be defined on grammars, they are compatible with the formal preference for the most probable language.

To test the hypothesis, we construct OT analyses for the subsystem of interest, pool the constraints that are active in each of them, and generate the factorial typology. The *r*-volumes of the language types should then predict the possible historical pathways between them. For tiny constraint systems the typology and ranking volumes can be calculated by hand. Previous pilot studies include the development of pronouns to reflexives, and the cycle from ergative to nominative case systems and back. Larger constraint systems require computational tools, such as PyPhon (Bane, Bowman, & Riggle 2011).

I assume that grammars are partially ranked constraint systems whose outputs define the grammatical utterances, and their ranking volumes measure their markedness and correlate inversely with their frequency and diachronic stability. Constraints constitute prior knowledge that enable language acquisition. They may be innate, or acquired through some maturation or bootstrapping process. For example, phonological feature space might be learned through babbling, and once learned, constraints could be systematically expanded into constraint families. All rankings are taken to be a priori equally probable. Change is assumed to proceed in minimal discrete steps, where a step is the fixing of a ranking relation between two constraints, or a reranking of them. That is, the learner moves from a current language state to one that is immediately adjacent in typological space.

I now turn to a small set of OT constraints that derive the basic word order typology, and underpin the diachronic pathways of innovations within it.

3.2 Building the OT analysis: Word-order constraints

Let us start with three basic word order constraints (a fourth will be added in section 6). Being OT constraints, they are all violable, and violated just in case a higher-ranked constraints forces the violation. The totality of rankings (3! = 6 in this minimal case) generates a typology, and, as we'll see, a frequency profile.

(3) a. HEAD-FINALITY: Heads follow their complements.

b. F$\prec$XP: Functional heads (operators) precede their complements.

[9]For arguments that these constraints drive language change see van Gelderen (2004, 2024).

c. HARMONY: If A dominates B, then A and B have the same headedness.

Importantly, these constraints are *grounded* in the syntax-semantic interface. Let us take them in order.

Constraint (3a) HEAD-FINALITY is just generalized subject-predicate order. Think of the core clause as a scenario, or mini-drama, where nominal arguments play roles such as Agent, Recipient, and Patient – the Thematic Roles (Theta-roles), organized in the familiar Theta-role hierarchy. In the semantic composition, predicates merge recursively with their direct arguments to discharge their Theta-roles. For example, if that the argument structure of *give* is (5),

(4) *give:* $\lambda z \lambda y \lambda$x [x CAUSE [y HAVE z]]

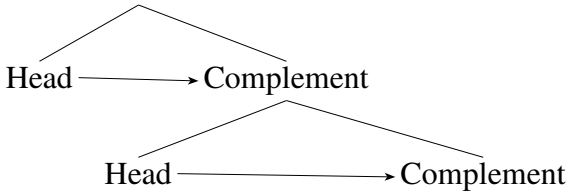
then the lowest role is the thing-given. It is discharged first, forming a new predicate with the verb, which is then applied to the next higher role, again forming a new predicate, until the highest role, the subject is reached. The external argument is the direct argument with the highest (external) Th-role. The nominal that fills the highest role is the *logical subject*, and the nominal that fills the highest structurally licensed ("direct") role is the *grammatical subject*. This generates the order and constituency of direct arguments. Thus right-headed order (S-O-V) is generalized subject-predicate order, as is S-IO-O-V when the indirect object is a direct argument rather than an oblique, as shown for *give* in (5):

(5) [Subject [Indirect Object [Object Verb]]]

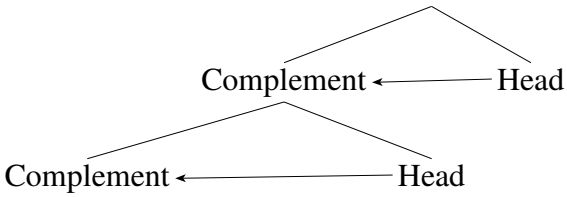
Constraint (3b) says that functional heads are preferentially initial. Unlike verbs, functional heads have no Thematic roles (and therefore no subjects or objects). They are not predicates but *operators*, subject to a special word order constraint that operators precede operands.

(3c) is the familiar Cross-categorial Harmony preference principle of Hawkins (1994), reformulated as an OT constraint. It permits stacked left-headed structures and stacked right-headed structures like (6):

(6) *a. Left-headed harmony:*



b. Right-headed harmony:



Cross-categorial harmony is what explains, among other things, the abovementioned generalization that auxiliary verbs tend to follow the content verb in OV languages and to precede it in VO languages. So far we have not made use of constraint ranking. Things become more interesting when we turn to disharmonic structures – left-headed structures stacked over right-headed structures, and vice versa. Both these disharmonic structures violate (3c). But constraint (3c) can be outranked by the other two constraints (3a) and (3b), and when they are, the grammar selects the former of these disharmonic structures as the optimum. The latter, though, remains underivable on any ranking. Let's consider a minimal structure where a functional head F governs a phrase consisting of a lexical head Y and a Complement XP, where we can take, for example, F=Aux, Y=V, XP=DP, or F=C, Y=Infl, XP=InflP, or F=Infl, Y=V, YP=VP. The rankings of the three constraints yield the six tableaux shown in (7)-(9). The arrow in the left column of each tableau points to the syntactic configuration which optimally satisfies the constraints on that ranking, which is its output.

(7) *Right-headed harmony (3 rankings)*

		HEAD-FINALITY	HARMONY	F≺XP
⇒ 1.	$[_{YP}$ XP Y] F			*
2.	F $[_{YP}$ Y XP]	**		
3.	F $[_{YP}$ XP Y]	*	*	
4.	*$[_{YP}$ Y XP] F	*	*	*

		HEAD-FINALITY	F≺XP	HARMONY
⇒ 1.	$[_{YP}$ XP Y] F		*	
2.	F $[_{YP}$ Y XP]	**		
3.	F $[_{YP}$ XP Y]	*		*
4.	*$[_{YP}$ Y XP] F	*	*	*

		HARMONY	HEAD-FINALITY	F≺XP
⇒ 1.	$[_{YP}$ XP Y] F			*
2.	F $[_{YP}$ Y XP]		**	
3.	F $[_{YP}$ XP Y]	*	*	
4.	*$[_{YP}$ Y XP] F	*	*	*

(8) *Left-headed harmony (2 rankings)*

		HARMONY	F≺XP	HEAD-FINALITY
1.	$[_{YP}$ XP Y] F		*	
⇒ 2.	F $[_{YP}$ Y XP]			**
3.	F $[_{YP}$ XP Y]	*		*
4.	*$[_{YP}$ Y XP] F	*	*	*

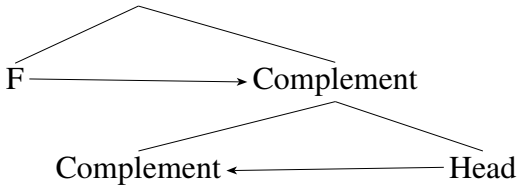
		F≺XP	HARMONY	HEAD-FINALITY
1.	$[_{YP}$ XP Y] F	*		
⇒ 2.	F $[_{YP}$ Y XP]			**
3.	F $[_{YP}$ XP Y]		*	*
4.	*$[_{YP}$ Y XP] F	*	*	*

(9) *Benign disharmony (1 ranking)*

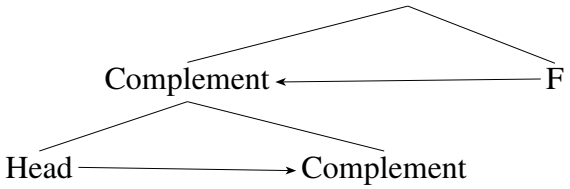
		F≺XP	HEAD-FINALITY	HARMONY
1.	$[XP\ Y]_{YP}$ F	*		
2.	F $[Y\ XP]_{YP}$		**	
⇒ 3.	F $[XP\ Y]_{YP}$		*	*
4.	*$[Y\ XP]_{YP}$ F	*	*	*

So the constraints predict that the left-headed structures headed by functional categories can dominate right-headed structures, but the reverse is not possible. Let us call the two configurations *benign* and *pernicious* disharmony, respectively.

(10) *a. benign disharmony:*



b. pernicious disharmony:



To illustrate: German has grammatical structures of the form (10a), such as *er wird Bücher lesen* 'he will read books'. But (10b) is ungrammatical in German (**er lesen Bücher wird*), and, importantly, so is its counterpart in *all* languages.

What we have just derived is similar to the Final-over-Final Constraint (FOFC) (Biberauer, Holmberg & Roberts 2014). But unlike the FOFC, it distinguishes between functional and lexical heads. That gives it some empirical advantages, to be outlined below. I think it is also conceptially superior, for it *derives* the directional asymmetry between benign and pernicious disharmony from the typologically well-supported and theoretically grounded asymmetry between functional and lexical categories.

It will be seen that five of the six possible rankings, those in (7) and (8), output the favored harmonic structures, and just one of them, (9), outputs the benign disharmony structure. The universally ungrammatical pernicious disharmony structure cannot be derived at all. That is, the *r*-volume of grammars is a measure of their typological frequency.

(11)

1.	Harmony	$[XP\ Y]_{YP}$ F	3 rankings
2.	Harmony	F $[Y\ XP]_{YP}$	2 rankings
3.	Benign disharmony	F $[XP\ Y]_{YP}$	1 rankings
4.	Pernicious disharmony	*$[Y\ XP]_{YP}$ F	0 rankings

The correlation of *r*-volume with typological preferences turns out to be general. Harmonic structures are more frequent and "natural" than disharmonic structures across languages, and pernicious disharmony is categorically ungrammatical. So we can leverage the quantitative aspect of the theory into a typology that can undergird a predictive account of drift. Here I present the first steps of a theory of word order change based on this idea.

4 OT linearization vs. FOFC

4.1 Quantitative predictions

Like the constraint system (3), the Final-over-Final Constraint (Biberauer, Holmberg, Roberts, *LI* 2014) prohibits pernicious disharmony. But unlike the FOFC, (3) derives the prohibition from the interaction of simpler grounded constraints that are needed anyway. Also, it yields more accurate empirical predictions in a range of cases. Let's look at the similarities and differences.

Germanic verb clusters can be right-headed or left-headed, or have a combination of both orders, but the overriding constraint is that pernicious disharmony does not occur. There are many differences among Dutch and German dialects in what orders they allow; for example, Swiss dialects have the harmonic Aux-V-O order shown in (12b), which is ungrammatical in standard German.[10]

(12) a. weil er den Zug sonst nicht [[erwischen können] hätte] *harmony*
because he the train else not [[catch-Inf can-Inf] had]
'because he couldn't have caught the train otherwise.'

b. wil er de Zug suscht nüme [hett [möge verwütsche]] *harmony*
because he the train else never [had [can catch]]
(Swiss, Stechow & Sternefeld 1988: 136)

c. weil er den Zug sonst nicht [hätte [erwischen können]] *benign disharmony*
because he the train else not [had [catch-Inf can-Inf]]

d. *weil er den Zug sonst nicht [[können erwischen] hätte] *pernicious disharmony*
. . . because he the train else not [can-Inf catch-Inf] had]]

(13) shows one tableau for each type; as before we exclude the two discontinuous orders in which the Auxiliary is separated from the Modal it is construed with (Verb-Aux-Modal and Modal-Aux-Verb), since they are rare, and in most dialects ungrammatical.

(13)

		HEAD-FINALITY	F≺XP	HARMONY
⇒ a.	erwischen können hätte		*	
b.	hätte können erwischen	**		
c.	hätte erwischen können	*		*
d.	*können erwischen hätte	*	*	*

		HARMONY	F≺XP	HEAD-FINALITY
a.	erwischen können hätte		*	
⇒ b.	hätte können erwischen			**
c.	hätte erwischen können	*		*
d.	*können erwischen hätte	*	*	*

[10]Note that these are not dependency diagrams; the arrows point from head to complement.

		F≺XP	HEAD-FINALITY	HARMONY
a.	erwischen können hätte	*		
b.	hätte können erwischen		**	
⇒ c.	hätte erwischen können		*	*
d.	*können erwischen hätte	*	*	*

As before, half of the tableaus derive the order (12a), which predicts that this order is typologically favored over other orders. The prediction could be tested on a sufficiently representative sample of languages. Individual languages may naturally favor different orders for any of several obvious reasons, such as adherence to a normative standard, language contact, or historical inertia. I have no frequency statistics for standard German, but (12a) seems to be more frequent than (12c) in the German dialect landscape (Schmid 2004: 154). Diachronically it is the original head-final word order, which still prevails in the higher layers of clause structure in most varieties of German. Disharmonic (12c) is a transitional pattern, which eventually yields to the full head-initial pattern found in Switzerland and in the Netherlands. In terms of constraint ranking, the originally dominant HEAD-FINALITY constraint becomes subordinated first to F≻XP, producing a mixed-headed system, and then becomes subordinated also to HARM, producing consistent head-finality. Of course, HEAD-FINALITY must outrank HARM for mixed headedness to be generated at all. So the idealized diachronic trajectory is:

(14) 1. (B) HEAD-FINALITY > F≻XP, HARM
2. (A) F≻XP > HEAD-FINALITY > HARM
3. (C) F≻XP, HARM > HEAD-FINALITY

This direction of change might seem surprising since the innovative head-initial order is the reverse of the strict head-final order of the rest of the verb phrase in the continental Germanic languages. Diachronic analysis of the verb clusters suggests that the reversal is a response to the rise of modals and auxiliaries as function words. Prior to that development, the ranking of F≻XP and HARM is moot, and only HEAD-FINALITY plaus an active role in verb clusters. The new function words are placed before their complements (F≻XP > HEAD-FINALITY > HARM), and the head-initial order is then generalized to main verbs (F≻XP, HARM > HEAD-FINALITY). Thus the verb cluster is a microcosm of the clausal domain, where the spread of head-initial VP is likewise triggered by the rise of head-initial complementizers and other heads and their activation of HARM.

The ungrammatical output (12d), with a head-initial constituent as complement of a head on its right ([Mod V] Aux) violates all three constraints. Therefore it cannot be the optimal candidate on any ranking; it is *harmonically bounded*. It is excluded in the verb clusters of at least all Germanic languages. In derivational syntactic theories it is excluded by the FOFC constraint; in the constraint-based theory proposed here it is excluded by the logic of the system itself. So far, no convincing instances of such verb clusters have been found in Germanic.[11]

Old English verb clusters with a modal governing a periphrastic tense show an analogous pattern (Haeberli & Pintzuk 2012). For subordinate clauses such as 'that he wants to have held' Haeberli & Pintzuk 2012 provide the following statistics (N=430):

[11]That includes Old English, where all the other orders are attested (Haeberli & Pintzuk 2012). The 2 3 1 verb clusters that have been claimed to exist in West Flemish and Afrikaans apparently have a different structure (Biberauer, Holmberg & Roberts 2014: 203 ff.).

(15)	þe he wyle [habban gehealden]	Modal Aux Verb	harmony	(71.2%)
	þe he [gehealden habban] wyle	Verb Aux Modal	harmony	(15.6%)
	þe he wyle [gehealden habban]	Modal Verb Aux	benign disharmony	(11.2%)
	þe he gehealden wyle habban	Verb Modal Aux	MV fronting	(1.6%)
	þe he habban wyle gehealden	Aux Modal Verb	MV extraposition	(0.5%)
	þe he [habban gehealden] wyle	*Aux Verb Modal	pernicious disharmony	(0.0%)

The three basic orders comprise 98% of the occurrences. The discontinuous examples with extraction do not not violate any of our three basic constraints, but rather a fourth constraint, tacitly assumed above, which requires the daughters of a constituent to be adjacent, and which is violated by preposing and extraposing elements from the core clause. On our assumptions, violations of that constraint must be enforced by dominant constraints which enforce discourse-driven displacement processes. The three-constraint system itself is obeyed without exception in the Old English data.

Old English Aux-Verb-Object combinations provide another test case. Old English Head-final IP (= TP = AuxP) cannot have a head-initial VP complement (Pintzuk 1991, Kiparsky 1997, Haeberli & Pintzuk 2012).

(16) a. þæt ic [mihte [God forbeodan]$_{VP}$]$_{TP}$
that I might God forbid
'that I might forbid God'

b. þæt ic mihte [forbeodan God]

c. þæt ic [God forbeodan] mihte

d. *þæt ic [forbeodan God] mihte

Verbs with particles likewise confirm the pattern:

(17) a. (þæt) se biscop [wolde [aheafan up [þæt cild]$_{DP}$]$_{VP}$]$_{TP}$. *harmony*
"(that) the bishop wanted to lift up the child."

b. (þæt) se biscop [[[þæt cild] up aheafan] wolde]. *harmony*

c. (þæt) se biscop [wolde [[þæt cild] up aheafan]]. *benign disharmony*

d.*(þæt) se biscop [[aheafan up [þæt cild]] wolde]. *pernicious disharmony*

A remarkable 7460 out of 7471 instances of disharmony (99.85%) in the York corpus are of the benign kind (Haeberli & Pintzuk 2012).

4.2 Word-order typology

So far we have seen that our constraints on different rankings account for all the attested word orders orders in a range of constructions, and exclude the unattested orders. But they go beyond the FOFC in that they predict their overall relative frequencies at the typological level. I now turn to other fundamental advantages of the constraint system (3) over the FOFC.

A key prediction of the constraints is that the basic word order is SOV (in accord with Haider 2000, and contra Kayne 1994, 2005), and that SVO arises by harmony with overt functional categories. That derives the typological generalization that languages with syntactically left-headed

lexical categories have syntactically left-headed functional categories. Languages in which all functional categories are either morphologically expressed, or syntactically right-headed, are V-final. The generalizations that clause-final complementizers entail OV and that final auxiliaries entail OV are special cases.

Consequently languages with rich inflection tend to have flat clausal structure and OV word order. The lack of left-headed projections of functional categories such as complementizers and auxiliary verbs goes with SOV word order. This includes languages like Turkish, where auxiliaries are verbs, and languages like Japanese, where they are affixes on the final verb.[12]

Like the FOFC, the constraints predict the following implicational (near-)universals.

(18) a. If CP is head-final, then TP is head-final. (If complementizers come at the end of the subordinate clause, then auxilaries follow their VP complements.)

b. If TP is head-final, then VP is head-final.

c. If CP is head-final, then VP is head-final – a consequence of (a) and (b).

Thus VO languages don't have final auxiliaries, and Aux-VP languages don't have final complementizers. (18a,b,c) appear to be exceptionless or nearly so (Dryer 1980, 1992: 102; 2009a, Hawkins 1990: 225, 256-7, 2004, Kayne 2000: 320-321). Two out of 599 languages in WALS that seem to have final complementizers and VO, Buduma (Chadic) and Guajajara (Tupi-Guarani), actually have SOV in subordinate clauses, so they are actially not exceptions (Newton 2007).[13]

Sign languages offer independent confirmation of this typological prediction. In so far as they have layered functional projections at all, they are predicted to have a basic SOV order. This seems indeed to be the default order for sign languages, with SVO the result of discourse-motivated displacement (Napoli & Sutton-Spence 2014). Newly created sign languages are SOV even where ambient spoken language is not SOV (Meir, Sandler, Padden, & Aronoff 2010). Al-Sayyid Bedouin Sign Language, which emerged untaught among deaf speakers, has dominant SOV word order, even though the Arabic of the hearing members of the community has SVO (Sandler, Meir, Padden & Aronoff 2005). Nicaraguan Sign Language also emerged on its own, and has a default SOV word order, while Spanish has SVO (Senghas, Kita & Özyürek 2004).

Still more dramatic is the experimental evidence from non-linguistic gesture tasks, including cognitive tasks that involve no communication at all. When hearing speakers of a variety of languages, including the SVO languages English and Spanish, are asked to describe vignettes with an agent and a patient, they regularly use Agent - Patient order (Goldin-Meadow & Mylander 1998, Goldin-Meadow et al. 2008, Hall et al. 2013, Kimmelman 2022).

4.3 Lexical vs. functional heads

Unlike the the FOFC, the constraint system (3) predicts that final functional heads are more restricted than final lexical heads. Thus (3) explains the generalization that left-headed nominal phrases cannot be governed by postpositions, as illustrated by the German contrast in (19).

[12]Classical Japanese did have some auxiliary verbs, for example, the original form of what is now the past tense suffix -ta.

[13]I take complementizers to be subordinating conjunctions, that is, functional heads that turn their complement clauses into arguments or modifiers. On this understanding, the final particles of Chinese are not complementizers (contra Julien 2002: 172, see Chan 2013). Question particles don't obey the constraint either. Bailey (2015) proposes that they are not complementizers, but disjunction markers.

(19) a.*[**der** **Miete** für die Wohnung] **wegen** violates FOFC and (3)
[the-GEN rent-GEN for the-ACC apartment-ACC] on account of
'on account of the rent for the apartment'

b. [**die** **Miete** für die Wohnung] **bezahlen** violates only FOFC
[the-ACC rent-ACC for the-ACC apartment-ACC] to pay
'to pay the rent for the apartment'

The Finnish contrast in (20) is similar. When the postposition is replaced by a verb, the result is grammatical though perhaps stylistically somewhat marked.[14]

(20) a. *[**osa-n** sii-tä] **läpi** ungrammatical, violates FOFC and (3)
[part-Gen it-Part] through
'through a/the part of it'

b. [**osa-n** sii-tä] **osta-maan** grammatical, violates only FOFC
[part-Acc it-Part] buy-Inf
'to buy a/the part of it'

FOFC does not distinguish between the (a) and (b) cases. But (3) predicts the grammaticality contrast between them, for (19a) and (20a) violate *all* the constraints in (3), and can therefore never be grammatical, whereas (19b) and(20b) are derivable on one ranking (the tableau is isomorphic to (9)) — see the comparative violation profiles in (21) (not a tableau!).

(21)

	F$\prec$XP	HEAD-FIN	HARMONY	
*[osaa$_N$ siitä$_{PP}$]$_{NP}$ läpi$_F$	⋆	⋆	⋆	pernicious disharmony
[osan$_N$ siitä$_{PP}$]$_{NP}$ ostamaan$_V$	✓	⋆	⋆	benign disharmony
ostamaan$_V$ [osan$_N$ siitä$_{PP}$]$_{NP}$	✓	✓	✓	harmony

Here are examples of type (20b) from the Internet (all of them by clearly native writers.)

(22) a. Sampo=han ilmeisesti o-n valmis piene-n osa-n sii-tä osta-maan.
Sampo=CL evidently be-3SG ready small-ACC part-ACC it-ELAT buy-INF
'Of course Sampo is evidently ready to buy a small part of it.'

b. Se laitet-tiin kahti-a, jotta joku suostu-i osa-n sii-tä otta-maan.
It make-PAST-PASS in.two so.that someone agree-PAST part-ACC it-ELAT take-INF
'It was cut in two, so that somebody agreed to take a part of it.'

c. Kunna-t men-i-vät … nuo tai merkittävä-n osa-n nii-stä myy-mään.
County-PL go-PAST-3PL … them-ACC or significant-ACC part-ACC they-ELAT sell-INF
'The counties started … to sell them or a significant part of them.'

d. Kuluttaja=kin joutu-u merkittävä-n osa-n hinna-sta maksa-maan Teosto-lle.
consumer.too ends.up significant-ACC part-ACC price-ELAT pay-INF Teosto-ALLAT.
'The consumer too ends up having to pay a significant part of the price to Teosto.'

e. Toivottava-sti ehdi-tään edes pien-tä osa-a nii-stä käyttä-mä-än.
hopefully manage-Pass at least small-Part part-Part they-Elat use-Inf-Illat
'Hopefully they will have time to use at least a small part of them.'

[14]The Accusative case on *osa* 'part' in (20) is assigned by the postposition and verb that governs it.

f. Osa-lla sii-tä makse-taan omistaja-n laina-n lyhennykse-t ja koro-t.
part-Adess it-Part pay-Pass owner-Gen loan-Gen payment-PlAcc and interest-PlAcc
'With a part of it the payments and interest of the owner's loan are paid.'

g. Osa-sta sii-tä teh-dään sähkö-ä.
part-Elat it-Part make-Pass electricity-Part
'From part of it electricity is made.'

Latin likewise allows all four order types with lexical heads: harmonic V[VO] and [OV]V, and disharmonic V[OV] and [VO]V. The last of these cases is the crucial one. According to (3) it is acceptable though marked, although it is an outright violation of the FOFC. And it is unquestionably grammatical; see (23d).[15]

(23) a. cum ad veritatem **coepi** [**revocare** rationem]$_{VP}$
'when I began to check my reasoning against reality'

b. [eo negotio M. Catonis splendorem **maculare**]$_{VP}$ **voluerunt**
'by that maneuver they wanted to stain Cato's good reputation'

c. **coepi** [regiones circumcirca **prospicere**]$_{VP}$
'I began to survey the areas all around'

d. [**exheredare** filium]$_{VP}$ **voluit**
'he wanted to disinherit his son'

4.4 Sublexical heads are not visible to syntax

On the lexicalist assumption that lexical structure below the word level is invisible in the syntax, perhaps because it is occulted by bracketing erasure or by some other process, our constraints predict that left-headed words in right-headed phrases are acceptable, even though they violate FOFC.

(24) a. [[bearbeiten] können] German
APPLIC-work be-able
'to be able to work on'

b. gad- bo:ŋ- te- dʒi Sora (Ramamurti 1931:49)
cut- buffalo- NONPAST- 3.PL
'they are cutting the buffalo'

For the same reason we predict that suffixes may be attached to head-initial complex stems. In contrast, syntactic theories of word formation predict that internal word structure obeys FOFC (Hein & Murphy 2022). The evidence clearly supports the former view, according to which morphology tolerates the disharmonic structures that syntax prohibits. Consider the word *dismemberment*, analyzable as [[dis-member]ment], from *dismember*, where *dis-* is the head because it turns the base *member* into a verb. The right-branching structure *[dis-[member-ment]] is excluded because because there is no noun **memberment* and *dis-* does not go on nouns anyway. Derived words with this structure are entirely productive, e.g. *antipodal, antisemitism, degeneration, degradation, devaluation, disclosure, disconnection, disavowal, dissimilarity, enchantment,*

[15] All examples in (23) are from Cicero's works, but each type can be abundantly illustrated from the usage of other authors. Unfortunately Latin has no postpositions that would allow testing the other half of the prediction.

encirclement, encryption, endangerment, enslavement, inflammation, procreation, unburdenment. With verbal heads: *do-nothingism, shunpiking, blame-the-victimism, push-button release, sitdown strike, tell-all memoir, turnkey project, get-rich-quick scheme*.

German presents essentially the same picture. *Entlastung* 'unburdening' is [[ent-last-]ung], from the verbal stem of *entlasten* 'to unburden'. It cannot be *[ent-[last-ung]], since there is no noun **Lastung* and *ent-* is never prefixed to nouns anyway, and *ent-* must be the head of *entlasten* since it turns the noun *Last* 'burden' into a verb. Similarly *Vermarktung* 'marketing', *zerbrechlich* 'fragile', *Begrüssung* 'greeting', *Enthauptung* 'decapitation', Other examples:

(25) a. Swedish: *fõrslavning* 'enslavement', *begravning* 'burial', *fõrnyelse* 'renewal', *befläckelse* 'sullying'

b. French: *tire-bouchonner* 'to corkscrew', with the adjective *tire-bouchonné* 'corkscrewed', and *portefeuilliste* "portfolio manager", *porte-parolat* 'representation', from *portefeuille* 'portfolio'.

c. Greek: *philo-soph-ía* 'philosophy' (lit. 'loving-wisdom-ness'), *philo-ksen-ía* 'hospitality', 'loving-stranger-ness', *mis-anthrop-ía* 'hating-people-ness'.

d. Sanskrit: *kurvad-rūpa-tva* 'formativeness', 'formative cause', from *kurvad-rūpa* 'formative' (lit. 'making-form').

4.5 [[F XP]$_{FP}$ Y] is benign disharmony

Another prediction is that constructions of the form [[F XP]$_{FP}$ Y] are benignly disharmonic and therefore acceptable. They are FOFC violations, but they can be generated by a ranking of (3). The reformulation of FOFC that limits it to "extended verbal projections" (Sheehan 2017: 65) makes them grammatical, but then the prepositional cases in (19) and (20) go unexplained.

(26) a. [das Buch] lesen — German
the book read
'read the book'

b. [nach Amerika] reisen
to Americs travel
'travel to America'

The tableau (27) shows the ranking that outputs (26).

(27)

		F≺XP	HEAD-FIN	HARMONY
⇒ A1.	das Buch lesen		*	*
A2.	lesen das Buch		**	
A3.	Buch das lesen	*		
A4.	lesen Buch das	*	*	*

In contrast, **lesen Buch das* is pernicious disharmony, since [[Y [XP F]$_{FP}$]$_{YP}$ violates all three constraints and therefore will always be bested by at least one other word order.

Structures of the form [[F XP]$_{FP}$ F], such as German *der Kinder wegen* 'for the sake of the children' are however still predicted to be ungrammatical, as they are by the FOFC.

The corresponding typological prediction is that postposed articles in SVO languages are not proper syntactic heads. They can be clitics, or suffixes which do not head a syntactic projection. The WALS data are consistent with this prediction.[16]

(28) a. VO languages with preposed articles: Ndyuka, Guarani, Koromfe, Iai, Tolai, West Futuna-Aniwa, Cornish, Vurës, Arosi, Siar, Owa, Mangarrayi, and Dagaare, see (29a)

b. VO languages with enclitic articles: Swedish, Bulgarian, Koyra Chiini, Husa, Chamicuro, and Fyem, see (29b)

(29) a. a ɔrri nye ayi Dagaare (Bodomo 1997: 50)
DEF berries these two
'these two berries'

b. kón awúr walá-mo Fyem (Nettle 1998: 23)
tree that big-DEF
'that big tree'

Mupun, Degema and Sisiqa, with VO and yet with DPs that can have postposed articles, are potential counterexamples. However, the glosses of the examples and the descriptions of their use suggest that they are 'weakly demonstrative definite determiners' (Lehmann 1995:55), a category half-way between a demonstrative and an article. One diagnostic of this intermediate category is that, unlike articles, it is weakly demonstrative, in the sense that it is not used to identify referents that are either unique, such as *the President, the sun*, *the Democratic Party*, or generic, such as *the average consumer,* or *the farmer who beats his donkey*.[17]

(30) Mupun (Frajzyngier 1993: 169, 278)

a. lua nə
meat DEF
'the meat' (mentioned before in the discourse)

b. n-taŋ takarda nə zak
1SG-read letter DEF again
'I read the letter again'

(31) Degema (Kari 2004: 53, cited from Bögel 2014)

Mi̩=món owéy ([nú̩ (ba̩w̩) e=kótú=n] **yo̩**)
1Sg=see=FE person that they 3Pl=call=FE DEF
'I saw the person who they called.'

(32) Sisiqa (Oceanic, Lynch, Ross & Crowley 2011)

a. aro to ko piso **ne**
1SG C 1SG:REAL break DEF
'It was I who broke it'

[16] A fourteenth language, Angami, has been erroneously classified as SVO in the WALS typology, but it seems not to be SVO.

[17] Babungo has postposed "demonstrative adjectives", which surely are not function words.

b. Pötamama te ko mia =ro **ne**
Pötamama C 1SG:REAL see =1SG: DEF
'it was Pötamama that I saw'

I conjecture that these are not articles but weak demonstratives. That Degema *yọ* is a reduced demonstrative is moreover suggested by the existence of a corresponding full demonstrative *owéey yọ́o* 'that person', *ewéey yọ́o* 'those persons' (Kari 1997).

In nearly all verb-final languages (97%) adpositions are postpositions (Dryer 2013, Huber 2013), though some have prepositions, German *nach Berlin gegangen* 'gone to Berlin'

Prepositional phrases can be complements of verbs on their right, as in (26b). This fits the prediction well. Cross-linguistically, PP and VP order harmonize strongly (Dryer 2011:

(33)

	Prep.	Postp.
VO	456	42
OV	14	472

The substantial number of VO languages with postpositions is however troubling, since this combination should yield pernicious disharmony. But it may be an artifact. What WALS and APiCS categorize as postpositions in VO languages are almost always location nouns (often bound nominal stems denoting locations or body parts with locative case inflection) with the located object in the genitive. While they are translational equivalents of the adpositions of familiar languages like English, they have the morphology and syntax of nouns in those languages, including genitive case on the complement. Examples are Dagaare (Bodomo 1997) and Ewe (Warburton et al. 1968: 48). In Dagaare 'front' is derived from 'forehead', and the resulting adposition mirrors the order of genitive and noun (since Dagaare is GenN, this results in a "postposition"). So the postpositions of these languages are nominal (as Dryer points out). In the case of locatives, they are forms of bound nominal stems that denote locations, and therefore not function words (operators) in the relevant sense.

5 Word-order change

5.1 Diachronic predictions

Having justified our minimal OT constraint system on typological and theoretical grounds, we return to its historical consequences. Recall the hypothesis that long-term drift and convergence (homoplasy) are bias-driven change towards more probable languages, where probability is measured by *r*-volume. I laid out a system of three grounded constraints (3) that delivers a theory of core phrase structure and linearization and makes it possible to quantify the *r*-volumes of the basic syntactic structures. These make a number of predictions about the direction of language change.

One of the key diachronic predictions that follow is that functional heads are lexicalized in top to bottom order. It is confirmed by the syntactic trajectory of Indo-European syntax as attested in texts and reconstructed by the comparative method. The proto-language had no *that*-type clauses (subordinate finite complement clauses), and relative clauses were adjoined. New subordinating complementizers developed in the daughter languages, all of them left-headed, in conformity with the F$\prec$XP constraint (emergence of the unmarked). Some languages then add lower functional projections such as AuxP (Haig 2006, 2008 for Iranian), MoodP, and PolP/ΣP (Laka 1990), headed by Polarity items (negative, affirmative, and emphatic), and a site of focused constituents

(cf. Condoravdi & Kiparsky 2002 on Greek and McCloskey 2010 on Irish, and perhaps fn. 6 on Kashmiri). These new functional categories are left-headed too, as expected. The final episode of this development was the generalization of the head-complement order from these superordinate constituents to the VP by the promotion of the HARMONY constraint. The Indo-Aryan languages, which gained initial complementizers only late through contact with Persian and other Iranian languages, retained the original OV order. All Indo-European languages that changed from OV to VO word order had acquired initial complementizers and some intermediate projection. For example, both developments took place in English and Scandinavian, whereas neither of them took place in continental West Germanic.

We are now in a position to explain the basic word order cycle. The OV > VO phase of the cycle is initiated by the rise of functional projections by grammaticalization, beginning with left-headed CP. It initiates a downward cascade of shifts from head-finality to head-initiality, terminates in the lexical categories (OV $>$ VO), as in Germanic (and other I.E. except Indo-Aryan), Uralic, Afro-Asiatic, and Niger-Congo. The second phase of the cycle is triggered by the loss of these functional projections by morphological attrition and fusion (univerbation), resulting in a reversion of VO to OV. As discussed below in sections (5.5) and (5.6), this phase is arguably revealed by comparative reconstruction in the Munda branch of Austro-Asiatic and Niger-Congo, and by internal reconstruction in Motu (Papuan branch of Oceanic), Athabascan, and Basque, where the former VO syntax is fossilized in word structure ("today's morphology is yesterday's syntax", as T. Givón's slogan has it). In this way the loss of syntactic functional heads through morphologization and their renewal through grammaticalization drives a long-term head-final $>$ head-initial $>$ head-final cycle.

The transition from one canonical system to another necessarily goes through non-canonical systems. A Sanskrit-type harmonic head-final syntax cannot turn into an English-type harmonic head-initial syntax in a single step. It proceeds in minimal steps — modeled in OT as one constraint reranking at a time. But the intermediate disharmonic stages strictly conform to universal constraints. For example, if a special constraint that requires initial complementizers (C$\prec$XP) dominates XP$\succ$Y and the other constraints rank lower, complementizers are initial and other heads are final, as in German. Other constraint rankings yield strict head-initiality and strict head-finality. The grammaticalization of initial functional heads triggers a downward cascade of head-final to head-initial shifts, terminating in the lexical projections (e.g. OV $>$ VO) shift driven by cross-categorial harmony. Further grammaticalization of functional heads to inflections triggers restoration of OV. But every episode in the pathway around the cycle is derivable by some constraint ranking. For example, German is V-final but C-initial, but a V-medial and C-final intermediate stage never arose and could not have arisen.

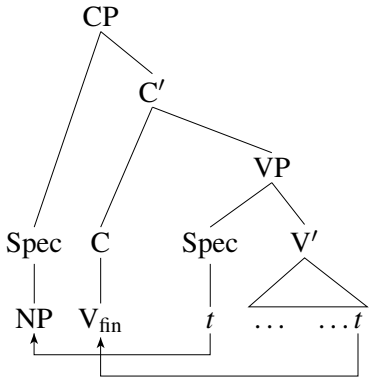
The implication for typological research is that existing languages are a fair sample of the possible variety of grammatical structure. No feature in the human linguistic repertoire has been permanently erased by language change, and whatever may be appear in the future is likely to have already existed in the past. Proto-languages were not qualitatively different from existing and historically recorded languages. That should not be surprising if we believe that the language faculty evolved as part of human nature, and it justifies the reliance on typological guardrails in historical reconstruction.

5.2 Germanic

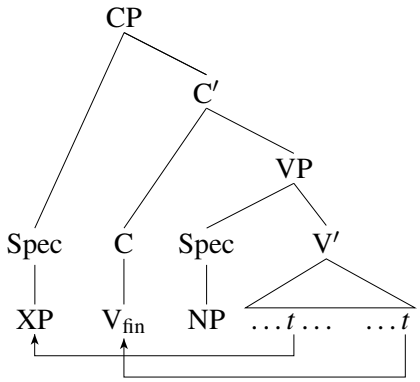
Germanic languages fall into two types. The first comprises German, Dutch, West Flemish, Afrikaans, and Frisian – essentially continental West Germanic with the exception of Yiddish.

This group has no V-movement in subordinate clauses. These languages are standardly (though not uncontestedly) assigned a less ramified constituent structure than the other Germanic languages. C takes VP complements, and subjects sit in Spec-VP. There is no TP (or PolP/ΣP) constituent above VP. The finite verb moves to a vacant C, and the subject or another constituent can be moved before it to yield V2 in main clauses. A sparse analysis of this type is viable for all languages of this group (Haider 1997, 2014, Vikner 2001, 2020, Fuss and Trips 2003, Sternefeld 2006: 322-343, Axel 2007: 104-110, Axel-Tober 2017, 2018).

(34) With subject fronting:



With object/adverb fronting:



The second group comprises all other Germanic languages now spoken. Below the CP layer that all Germanic languages share (including Gothic, Eythórsson 1996), they have additional functional projections. The generalization is that all and only the Germanic languages that developed such additional functional projections switched from OV base order to VO.

At least one such functional projection seems to have emerged already in Old English. Wh-pronouns, *ne*, and *þa* were placed in Spec-CP, with the finite verb immediately following in the C head. Topics could be adjoined to the highest clausal projection (hanging topics, left dislocation). Below the CP layer, there were at least two positions where subjects could be placed — a higher position for discourse-given arguments, regularly housing subject pronouns, occasionally topical definite DP subjects and object pronouns, but never adverbs or PPs, and a lower position for discourse-new, generic/indefinite, and focused nominal subjects (Haeberli 2000, 2005, 2007, van Kemenade & Los 2006, van Kemenade, Milicev & Baayen 2008, van Kemenade 2009, van Kemenade & Milicev 2012, van Kemenade & Links 2020). Haeberli identifies the lower position as SpecTP, and until that constituent became obligatory, it could simply have been SpecVP, as in German. He labels the higher position SpecAgrP, though it is not clear why agreement would what be demarcates old information from new information, nor even whether there is any empirical evidence for a separate syntactic AgrP constituent in Old English. A clue to the nature of the two positions comes from the fact that when discourse particles like *la, nu*, *þa, þonne*, and the pleonastic negation *ne, ná* occur with them, they are usually placed between them. These particles have affirmative, negative, and inferential functions; the positive ones can be glossed as "then", "so", "indeed" (German "wohl"). *La* normally occurs in questions and conveys surprise or a negative implicature (like German "denn").[18]

[18]Most of these items also have other functions than as discourse particles: *la* is a vocative "O" and an interjection "Oh!", *þa, þonne.* and *nu* are temporal adverbs "then", "now", and *þonne* also serves as a relativizing conjunction "when".

(35) a. Ne beheolde ðu **la** minne ðeowan iob . . . ?
not considered you then my servant Job
'Have you not considered my servant Job?' (*ÆHom* 35:261)

b. Hu mæg **la** se blinda lædan þone blindan?
how may indeed the blind lead the blind
'How can the blind lead the blind? (*ÆHom* 13.124)

c. hwi læddest þu **la** us of Egipta lande?
why led thou indeed us from Egypt land
'Why did you lead us from the land of Egypt?' (*ÆHom* 21:304)

d. Eart þu **la** furðor þonne ure fæder Iacob? (*ÆHom* 23.693)
'Are you then greater than our father Jacob?'

e. Hwæt mæg ic **la** nu don?
'So what can I do now'
(*ÆHom* 10.2371)

f. Eart þu **la** God? *ÆLives* 72.2209
are you then God?
'Then are you God?'

g. he **þonne** þe eac from tinttregum genereð
he then you also from torments saved
'then he also saved you from torments' (*Bede* 2:9.132.29)

h. Ne eom ic **ná** Crist
not am I not Christ
'I am not Christ' (*John 1.20*, ed. Skeat)

i. Ne geortriéwe ic **ná** Gode þæt hé ús ne mæge gescildan
not distrust I not God that he us not can shield
'I do not distrust God that he can save us' (*Orosius* 2.5)

If we assume that these items head a PolP constituent (see above), then arguments are to the left of it, possibly adjoined to it, and "new information" arguments on its right.

5.3 Latin

The early stages of drift from mixed head-initial and head-final Latin word order to consistently head-initial Romance is well studied (Ledgeway 2012, with references to much previous work). The earliest Latin already has strictly head-initial CP and PP, with other categories showing variable headedness, and relative clauses still of the correlative type. There is a steady increase of head-initiality throughout Latin, including Verb-Object, Auxiliary-Main Verb, Noun-Adjective, Noun-Genitive. The enclitic *-que* 'and' is replaced by the free-standing conjunction *et*. This is another hallmark of the new syntax, since *que*-type conjunctions are typical of head–final languages and *et*-type conjunctions are typical of head–initial languages (Stassen 2000; 2001).

The exact chronology of the drift is difficult to determine, because late Latin had a strong literary standard, verging on diglossia. It is clear that it went to completion in the early Romance vernaculars. Adams (1976a: 88) even claims that "Latin is basically VO in type from the time of the earliest texts.", since "the classical language does not represent a chronologically linear variety of Latin as such but, rather, a highly stylized, literary variant of the language".

5.4 Greek

The rise of syntactically projected functional structure in Greek can be tracked by concurrent changes in word order, auxiliation, and clitic placement.

The head-first word orders Verb-Object, Noun-Genitive, and Preposition-Object emerge from the archaic period to postclassical Greek (Seržant & Rafiyenko 2020). SOV is the unmarked order in Homeric Greek. SVO and VSO then increase in frequency, until in Koine Greek, SOV becomes the marked order; indeed, it becomes ungrammatical with full DP objects (Lavidas 2015), see also Kirk (2012) on New Testamant Greek.

Whether there is any syntactic projection between CP and VP already in Homeric Greek is questionable. Negation and modal particles such as *án* may be functional heads, but this remains to be shown. Late Greek sees the beginnings of an auxiliary system, with periphrastic perfects, and finally periphrastic futures *θelo ina* > *θa*. I assume that this reflects the rise of a separate auxiliary head in the syntax.

According to Lavidas (2015), Classical Greek has only V to C movement, whereas Koine Greek has V to T movement and "meets all criteria for V-to-T languages".

A diagnostic for the emergence of a more articulated phrase structure in post-Homeric Greek is the placement of object clitics. Their landing site changes from the Wackernagel second position in Homeric to a seemingly lower position in later Greek. Arguably the clitic position remains unchanged, in that object clitics remain at the beginning of their VP, and what changes is actually the functional structure above the VP. In Homer's Greek, where no intermediate syntactic functional category exists yet, the VP-initial position, to which clitics are adjoined, coincides in main clauses with the CP-initial position, where the clitics undergo prosodic inversion if necessary to satisfy their enclisis requirement. This is to say that Homeric clitics are second position (Wackernagel) clitics. As mentioned above, a second functional projection, ΣP, headed by (lower) negation, assertion markers and modal elements arose. After the emergence of this higher intermediate projection (left-headed like all innovating functional structure, therefore inducing verb-medial order), we can maintain exactly the same cliticization rule for all stages, viz. that clitics go to the beginning of VP, which now is no longer the left-peripheral Wackernagel position:

(36) a. $_{\text{CP}}$[… … $_{\text{VP}}$[X *clitic* …]] Homeric Greek

b. $_{\text{CP}}$[… $_{\Sigma\text{P}}$[… $_{\text{VP}}$[X *clitic* …]]] Koine, New Testament Greek

5.5 VO > OV

Our hypothesis attributes VO order to left-headed functional projections that induce VO by HARMONY. It follows that the the loss of those functional projections sets in motion a return to OV, completing the word order cycle. Historical linguistics suggests that such loss results from the incorporation of syntactic functional material into the verb morphology (as in Japanese), or from phonological weakening and deletion. In either case, it is likely to be more rapid than the buildup of new heads by grammaticalization. That might explain why no clear cases of ongoing reversals of this type have been documented. However, very clear historical evidence for such reversal comes from the internal reconstruction of morphology in languages that have undergone it.

Perhaps the clearest instance of VO > OV is found in the Austro-Asiatic family. It has two main branches, Mon-Khmer and Munda. The Mon-Khmer languages are predominantly isolating (that is, their words are mostly monomorphemic) and they have SVO syntax. This is also the

structure reconstructed for Proto-Austro-Asiatic. The Munda languages, though, which include Santali, Mundari, Sora, and Kharia, have developed complex words, rich agreement inflection, and a fairly rigid right-headed S-IO-O-V-Aux clausal order. The auxiliary verb is generally the finite verb of the clause and appears in clause-final position following the lexical verb in the sentence. Anderson (2007) posits SOV constituent order for Proto-Munda and for all its sub-branches.

The evidence for positing earlier SVO for Munda and Proto-Austro-Asiatic is the clear layer of SVO preserved in word structure, including head-initial negative, causative and desiderative prefixes, and noun incorporation:

(37) a. ŋɛn əd- **məl- jom- jɛl- yə-** aj- t- en- ay
I NEG- want- eat- meat- fish- all- PST- TR- 1SG.SUBJ
'I don't want to eat all the fish' (Sora, data from Ramamurti 1931, via Anderson 2007.)

b. aliŋ baro tʃa ŋu- te= liŋ sanaŋ- tan- a
we.DU two tea drink- ABL= 1DU DESID- PROG- FIN
'We two wish to drink tea' (Ho; Anderson 2016)

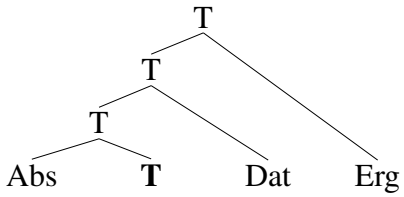
Morphological noun-incorporation preserves the old VO word order. Incorporation is semantically transparent, as the examples in (37) illustrate, but it is clearly lexical rather than syntactic, as shown by the fact that incorporated nouns are truncated to monosyllables: in (38a), *kid-* 'tiger' is reduced from its free-standing counterpart *kina-* seen in (38b). Not all nouns have a monosyllabic truncated form, and those that do not have one do not incorporate, since morphology must satisfy constraints on word structure.

(38) a. **ŋam- kid-** te- n- ai
seize- tiger NONPAST- INTRANS CISLOC.1
'I will seize the tiger'

b. **kina-** n **ŋam-** t- am
tiger- NOUN seize- NONPAST- 2
'The tiger will seize you.'

A second plausible case is Basque. Modern Basque is SOV, but the archeology of the Basque verb is believed to disclose its former SVO syntax. In the verb complex, the auxiliary head is in second position (Arregi & Nevins 2012).

(39)



The SVO > SOV shift in Basque is thought to be a result of a passive-to-ergative shift (Trask 1977: 207, Gomez & Sainz 1995, Verdú 2010: 121). The auxiliary complex reflects the old word order Subject-Verb-Object-Oblique (where the modern Absolutive, the unmarked case, was a Nominative). By the same reasoning, the causative prefix *ra-* can be identified as an SVO remnant frozen in the verb morphology. The Athabascan verb is a bipartite structure that very likely goes

back to a comples Aux + Verb constituent. It has been supposed that the Indo-European past tense prefix ("augment") *h_1e- is a downgraded Tense head (Hajnal 2016), perhaps cognate with the invariant *-*e* portion of the perfect Sg. *-*He,* *-*tHe,* *-*e*.

More straightforward is the case of Motu (Austronesian, western Melanesia). Blust (2024) summarizes the argument: "The structure of the verb complex provides a clue that the current SOV order of sentence constituents has developed from an earlier SVO order." "The verb complex consists of a prefixed subject marker, the verb stem, and a suffixed object marker, together with free nouns or pronouns marking subject and object, producing structures such as 'the man the dog he-kicked-it' for 'the man kicked the dog.'"

5.6 Bantu and Niger-Congo

If we knew the Germanic languages only in their current state, and applied the standard procedures of the comparative method to their clausal syntax, we might well reconstruct a head-initial proto-language. For that would be the reconstruction that requires the fewest change events, and the simplest ones, which sound method requires us to prefer, other things being equal. Since the Scandinavian and English branches are completely head-initial, and the West Germanic languages (German, Dutch and their close relatives) are head-initial except for their head-final finite subordinate clauses and some postpositions, those guidelines favor reconstructing a fully head-initial proto-syntax, with no change in Scandinavian and English, and just a reversal in the West Germanic VP and PP. The alternative reconstruction, positing a head-final proto-syntax with an across-the-board reversal to head-initiality in English and Scandinavian, and a partial reversal in West Germanic in all categories except VP and PP, would be rejected because it posits more drastic and numerous changes in the daughter languages than the head-initial reconstruction does.

However, while the rise of a disharmonic order from a consistent left-headed order is easy to state, it is hard to explain. In the light of naturalness considerations such as those formalized here in terms of ranking volume, the head-final reconstruction would gain some plausibility. And indeed it would be the hypothesis that the actual historical data confirms.

Such a thought experiment comes up for real in the reconstruction of Bantu and the larger Niger-Congo family to which Bantu belongs. The great majority of Bantu languages are consistently head-initial, with Prepositions, Verb-Object, Aux-Verb, C-Comp, Verb-Oblique, and Verb-Adverb. Most Bantuists naturally reconstruct this clause structure for Proto-Bantu. But a few of the languages step out of line. Tunen (a.k.a. Nen) is mostly head-initial, but it has a head-final VP, regardless of tense and information structure, viz. Subj-Tense-IO-DO-V (Gensler 1994: 6, Kerr 2024).[19] If we follow majority rule and reconstruct a consistent head-initial Subj-Aux-V-Obj constituent structure, it is not clear why some languages would have reversed the VP's headedness to get a disharmonic structure (see Kerr 2024, section 4.3 for a review of hypotheses). On the opposite assumption, Bantu would have traversed the well-established pathway that we know from the Indo-European languages, and Tunen would then be a language that preserves the pre-final stage of the process, as German does.

Internal reconstruction seems to offer some support for this alternative scenario. The typical Bantu prefix-root stem constellation of the form Subject-Tense/Aspect-Object-V-… looks like an S-Aux-O-V clause in miniature.[20]

[19] In the Tunen dialect described by Mous 2014, information structure seems to play a bigger role, but the basic order in the VP is still recognizably OV.

[20] In the same way, the new suffixed perfect and future tenses in Romance, and the weak preterite in Germanic, all derived from periphrastic forms, are *Leitfossils* of former V-Aux word order.

(40) a. á- cha- ri- téng- és- a
3Sg.Subj- FUT- 3Sg.Obj buy- CAUS- FV
'He/she will sell it.' Shona (Myers 1990:27)

b. k'i-tl'a-kw'ál-él-él-a ŋwana bats'álí lūkw'áló
1s-FUT-write-APP-APP-FV child parents letter
'I'll write a letter to the parents for the child.' Tswana (Nurse et al. 2016: 63)

The morphological structure of the Bantu stem was proposed as a fossilized residue of original S-Aux-O-V syntax even before the discovery of actual S-Aux-O-V systems in Bantu (Givón 1971, Hyman 1975, Williamson 1986, Nurse 2007).[21]

S-Aux-O-V and even completely head-final syntax is found also outside of Bantu in other branches of the larger family to which Bantu belongs, the Niger-Congo languages (the major branch of Greenberg's putative Niger-Kordofanian phylum).[22] In a pattern reminiscent of West Germanic, Kisi (in the Mel subfamily) has Subj-Aux-IO-DO-Verb-X alternating with Subj-Verb-IO-DO-X, presumably with the verb fronted to the Aux (Tense) position if there is no overt auxiliary to fill it (Creissels 2018: 781).

(41) a. À wá ndú kòówáŋ kìóó
3PL PAST.PROG 3SG medicine give
'They were giving him medicine.'

b. Ò ké yá tòòlúlán
3Sg give 1Sg support
'She gave me support.'

In the Kru subfamily and in Senufo, the fronting of the main verb to a vacant auxiliary slot is paired with another notable feature, the restriction of the VP to a single direct object preceding the verb, with all other arguments and adjuncts following it. The verb is placed after the direct object when there is an auxiliary, and before the direct object when there is no auxiliary (S-Aux-O-V-X but S-V-O-X). In Koopman's Germanic-style synchronic analysis of this type of variable verb positioning in Vata (1984: 28), the main verb moves from its basic VP-final position to an empty IP (TP, AuxP) head unless it is filled by an auxiliary.

(42) a. ṅ **ká** nɑ̀ gòlì mlÍ́ pùtù̀ **sà** Vata
I FUT-A my mounds in grass remove
'I will clear the weed from my mounds.'

b. ɑ́ **lɛ̀** bÌ̀ sɑ̀ká Vata
I eat now rice
'I am eating rice right now.'

[21]The prefixes between the subject and object mark a number of tense/aspect categories, varying by language, which would on this account be the residue of an old auxiliary. After the root comes an "Extension" consisting of argument-structure modifing suffixes (diatheses in an extended sense), the most common of them being causative, applicative, reciprocal, and passive, in that order if there is more than one, and an obligatory "Final Vowel" suffix that marks perfect and imperfective outer aspect (Nurse et al. 2016: 66). This would then be the oldest part of Bantu word structure. It is followed by a layer of more loosely attached locative and other suffixes or clitics.

[22]There is some disagreement among specialists on the nomenclature and on the subclassification; the phylum itself has been adopted as a working hypothesis but has not been definitively established by the comparative method (Güldemann 2018: 107-234).

Koopman's synchronic analysis of the Kru system offers a simple diachronic derivation of it from a reconstructed OV base order, analogous to the one which is well established for the development of Germanic from Indo-European. The original order would thus have been S-O-V, and after the rise of auxiliaries, S-Aux-O-V, with subsequent fronting of V to Aux, resulting in S-V-O. The rigid verb-final syntax of Ịjọ is then explained by the fact that it lacks auxiliary verbs, and therefore has no Aux projection to host a fronted V. If Ịjọ once had auxiliaries, they must have been lost (perhaps by merging morphologically with the main verb) before V-to-Aux fronting entered the language.[23]

The Dogon languages are consistently head-final, with SOV and predicate-copula word order, postpositions, and final complementizers (Heath 2008, Chs. 2 and 17). There is no clear indication of an auxiliary, though the suffixed negation and aspect morphology might reflect former postverbal auxiliaries. Possessors (including subjects and objects of nominalizations) precede their nominal heads.[24]

(43) a. [[mì dêː] lè] búːdù óː-sà-m
[[1SgP.L father.HL Dat] money give-Reslt-1SgS
'I gave the money to my father' (Heath 2008:18)

b. ñăː ñèː-lí-∅ má wá
meal eat-PerfNeg-3Sg Q Quot
'[it was asked,] "did he/she not eat?"' (Heath 2008:119)

Another SOV group of Niger-Congo languages is Ịjọ (Williamson 1986). These examples are cited from the Biseni variety (Worufah & Obikudo 2023).

(44) a. Akpoebi fún tụ̀bo-ni Laye piri-ma
Akpoebi book buy-for Laye give-PST
'Akpoebi bought a book for Laye'

b. Wá Laye ari-má
We Laye see-PST
'We saw Laye'

If the conjecture that Bantu verb morphology reveals original S-O-V syntax is correct, the development would have followed an IE-type trajectory in which left-headed functional categories are added – first functional heads (complementizers and then auxiliaries), resulting in benign disharmony, and finally a switch in the orientation of the VP, restoring a fully harmonic syntax, but left-headed this time: S-O-V > Comp-S-O-V > Comp-Aux-S-O-V > Comp-Aux-S-V-O.

Two other typologically notable features are found in Niger-Congo. Mande languages restrict the core VP to a verb with no more than a single direct object before it (e.g. (45a)); everything else is placed after the verb, including all oblique objects (45b) and adverbials/prepositional phrases (Creissels 2005, Nikitina 2008).

(45) a. Mŏo-lù yè báa tèyi.
person.D-PL CPL.TR river.D cross
'The people crossed the river' (Mandinka, Creissels 2018: 277)

[23]The post-root part has undergone especially heavy attrition in parts of the Bantu family, and has even been lost completely in some distant relatives, notably Yoruba.

[24]The other well-described Dogon language Tommo So, works the same way (McPherson 2013: 13-16), except that its noun phrase order is much freer, although the domains of tone spread indicate a highly articulated nominal constituent structure.

b. Móo-lù tèyí-tà bâa lá.
person.D-PL cross-CPL.INTR river.D POSTP
'The people crossed the river'

The several types of OV syntax in West Africa are consistent with a Germanic/Romance/Greek-type scenario, where a harmonic right-headed SOV language acquires high left-headed functional categories, which then trigger VO by syntactic harmony. Some researchers, however, reject the derivation of S-V-O from O-V, and the internal reconstruction of earlier syntax from verb structure. They hold Niger-Congo to have been originally S-V-O or S-V-O-X (Heine 1976, Claudi 1993, Güldemann 2018: 149). Claudi (1994) suggests for Kru that deverbal nominals with preceding objects were reanalyzed as verbs, resulting in two verbs, of which the first V then became an auxiliary (S-V-O-N_V $\rightarrow$ S-V-O-V $\rightarrow$ S-Aux-O-V). It is hard to make sense of the intermediate stage resulting from the first step in this development, and objections to the second stage have been raised by Bearth (1995) and by Kastenholz (2003). Nikitina attributes Mande S-O-V-X word order to reanalysis of noun phrases containing nominalized verbs as verb phrases, or complements of auxiliaries.[25] See further Gensler (1994, 1997), Gensler & Güldemann (2003), Güldemann (2008:159-62), Creissels et al. (2008:127).

6 Hill-climbing

A standard assumption in OT is that markedness is visibility. It has been assumed that every constraint that applies to a feature or category has a counterpart that applies to its marked value or member. So the general constraint that requires F$\prec$XP order in functional categories can be paired with a special constraint C$\prec$XP, which requires that order specifically in the CP projection. The new constraint does not change the previous results, but it provides a more fine-grained differentiation of functional heads. Let us explore the trajectories in grammatical space that it predicts – still far from the full picture, but illustrative of the power of the approach.

Setting aside discontinuous constituency, there there are eight combinations of Complementizer, Auxiliary, Main Verb, and Object, of which four are potentially grammatical. Given the four constraints – (3) and the additional one that we just added – illustrative tableaus and ranking volumes are shown in (46)-(49). (The dotted column-separating lines signify absence of crucial constraint ranking.)

(46) *A1: Strictly head-initial (10 grammars)*

		C$\prec$XP	F$\prec$XP	HARM	HEAD-FIN
$\Rightarrow$ A1.	Comp [Aux [MV Obj]]				***
A2.	Comp [Aux [Obj MV]]			*	**
*A3.	*Comp [[MV Obj] Aux]		*	*	**
A4.	Comp [[Obj MV] Aux]		*	*	*
*B1.	[Aux [MV Obj]] Comp	*	*	*	**
*B2.	[Aux [Obj MV]] Comp	*	*	*	*
*B3.	*[[MV Obj] Aux] Comp	*	**	*	*
B4.	[[Obj MV] Aux] Comp	*	**		

[25] Aboh 2004 and Kandybowicz 2008 derive OV orders, in Gbe and Nupe respectively, from VO synchronically by leftward object shift. Their analyses are is couched in a Kaynian framework in which *all* Complement-Head orders are a priori derived by movement. Whether this is right is a question full of interest, but fortunately we can sidestep it. It is the output order and constituency that counts, because that is what harmony, the FOFC, our constraints, and typological generalizations are all defined on, no matter how they are syntactically derived.

(47) *A2: Head-initial CP and IP (4 grammars)*

		C≺XP	F≺XP	HEAD-FIN	HARM
A1.	Comp [Aux [MV Obj]]			***	
⇒ A2.	Comp [Aux [Obj MV]]			**	*
*A3.	*Comp [[MV Obj] Aux]		*	**	*
A4.	Comp [[Obj MV] Aux]		*	*	*
*B1.	[Aux [MV Obj]] Comp	*	*	**	*
*B2.	[Aux [Obj MV]] Comp	*	*	*	*
*B3.	*[[MV Obj] Aux] Comp	*	**	*	*
B4.	[[Obj MV] Aux] Comp	*	**		

(48) *A4: Head-initial CP (2 grammars)*

		C≺XP	HEAD-FIN	F≺XP	HARM
A1.	Comp [Aux [MV Obj]]		***		
A2.	Comp [Aux [Obj MV]]		**		*
*A3.	*Comp [[MV Obj] Aux]		**	*	*
⇒ A4.	Comp [[Obj MV] Aux]		*	*	*
*B1.	[Aux [MV Obj]] Comp	*	**	*	*
*B2.	[Aux [Obj MV]] Comp	*	*	*	*
*B3.	*[[MV Obj] Aux] Comp	*	*	**	*
B4.	[[Obj MV] Aux] Comp	*		**	

(49) *B4: Strictly head-final (8 grammars)*

		HEAD-FIN	C≺XP	F≺XP	HARM
A1.	Comp [Aux [MV Obj]]	***			
A2.	Comp [Aux [Obj MV]]	**			*
*A3.	*Comp [[MV Obj] Aux]	**		*	*
A4.	Comp [[Obj MV] Aux]	*		*	*
*B1.	[Aux [MV Obj]] Comp	**	*	*	*
*B2.	[Aux [Obj MV]] Comp	*	*	*	*
*B3.	*[[MV Obj] Aux] Comp	*	*	**	*
⇒ B4.	[[Obj MV] Aux] Comp		*	**	

On the assumption that learners seek the most probable grammars and that they revise their provisional analyses one step at a time, this still very simplified model yields the following idealized trajectories for acquisition and change:

(50)
- Acquisition of B4: A1 → A2 → A4 → B4.
- Change from B4 to A1: B4 > A4 > A2 > A1.

(50) proceeds as shown in (51):

(51)

Word Order	R-volume	
A1. Comp Aux MV Obj	10/24	
A2. Comp Aux Obj MV	4/24	HARM promoted
*A3. Comp MV Obj Aux	0/24	
A4. Comp Obj MV Aux	2/24	F≺XP promoted
*B1. Aux MV Obj Comp	0/24	
*B2. Aux Obj MV Comp	0/24	
*B3. MV Obj Aux Comp	0/24	C≺XP promoted
B4. Obj MV Aux Comp	8/24	HEAD-FIN > C≺XP, F≺XP

In the long term, languages move from one optimum to another (here B4 to A1), or in terms of the hill-climbing metaphor, from hilltop to hilltop across intervening valleys. The question is how this happens. The bias-driven model of change posits that learners over-optimize in the course of acquisition, projecting grammars with the largest *r*-volume even though they generate languages that are different from the target norm. Their innovative outputs (especially at the final stages of acquisition where they are not saliently distinct from the ambient norm) will be similar to those of other learners in the speech community because they are projections of the same biases, and may be adopted by other learners. This pushes the language in the direction of a new, more optimal grammar. The result at each stage is a new age-graded sociolect which differs minimally from the ambient norm. In the long term the cumulative effect of such small changes is drift.

This mechanism allows valley-crossing transitions like (50.) A miniature example may clarify this point. When enclitic =*que* 'and' and =*cum* 'with' were replaced by the free-standing *et* and *cum* in Latin (a development that was part of the drift to head-initiality), the old form *nobiscumque* 'us-with-and' (right-headed harmony)[26] was replaced by *et cum nobis* 'and with us' (left-headed harmony). But this change spawned a third, mixed intermediate structure *et nobiscum*.

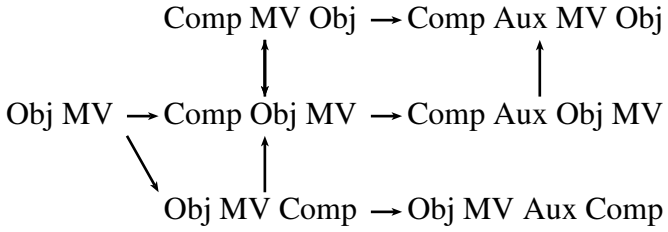
(52)

a.	et cum nobis	left-headed harmony
b.	nobiscumque	right-headed harmony
c.	et nobiscum	benign disharmony
d.	*cum nobisque	pernicious disharmony

Being an instance of benign disharmony, it is grammatical, albeit a complication. But it is easy to see how it might nevertheless have entered the language as a transitional hybrid between the old head-final form and the new head-initial one.

The *r*-volumes in (46)-(49) define a rudimentary holey fitness landscape (Gavrilets 2004), and the paths by which a language can traverse it in the first part of the cycle.

(53)



[26]E.g. *cum Caesare nobis cumque* (Cic. *Fam.* 9.9.1).

7 Conclusion: The word order cycle

When violable constraints are formalized and ranked in an OT grammar they interact to generate both hard constraints on grammatical structure and preferences that impact the types of diachronic change that occur. When languages gain functional projections by grammaticalization, they undergo a downward cascade of head-final to head-initial shifts, terminating in the lexical categories (e.g. OV > VO). When languages lose functional projections by morphological attrition and univerbation, they revert to OV. The loss of syntactic functional heads through morphologization and their renewal through grammaticalization drives a long-term head-final > head-initial > head-final cycle.

This approach to change provides a rationale for long-term unidirectional change within a language ("drift"), and independent parallel changes across different languages and families ("convergence"). Because change proceeds in minimal discrete steps, the transition between stable optima necessarily passes through "mixed" disharmonic grammars with inconsistent headedness.

References

ABOH, ENOCH. 2004. *The morphosyntax of complement-head sequences.* Oxford: OUP.

ALLEN, CYNTHIA L. 1995. *Case marking and reanalysis: Grammatical relations from Old to Early Modern Linguistics.* Oxford: Clarendon Press.

ANDERSON, GREGORY D. S. (2007). *The Munda Verb: Typological Perspectives.* Berlin: De Gruyter Mouton.

ANDERSEN, HENNING. 1973. Abductive and deductive change. *Language* 49:567–595.

ANDERSEN, HENNING. 2001. Actualization and the (uni)directionality of change. In Andersen, Henning (ed.) *Actualization: Linguistic change in progress,* 21–57. Amsterdam: Benjamins.

ANDRASON, ALEXANDER, JOHN SULLIVAN & JUSTYNA OLKO. 2023. Language simplification in endangered languages? – Inflectional categories of nouns and verbs in Eastern Huasteca Nahuatl and Wymysorys. *Studies in Language* 47(1). 190-241.

ANTTILA, ARTO. 1997. Deriving variation from grammar. In F. Hinskens, R. van Hou, & L. Wetzels (Eds.), *Variation, change and phonological theory*, pp. 35–68. Amsterdam: John Benjamins.

ANTTILA, ARTO. 2007. Variation and optionality. In P. de Lacy (Ed.), *The Cambridge handbook of phonology* 519–536. Cambridge: Cambridge University Press.

ARKADIEV, P. & F. GARDANI (eds.) 2020. *The Complexities of Morphology*. Oxford: OUP. https://doi.org/10.1093/oso/9780198861287.003.0004

ARREGI, KARLOS & ANDREW NEVINS. 2012. *Morphotactics: Basque auxiliaries and the structure of spellout.* Dordrecht: Springer.

AXEL, KATRIN. 2007. *Studies on Old High German syntax. Left sentence periphery, verb placement, and verb-second.* Amsterdam/Philadelphia: Benjamins.

AXEL-TOBER, K. 2017. The development of the declarative complementizer in German. *Language* 93(2), e29–e65. https://doi.org/10.1353/lan.2017.0030

BAERMAN, MATTHEW, DUNSTAN BROWN & GREVILLE CORBETT. 2015. What is morphological complexity? In M. Baerman, D. Brown & G. Corbett, *Understanding and measuring morphological complexity*, 3–10. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780198723769.001.0001

BAILEY, LAURA. 2015. Word order and the syntax of question particles. In Sylvie Hancil, Alexander Haselow, Margje Post (eds.) *Final particles*. Berlin: De Gruyter Mouton.

BANE, MAX & JASON RIGGLE. 2008. Three correlates of the typological frequency of quantity-insensitive stress systems. In *Proceedings of the tenth meeting of the ACL special interest group on computational morphology and phonology,* 29–38. Columbus: Association for Computational Linguistics.

BANE, MAX. 2011. Deriving the structure of variation from the structure of non-variation in the English Dative alternation. In Mary Byram Washburn et al. (eds.), Proceedings of the 28th West Coast Conference on Formal Linguistics, 42-50. Somerville, MA: Cascadilla Proceedings Project.

BANE, MAX, SAM BOWMAN, & JASON RIGGLE. 2011. *pyphon 1.5. software package.* `https://code.google.com/archive/p/clml/`.

BATTYE, ADRIAN & IAN ROBERTS. 1995. (Eds.) *Clause structure and language change.* (Oxford studies in comparative syntax.) New York/Oxford: Oxford University Press.

BAUER, BRIGITTE. 1995. *The emergence and development of SVO patterning in Latin and French: Diachronic and psycholinguistic perspectives.* New York: Oxford University Press.

BAYER, JOSEF. 1999. Final complementizers in hybrid languages. *Journal of Linguistics* 35(2): 233-271.

BEARTH, THOMAS. 1995. Nominal periphrases and the origin of the predicative marker in Mande languages – an alternative view. *Afrikanistische Arbeitspapiere* 41: 89–117.

BERGSLAND, KNUT. 1997. *Aleut Grammar.* Fairbanks: Alaska Native Language Center.

BERWICK, ROBERT. 1985. *The acquisition of syntactic knowledge.* Cambridge, MA: MIT Press.

BIBERAUER, THERESA & IAN ROBERTS. 2009. Subjects, tense and verb-movement. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.) *Parametric Variation: Null Subjects in Minimalist Theory*, Ch. 7, 263-302. `https://doi.org/10.1017/CBO9780511770784.008`. Cambridge: CUP.

BIBERAUER, THERESA, ANDERS HOLMBERG & IAN ROBERTS. 2014. A syntactic universal and its consequences. *Linguistic Inquiry* 45:169-225.

BIBERAUER, THERESA & MICHELLE SHEEHAN. 2013. *Theoretical approaches to disharmonic word orders.* Oxford: Oxford University Press.

BIERWISCH, MANFRED & ROBERT SCHREUDER. 1992. From concepts to lexical items. *Cognition* 42: 23-60.

BLUST, ROBERT. 2003. Three notes on early Austronesian morphology. *Oceanic Linguistics* 42(2): 438-78. `https://www.jstor.org/stable/3623246?seq=1`

BLUST, ROBERT. 2024. Structural characteristics of Austronesian languages. In *Austronesian languages. Encyclopedia Britannica.* `https://www.britannica.com/topic/Austronesian-languages/Morphology-and-`

BODOMO, ADAMS. 1997. *The structure of Dagaare.* Stanford, CA: CSLI Publications.

BÖGEL, TINA. 2014. Degema and the string interface, In Miriam Butt and Tracy Holloway King (eds.) *Proceedings of the LFG14 Conference.* Stanford, CA.: CSLI Publications.

BORER, HAGIT. 2005. *Structuring sense*. Oxford: OUP.

CATASSO, NICHOLAS, CHIARA BASTIANI & MARCO CONIGLIO. 2022. *Language Change at the Interfaces.* Amsterdam: Benjamins.

CHAN, BRIAN HOK-SHING. 2013. Sentence-final particles, complementizers, antisymmetry, and the Final-over-Final Constraint. In Theresa Biberauer & Michelle Sheehan (eds.): *Theoretical approaches to disharmonic word orders.* Oxford: Oxford University Press, 407–444. https://doi.org/10.1017/CBO9780511770784.008

CHOMSKY, NOAM & MORRIS HALLE. 1968. *The sound pattern of English.* Bew York: Harper & Row.

CHOMSKY, NOAM. 1995. *The Minimalist Program.* Cambridge, Mass.: MIT Press

CLARK, ROBIN & IAN ROBERTS. 1993. A computational approach to language learnability and language change. *Linguistic Inquiry* 24: 299-345.

CLAUDI, ULRIKE 1993. *Die Stellung von Verb und Objekt in Niger-Kongo Sprachen: Ein Beitrag zur Rekonstruktion historischer Syntax.* Köln: Institut für Afrikanistik.

CLAUDI, ULRIKE 1994. Word order change as category change: the Mande case. In W. Pagliuca (ed.) *Perspectives on Grammaticalization,* 191–232 Amsterdam: Benjamins.

CONDORAVDI, CLEO & PAUL KIPARSKY. 2002. Clitics and clause structure. *Journal of Greek Linguistics* 2(1), 1-39. https://doi.org/10.1075/jgl.2.02con

CREISSELS, D. 2005. SOVX constituent order and constituent order alternations in West African languages. *Annual Meeting of the Berkeley Linguistics Society* 37-52.

CREISSELS, D. 2018. Current issues in African morphosyntactic typology. In Tom Güldemann (ed.) *The languages and linguistics of Africa,* 714-821. Berlin/Boston: De Gruyter Mouton.

CULBERTSON, JENNIFER, PAUL SMOLENSKY, GÉRALDINE LEGENDRE. 2012. Learning biases predict a word order universal. *Cognition* 122: 306–329.

DAHL, ÖSTEN. 2004. *The growth and maintenance of linguistic complexity.* Amsterdam: Benjamins.

DE SMET, HENDRIK. 2009. Analysing reanalysis. *Lingua* 119: 1728–1755.

DE SMET, HENDRIK & MARKEY, M.A. 2021. The spark or the fuel? On the role of ambiguity in language change. *Journal of Historical Syntax*, 5(32-39), pp.1-24.

DEUTSCHER, GUY. 2002. On the misuse of the notion of 'abduction' in linguistics. *Journal of Linguistics* 38(3): 469-485. https://doi.org/10.1017/S002222670200169X

DIXON, R.M.W. 1984. *Ergativity.* Cambridge: CUP.

DRESHER, B.E. 1999. Charting the learning path: Cues to parameter setting. *Linguistic inquiry* 30(1):27-67.

DRYER, MATTHEW S. 1992. The Greenbergian word order correlations. *Language* 68: 81-138.

DRYER, MATTHEW S. 2003. Significant and non-significant implicational universals. *Linguistic Typology* 7(1): 108-128. https://doi.org/10.1515/lity.2003.007

DRYER, MATTHEW S. 2011. "Order of adposition and noun phrase". In: Dryer, Matthew S. & Haspelmath, Martin (eds.). The World Atlas of Language Structures Online. Munich: Max Planck Digital Library. http://wals.info.

ENGELHARDT, PAUL E. & LUNA FILIPOVIĆ & JOHN A. HAWKINS. 2024. Prediction in SVO and SOV languages: processing and typological considerations. *Linguistics* 62(2): 349–383.

EYTHÓRSSON, THÓRHALLUR. (1996). Functional categories, cliticization, and verb movement in the early Germanic languages. In Thráinsson et al. (1996: 109–39). Galbraith, Daniel. 2023. *Optimal Linking Grammar: A Theory of Morphosyntax.* Cambridge, Cambridge University Press.

FAARLUND, JAN TERJE. 2021. A cluster of changes: Norwegian word order. In Meklenborg, Lohndal & Østby (eds.) *Syntax, semantics and acquisition: In honor of Hans Petter Helland, Oslo Studies in Language* 12(1): 1-20. http://www.journals.uio.no/osla.

FRAJZYNGIER, ZYGMUNT. 1993. *A Grammar of Mupun.* Berlin: D. Reimer Verlag.

FUSS, ERIC & CAROLA TRIPS. 2003. Variation and change in Old and Middle English — on the validity of the Double Base Hypothesis. *The Journal of Comparative Germanic Linguistics* 4. 171–224.

GALBRAITH, DANIEL. 2023. *Optimal Linking Grammar: A Theory of Morphosyntax*. Cambridge: Cambridge University Press.

GARRETT, ANDREW & KEITH JOHNSON. 2013. Phonetic bias in sound change. In Alan C. L. Yu (ed.) *Origins of sound change: Approaches to phonologization*, 51-97. Oxford: OUP.

GAVRILETS, SERGEY. 2004. *Fitness landscapes and the origin of species.* Princeton: Princeton University Press.

GELDEREN, ELLY VAN. 2004. Grammaticalization as economy (Linguistik aktuell/Linguistics today 71). Amsterdam: John Benjamins.

GELDEREN, ELLY VAN. 2009. Renewal in the left periphery: Economy and the complementiser layer. *Transactions of the Philological Society* 107(2): 131–195.

GELDEREN, ELLY VAN. 2024. *The linguistic cycle: Language change and the language faculty.* Oxford: OUP.

GELL-MANN, MURRAY & SETH LLOYD. 2004. Effective Complexity. *Santa Fe Institute Working Paper* 2003-12-068.

GELL-MANN, MURRAY & MERRITT RUHLEN. 2011. The origin and evolution of word order. *PNAS* 108(42): 17290–17295. https://doi.org/10.1073/pnas.1113716108

GENSLER, ORIN D. 1994. On Reconstructing the Syntagm S-Aux-O-V-Other to Proto-Niger-Congo. *Proceedings of the 20th Annual Meeting of the Berkeley Linguistics Society, Special Session on Historical Issues in African Linguistics*, 1-20. Berkeley, CA.

GENSLER, ORIN D. 1997. Review article. *Journal of African Languages and Linguistics*, 18(1): 57-94. https://doi.org/10.1515/jall.1997.18.1.57

GENSLER, ORIN D. & TOM GÜLDEMANN. 2003. S-Aux-O-V-Other in Africa: Typological and areal perspective.
https://www.iaaw.hu-berlin.de/de/region/afrika/afrika/linguistik/mitarb

GIVÓN, TALMY. 1971. *On the verbal origin of the Bantu verb suffixes.Studies in African Linguistics* 2(2): 145-163.

GOLDIN-MEADOW, SUSAN, & C. MYLANDER. 1998. Spontaneous sign systems created by deaf children in two cultures. *Nature* 391: 279–281.

GOLDIN-MEADOW, SUSAN, WING CHEE SO, ASLI ÖZYÜREK & CAROLYN MYLANDER. 2008. The natural order of events: How speakers of different languages represent events nonverbally. *PNAS* 105(27): 9163-9168. https://doi.org/10.1073/pnas.0710060105

GOLDSMITH, JOHN. 2011. The evaluation metric in generative grammar.
https://www.academia.edu/108890414/The_evaluation_metric_in_generative_

GOLDSTEIN, D.M. 2019. Language Change and Linguistic Theory: The case of archaic Indo-European conjunction. *Transactions of the Philological Society* 117(1): 1–34.

GOMEZ RICARDO & KOLDO SAINZ. On the origin of the finite forms. In Hualde, Ignacio José, Joseba A. Lacarra, & R.L. Trask (eds.) *Towards a history of the Basque Language*, 235-274. Amsterdam: Benjamins.

GÜLDEMANN, TOM. 2018. Historical linguistics and genealogical language classification in Africa. In Tom Güldemann (ed.) *The languages and linguistics of Africa,* 58-444. Berlin/Boston: De Gruyter Mouton.

HAEBERLI, ERIC. 2000. Adjuncts and the syntax of subjects in Old and Middle English. In Pintzuk, S., Tsoulas, G., Warner, A. (Eds.), *Diachronic Syntax.* Models and Mechanisms. Oxford University Press, Oxford, pp. 109–131.

HAEBERLI, ERIC. 2005. Clause type asymmetries in Old English and the syntax of verb movement. In: Batllori, M., Roca, F. (eds.) *Grammaticalization and Parametric Variation*. Oxford University Press, Oxford, pp. 267–283.

HAEBERLI, ERIC & S. PINTZUK. 2012. In Revisiting Verb (Projection) Raising. In D. Jonas, J. Whitman & A. Garrett, (Eds.), *Grammatical Change: Origins, Nature, Outcomes*, p. 219-238. Oxford : Oxford University Press.

HAHN, MICHAEL & YANG XU. 2022. Crosslinguistic word order variation reflects evolutionary pressures of dependency and information locality. *PNAS* 119 (24) e2122604119. https://doi.org/10.1073/pnas.2122604119

HAIDER, HUBERT. 1997. Typological implications of a directionality constraint on projections. In A. Alexiadou and T. A. Hall (eds.) *Studies on Universal Grammar and typological variation*. Amsterdam: Benjamins. 17–33.

HAIDER, HUBERT. 2000. OV is more basic that VO. In Peter Svenonius (ed.) *The Derivation of VO and OV.* Amsterdam: Benjamins.

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

HAIDER, HUBERT. 2021. Grammar change – A case of Darwinian cognitive evolution. *Evolutionary Linguistic Theory* 3(1). 6–55.

HAIDER, HUBERT. 2023. SVO Attractor in the declarative-to-procedural shift in grammar evolution. *Acta Linguistica Academica* 70(2): 195–218. 10.1556/2062.2023.00642

HAIG, GEOFFREY. 2006. Turkish influence on Kurmanji: Evidence from the Tunceli dialect. In: Johanson, Lars & Christiane Bulut (eds.) *Turkic-Iranian contact areas. Historical and linguistic aspects,*, 283-299. Wiesbaden: Harrassowitz.

HAIG, GEOFFREY. 2008. *Alignment change in Iranian languages. A Construction Grammar approach.* Berlin: Mouton.

HAJNAL, IVO. 2016. Induktive versus abduktive Rekonstruktion. *Indogermanische Forschungen* 121: 435-453.

HAIG, GEOFFREY. 2018. The grammaticalization of object pronouns: why differential object indexing is an attractor state. *Linguistics* 56(4):781–818

HALE, MARK. 1998. Diachronic syntax. *Syntax* 1: 1–18.

HALE, MARK, MADELYN KISSOCK, & CHARLES REISS. 2006. Microvariation, variation, and the features of universal grammar. *Lingua* 117: 645-65.

HALE, MARK, MADELYN KISSOCK & CHARLES REISS. 2015. An I-Language approach to phonologization and lexification. In Honeybone, Patrick & J. Salmons (ed.) Handbook of Historical

Phonology, 337–358. Oxford: OUP.
https://doi.org/10.1093/oxfordhb/9780199232819.013.027

HALL, MATTHEW L., RACHEL I. MAYBERRY, & VICTOR S. FERREIRA. 2013. Cognitive constraints on constituent order: Evidence from elicited pantomime. *Cognition.* 129(1): 1–17. 10.1016/j.cognition.2013.05.004

HALPERN, AARON AND JOSEP M. FONTANA. 1994. X^0 and X^{max} clitics. *WCCFL* 12:251–266.

HARRIS, ALICE C. 2003. Cross-linguistic perspectives on syntactic change. In Brian D. Joseph & Richard D. Janda (eds.), *The handbook of historical linguistics*, 529–551. Oxford: Blackwell.

HASPELMATH, MARTIN. 1998. Does grammaticalization need reanalysis? *Studies in Language* 22(2): 315-351. DOI10.1075/sl.22.2.03has

HARRIS, A. & LYLE CAMPBELL. 1995. *Historical Syntax in Cross-linguistic Perspective.* New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511620553

HAWKINS, J. A. 1990. A parsing theory of word order universals. *Linguistic Inquiry* 21: 223–61.

HAWKINS, J. A. 1994. *A performance theory of order and constituency.* Cambridge: CUP.

HAWKINS, J. A. 2004. *Efficiency and Complexity in Grammars.* Oxford: OUP.

HEATH, JEFFREY. 2008. *Jamsay grammar.* Berlin, New York: Mouton.

HEIN, JOHANNES & ANDREW MURPHY. 2022. VP-Nominalization and the Final-over-Final Condition. *Linguistic Inquiry* 53(2): 337–370.
https://doi.org/10.1162/ling_a_00407

HEINE, BERND. 1976. *A typology of African languages based on the order of meaningful elements.* Berlin: Dietrich Reimer.

HODGE, CARLETON. 1970. The linguistic cycle. *Language Sciences* 13: 1–7.

HOOK, PETER E., AND O. KOUL. 1996. Kashmiri: An SOV language with V-2 word order 1. In V. S. Lakshmi and A. Mukerjee (eds.) *Word Order in Indian Languages.* Hyderabad: CASL-Osmania University: 95-106.

HOOK, PETER EDWIN & OMKAR NATH KOUL. n.d.a. Relaxing the V-2 Constraint on Kashmiri Word Order. https://linguistics.uok.edu.in/Files/f6ec3740-422d-4ac1-9f52-d

HOOK, PETER EDWIN & OMKAR NATH KOUL. n.d.b. Relaxing the V-2 Constraint on Kashmiri Word Order. https://d1wqtxts1xzle7.cloudfront.net/34132859/V2_and_Verb_Comp

HUBER, MAGNUS 2013. Order of adposition and noun phrase. In Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds.) 2013. *The Atlas of Pidgin and Creole Language Structures.* Oxford: Oxford University Press.

HYMAN, LARRY. 1975. On the change from SOV to SVO: Evidence from Niger-Congo. In Charles N. Li (ed.) *Word order and word order change.* Austin: University of Texas Press.

JÄGER, GERHARD. 2019. Language change as a random walk in vector space. https://profgerhard.de/

JÄGER, GERHARD. 2025. Computational Typology. https://arxiv.org/pdf/2504.15642

JANDA, RICHARD AND BRIAN JOSEPH. In press. This volume.

JOSEPH, BRIAN. 2006. On projecting variation back into a proto-language, with particular attention to Germanic evidence. In T. Cravens (ed.) *Variation and Reconstruction*. Amsterdam: Benjamins.

JOSEPH, BRIAN. 2013. Demystifying drift, A variationist account. In Martine Robbeets & Hubert Cuyckens (eds.), *Shared Grammaticalization*, 43-66.

KANDYBOWICZ, JASON. 2008. *The grammar of repetition.* Amsterdam: Benjamins.

KARI, ETHELBERT E. 1997. *Degema*. LINCOM Europa.

KARI, ETHELBERT E. 2004. *A Reference Grammar of Degema.* Köln: Köppe.

KASTENHOLZ, RAIMUND. 2003. Auxiliaries, grammaticalization, and word order in Mande. *Journal of African Languages and Linguistics* 24: 31–53.

KAYNE, RICHARD S. 1994. *The Antisymmetry of Syntax.* Cambridge, MA: MIT Press.

KAYNE, RICHARD S. 2005. 'Antisymmetry and Japanese', In Richard S. Kayne, *Movement and Silence*. New York: Oxford Academic.
https://doi-org.stanford.idm.oclc.org/10.1093/acprof:oso/9780195179163.

KEMENADE, ANS VAN. 2011. Secondary negation and Information Structure organization in the History of English. In Pierre Larrivee & Richard Ingham (eds.), The evolution of negation: Beyond the Jespersen Cycle, 77–114. Berlin: Mouton de Gruyter. DOI: *https://doi.org/10.1515/9783110238617.77*

KEMENADE, ANS VAN & BETTELOU LOS. 2006. Discourse adverbs and clausal syn- tax in Old and Middle English. In Ans van Kemenade & Bettelou Los (eds.), *The handbook of the history of English*, 224–248. London: Blackwell.
https://doi.org/10.1002/9780470757048.ch10

KEMENADE, ANS VAN, TANJA MILICEV & HARALD BAAYEN. 2008. The balance between discourse and syntax in Old and Middle English. In Maurizio Gotti, Marina Dossena & Richard Dury (eds.), *English Historical Linguistics 2006, Vol I: Syntax and Morphology*, 3–22. Amsterdam: Benjamins.

KEMENADE, ANS VAN. 2009. Discourse relations and word order change. In Roland Hinterhölzl & Svetlana Petrova (eds.), Information structure and language change, 91–120. Berlin: Mouton de Gruyter.

KEMENADE, ANS VAN & TANJA MILICEV. 2012. Syntax and discourse in Old and Middle English word order. In Dianne Jonas, John Whitman & Andrew Garrett (eds.), *Gram- matical change: Origins, nature, outcomes*. 239–254. Oxford: Oxford University Press.

VAN KEMENADE, ANS VAN & LINKS, M. 2020. Discourse particles in early English: Clause structure, pragmatics and discourse management. *Glossa* 5(1):3.
https://doi:https://doi.org/10.5334/gjgl.1020

KEISER, STEVEN HARTMAN. 2009. When 'speech islands' aren't islands: Parallel independent development, drift, and minimal levels of contact for diffusion. *Diachronica* 26(1): 1-35.
https://doi.org/10.1075/dia.26.1.01kei

KERR, ELIZABETH J. 2024. On OV and VO at the Bantu /Bantoid borderlands. In Yaqian Hong et al. (eds.) ACAL in SoCAL: Selected papers from the 53rd Annual Conference on African Linguistics, 297–345. Berlin: Language Science Press.
https://doi.org/10.5281/zenodo.11210464

KIMMELMAN, VADIM. 2022. Argument structure in sign languages. *Annual Review of Linguistics.* 8: 19–38.
https://doi.org/10.1146/annurev-linguistics-031220-122519

KIPARSKY, PAUL. 1993. An OT approach to phonological variation. Rutgers Optimality Workshop, 1993. https://web.stanford.edu/~kiparsky/Papers/nwave94.pdf.

KIPARSKY, PAUL. 1995. Indo-European origins of Germanic syntax. In Adrian Battye & Ian Roberts (eds.) *Clause structure and language change.* (Oxford studies in comparative syntax), 200–226. New York/Oxford: Oxford University Press.

KIPARSKY, PAUL. 1996. The Shift to Head-Initial VP in Germanic. In H. Thrainsson, J. Peter, and S. Epstein (eds.), *Comparative Germanic Syntax.* Kluwer, 1996.

KIPARSKY, PAUL. 1997. The Rise of Positional Licensing. In Ans van Kemenade and Nigel Vincent (eds.), *Parameters of morphosyntactic change.* Oxford: OUP.

KIPARSKY, PAUL. 2002. Disjoint Reference and the Typology of Pronouns. In Ingrid Kaufmann and Barbara Stiebels (eds.), *More than Words. Studia Grammatica* 53, 179-226. Berlin: Akademie Verlag.

KIPARSKY, PAUL. 2012. Greek anaphora in cross-linguistic perspective. *Journal of Greek Linguistics* 12: 84-117. `www.ingentaconnect.com/content/brill/jgl`

KIRK, A. (2012). *Word order and information structure in New Testament Greek.* Ph.D. dissertation, Leiden University. LOT series. `https://hdl.handle.net/1887/20157`.

KISSEBERTH, CHARLES. 1970. On the functional unity of phonological rules. *Linguistic Inquiry* 1.291-306.

KOENEMAN, OLAF, & ZEIJLSTRA, HEDDE. 2014: One law for the rich and another for the poor. The Rich Agreement Hypothesis rehabilitated. *Linguistic Inquiry* 45(4): 1–45.

KOOPMAN, HILDA. 1984. *The Syntax of Verbs: From Verb Movement Rules in the Kru Languages to Universal Grammar.* Dordrecht: Foris Publications.

KOUL, OMKAR N. & KASHI WALI. 2006. *Modern Kashmiri Grammar.* Springfield: Dunwoody Press.

KROCH, ANTHONY. 1989. Reflexes of grammar in patterns of language change. *Language variation and change*, 1: 199-244.

KROCH, ANTHONY. 2001. Syntactic Change. In M. Baltin and C. Collins (eds.) *The Handbook of Contemporary Syntactic Theory,* 699–729. Malden, MA: Blackwell.

LABOV, WILLIAM. 1994. *Principles of language change*, Vol. 1. Oxford: Blackwell.

LABOV, WILLIAM. 2001. *Principles of linguistic change*, Vol. 2: Social factors. Oxford: Blackwell.

LABOV, WILLIAM. 2007. 'Transmission and diffusion', *Language* 83: 344–87.

LAKA, ITZIAR. 1990. *Negation in syntax: On the nature of functional categories and projections.* PhD Dissertation, MIT.

LAKOFF, ROBIN. 1972. Another look at drift. In Robert P. Stockwell & Ronald Macaulay eds. *Linguistic change and generative theory.* Bloomington: Indiana University Press. 172-98.

LANGACKER, RONALD W. 1977. Syntactic reanalysis. In Charles N. Li (ed.), *Mechanisms of syntactic change*, 57–139. Austin: University of Texas Press.

LAVIDAS, N. 2015. How does a basic word order become ungrammatical? SOV from Classical to Koiné Greek. *Studies in Greek Linguistics* 35: 323-335.

LEDGEWAY, ADAM. 2012. *From Latin to Romance: Morphosyntactic typology and change.* Oxford: Oxford University Press.

LEHMANN, CHRISTIAN. 1995. *Thoughts on Grammaticalization.* LINCOM Studies in Theoretical Linguistics 1. München: LINCOM Europa.

LI, CHARLES N. 1977. (ed.) *Mechanisms of Syntactic Change.* Austin, TX: University of Texas Press.

LIGHTFOOT, DAVID. 1979. *Principles of Diachronic Syntax.* Cambridge: CUP.

LIGHTFOOT, DAVID. 1991. *How to set parameters: Arguments from language acquisition.* Cambridge: CUP.

LIGHTFOOT, DAVID. 1999. The development of language: Acquisition, change, and evolution. Oxford: Wiley-Blackwell.

LIGHTFOOT, DAVID. 2003. Grammaticalisation: Cause or effect? In Hickey, Raymond (ed.) *Motives for language change*. Cambridge: CUP.

LIGHTFOOT, DAVID & MARIT WESTERGAARD. 2007. Language Acquisition and Language Change: Inter-relationships. *Language and Linguistics Compass* 1(5): 396–415. 10.1111/j.1749-818x.2007.00023.x

LIGHTFOOT, DAVID. 2010. Language acquisition and language change. *Wiley Interdisciplinary Reviews: Cognitive Science* 1(5)5: 677-684.

MADARIAGA, NEREA. 2017. Reanalysis. *The Cambridge handbook of historical syntax*, 70-91.

MANZINI, M. RITA & KENNETH WEXLER. 1987. Parameters, Binding Theory, and Learnability. *Linguistic Inquiry*, Summer 18(3): 413-444.

MARTIN, ALEXANDER & JENNIFER CULBERTSON. 2020. Revisiting the Suffixing Preference: Native-Language Affixation Patterns Influence Perception of Sequences. *Psychological Science* 31(9). https://doi.org/10.1177/0956797620931108.

MAURITS, LUKE AND THOMAS L. GRIFFITHS. 2014. Tracing the roots of syntax with Bayesian phylogenetics. *Proceedings of the National Academy of Sciences* 111: 13576–13581.

MCCLOSKEY, JAMES. 1996. On the scope of verb movement in Irish. *Natural Lan- guage and Linguistic Theory* 14(1): 47–104.

MCCLOSKEY, JAMES. 2017. Ellipsis, polarity, and the cartography of verb-initial orders in Irish. In Enoch Aboh, Eric Haeberli, Genoveva Puskás & Manuela Schönenberger (eds.), *Elements of comparative syntax: Theory and description.* (Studies in Generative Grammar 127), 99–151. Berlin/Boston: De Gruyter Mouton.

MCCLOSKEY, JAMES. 2025. Polarity sensitivity and fragments in Irish. In Carnie, Andrew, Diane Ohala, Dee Hunter, Samantha Prins, Michael Hammond & Luis Irizarry (eds.). *Foundational approaches to Celtic linguistics* 5-61. (Current Issues in Celtic Linguistics 1). Berlin: Language Science Press.

MCPHERSON, LAURA. 2013. *A Grammar of Tommo So*. Berlin: De Gruyter Mouton.

MCWHORTER, JOHN. 2007. *Language interrupted: Signs of non-native acquisition in standard language grammars.* Oxford: Oxford University Press. https://academic-oup-com.stanford.idm.oclc.org/book/36358/chapter/31988

MEILLET, ANTOINE. 1918. Convergence des développements linguistiques. *Revue philosophique* 85: 97–110.

MEIR, I., SANDLER, W., PADDEN, P., AND ARONOFF, M. 2010. Emerging sign languages, in M. Marschark and P. E. Spencer (eds.) *Oxford Handbook of Deaf Studies, Language, and Education, Vol. 2* 267–280. Oxford: OUP.

MIESTAMO, MATTI, SINNEMÄKI, KAIUS & KARLSSON, FRED (eds.) 2008. *Language complexity: Typology, contact, change.* Amsterdam: John Benjamins.

MITCHELL, MELANIE. 2009. *Complexity: A guided tour.* Oxford: Oxford University Press.

MOUS, MAARTEN. 2022. Tone and the verbal systems of Nyokon. *Journal of West African Languages*, 49(2): 1-33.

MOUS, MAARTEN. 2014. Full Object-Verb Order in the Mbam languages of Cameroon. *ZAS Papers in Linguistics* 57: 72-94.

MYERS, SCOTT. 1990. *Tone and the Structure of Words in Shona.* Outstanding dissertations in linguistics. Garland.

NAPOLI, D.J. AND SUTTON-SPENCE, R. 2014. Order of the major constituents in sign languages: Implications for all language. *Frontiers in psychology* 5:1–18, 2014.)

NETTLE, DANIEL. 1998. *The Fyem language of northern Nigeria.* Lincom Europa.

NEWMEYER, FREDERICK J. 2000. On the reconstruction of 'Proto-World' word order. In Chris Knight, Michael Studdert-Kennedy & James Hurford (eds.) *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form.* Cambridge: CUP.

NIKITINA, TATIANA. 2008. *The mixing of syntactic properties and language change.* Stanford: Stanford University dissertation.

NURSE, DEREK, SARAH ROSE AND JOHN HEWSON. 2016. *Tense and Aspect in Niger-Congo* Tervuren: Royal Museum for Central Africa.

PETZELL, ERIK MAGNUSSON. 2011. OV-ordföljd i svenskans historia. *Arkiv för nordisk filologi* 126. S. 141–191.

PETZELL, ERIK MAGNUSSON. 2016. Low German with a Swedish twist – Contact induced word order transfer in the 15th century. *Ampersand* 3: 143–150.

PIÑÓN, CHRISTOPHER. 1993. SigmaP and Hungarian. *WCCFL* 11:388–404.

PINTZUK, SUSAN. 1991. *Phrase Structures in Competition: Variation and Change in Old English Word Order.* Ph.D. dissertation. Philadelphia: University of Pennsylvania.

PINTZUK, SUSAN & TAYLOR, ANN. 2006. The loss of OV order in the history of English. I: Ans van Kemenade & Bettelou Los (eds.): *Blackwell Handbook of the History of English,* 249–278. Oxford: Blackwell.

PLATZACK, CHRISTER. 1995. The loss of verb second in English and French. I: Adrian Battye & Ian Roberts (eds.): *Clause structure and language change.* (Oxford studies in comparative syntax.) New York/Oxford: Oxford University Press. S. 200–226.

POLINSKY, MARIA, MICHAEL T. PUTNAM & JOE SALMONS. Forthcoming. Linguistic complexity in heritage languages: An introduction. In Polinsky, Maria, Michael T. Putnam (eds.). *Formal approaches to complexity in heritage language grammars.* Berlin: Language Science Press.

POPJES, JACK & JO POPJES. 1986. Canela-Krahô, In Desmond C. Derbyshire & Geoffrey K. Pullum. *Handbook of Amazonian Languages,* 129-199. Berlin: De Gruyter Mouton.

PRINCE, ALAN S. & PAUL SMOLENSKY. 2002. *Constraint Interaction in Generative Grammar*. DOI:https://doi.org/10.7282/T34M92MV

RIGGLE, JASON. 2010. Sampling rankings. https://https://roa.rutgers.edu/files/1075-05

ROBERTS, IAN. 2007. *Diachronic syntax.* Oxford: Oxford University Press.

ROBERTS, IAN. 2017. Harmony, symmetry and dominance in word order universals. I: Michelle Sheehan, Theresa Biberauer, Ian Roberts & Anders Holmberg: *The Final-over-Final Condition. A syntactic universal.* (Linguistic Inquiry Monographs 76.) Cambridge, MA: The MIT Press. S. 27–42.

LYNCH, JOHN, MALCOLM D. ROSS, & TERRY CROWLEY (eds.) 2002. *The Oceanic languages.* Abingdon: Routledge.

ROUSSOU, ANNA & IANTHI-MARIA TSIMPLI. 2006. On Greek VSO again! *Journal of Linguistics* 42(2) 317-354 https://doi.org/10.1017/S0022226706003914

SALVI, GIAMPAOLO. 2012. On the nature of the V2 system of medieval Romance. In Laura Brugè, Anna Cardinaletti, Giuliana Giusti, Nicola Munaro & Cecilia Poletto (eds.), *The cartography of syntactic structures.* vol. 7, functional heads, 103–111. Oxford; New York: Oxford University Press.

SAMPSON, GEOFFREY, DAVID GIL & PETER TRUDGILL (EDS.) 2009. *Language complexity as an evolving variable.* Oxford: OUP.

SANDLER, WENDY, IRIT MEIR, CAROL PADDEN, & MARK ARONOFF. 2005. The emergence of grammar: Systematic structure in a new language. *PNAS* 102(7): 2661-2665.

SANGFELT, ADRIAN. 2019. *Syntaktiska strukturer i tiden. OV- och bisatsledföljd i svenskans historia.* (Syntactic Structures in Time. OV and Subordinate Clause Word Order in the History of Swedish). Studia philologiae Scandinavicae Upsaliensia 20. 332 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN978-91-513-0792-3.

SAPIR, EDWARD. 1921. *Language: An introduction to the study of speech.* New York: Harcourt Brace & World.

SENGHAS, ANN, SOTARO KITA & ASLI ÖZYÜREK. 2004. Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science* 305. 1779–1782.

SERŽANT, ILJA A. & GEORGE MOROZ. 2022. Universal attractors in language evolution provide evidence for the kinds of efficiency pressures involved. *Humanities and Social Sciences Communications* 9:58. https://doi.org/10.1057/s41599-022-01072-0

SERŽANT, ILJA A. & DARIYA RAFIYENKO. 2020. Diachronic evidence against source-oriented explanation in typology: Evolution of prepositional phrases in Ancient Greek. *Language Dynamics and Change* 11(2): 167-210.

SHEEHAN, MICHELLE, 2017. FOFC and Linearization. In Sheehan et. al. (2017).

SHEEHAN, MICHELLE, THERESA BIBERAUER, IAN ROBERTS & ANDERS HOLMBERG. 2017. *The Final-over-Final Condition. A syntactic universal.* (Linguistic Inquiry monographs 76.) Cambridge, MA: The MIT Press.

SHLONSKY, U. 1997. *Clause structure and word order in Hebrew and Arabic: An essay in comparative Semitic syntax.* Oxford: OUP.

SIEWIERSKA, ANNA & DIK BAKKER. 2012. Case and alternative strategies: Word order and agreement marking. Andrej L. Malchukov & Andrew Spencer (eds.) *The Oxford Handbook of Case*, Ch. 19. Oxford: OUP.

STASSEN, LEON. 2001. Coordination. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, &; Wolfgang Raible. *Language Typology and Language Universals*, 2: 1105-111. De Gruyter Mouton.

STERNEFELD, WOLFGANG. 2006. *Syntax. Eine morphologisch motivierte generative Beschreibung des Deutschen.* Tübingen: Stauffenburg.

STERNEFELD, WOLFGANG. 2006. *Syntax. Eine morphologisch motivierte generative Beschreibung des Deutschen.* Tübingen: Stauffenburg.

STORTO, L. 1999. *Aspects of Karitiana Grammar.* PhD dissertation, MIT Linguistics Department.

TAYLOR, ANN. 2003. The distribution of object clitics in Koine Greek. In Mark R.V. Southern (ed.) *Indo-European perspectives*, 285–315. Washington, D.C.: Institute for the Study of Man.

TIMBERLAKE, A. 1977. Reanalysis and actualization in syntactic change, in C. N. Li (ed.), *Mechanisms of syntactic change.* Austin: University of Texas Press, 141–77.

TRASK, ROBERT L. 1977. Historical syntax and Basque verbal morphology: Two hypotheses. William E. Douglass, Robert Etulain, & William H. Jacobsen (ed.) *Anglo-American contributions to Basque studies. Essays in honor of Jon Bilbao,* 203-217. Reno: University of Nevada.

VENNEMANN, THEO. 1975. An exploration of drift. In by Charles N. Li (ed.) *Word order and word order change.* 270-305. Austin: University of Texas Press.

VERDÚ, JOSÉ FERNANDO DOMENE. 2010. *La lengua Vasca: originalidad y riqueza de una lengua diferente*. San Vicente: ECU.

VIKNER, STEN. 2001. *Verb movement variation in Germanic and optimality theory.* Habilitationsschrift, Universität Tübingen.

VIKNER, STEN. 2020. The placement of finite verbs. Michael T. Putnam & B. Richard Page (eds.), *The Cambridge Handbook of Germanic Linguistics*, Ch. 16. Cambridge: Cambridge University Press.

VILKUNA, MARIA. (1998) Word order in European Uralic. In A. Siewierska (ed.) *Constituent Order in the Languages of Europe.* Mouton de Gruyter, Berlin, 173–233. 10.1515/9783110812206.173

WARBURTON, IRENE et al. 1968. *Ewe basic course*. Bloomington: Indiana University.

WEINREICH, URIEL, WILLIAM LABOV & MARVIN HERZOG 1968. Empirical foundations for a theory of language change, in W. Lehmann and Y. Malkiel (eds.), *Directions for historical linguistics.* 95–189. Austin: University of Texas Press.

WEISS, HELMUT. 2019. Rebracketing (Gliederungsverschiebung) and the Early Merge Principle. *Diachronica* 36(4): 509-545.

WEISS, HELMUT. 2020. Where do complementizers come from and how did they come about? A re-evaluation of the parataxis-to-hypotaxis hypothesis. *Evolutionary Linguistic Theory* 2(1): 30–55. https://doi.org/10.1075/elt.00014.wei

WHITMAN, JOHN. 2012. Misparsing and syntactic reanalysis. In Kemenade, Ans van & Nynke de Haas (ed.) *Historical linguistics 2009*. Amsterdam: Benjamins.

WILLIAMSON, KAY. 1986. Niger-Congo: SVO or SOV? *Journal of West African Languages* 16: 5–14.

WORUFAH, EBINI-IPIRI RAPHAEL & EBITARE FABIAWARI OBIKUDO. 2023. A grammatical analysis of Biseni clause structure. *Periodica. Journal of Modern Philosophy, Social Sciences and Humanities* 22: 1-21.

YANG, CHARLES. 2000. Internal and external forces in language change. *Language Variation and Change* 12: 231–250.

YANG, CHARLES. 2016. *The price of linguistic productivity.* Cambridge, MA: MIT Press.