

Word order evolves at similar rates in main and subordinate clauses

Corpus-based evidence from Indo-European

Yingqi Jing¹, Paul Widmer² and Balthasar Bickel²

¹ Uppsala University | ² University of Zurich

In syntactic change, it remains an open issue whether word orders are more conservative or innovative in subordinate clauses compared with main clauses. Using 47 dependency-annotated corpora and Bayesian phylogenetic inference, we explore the evolution of S/V, V/O, and S/O orders across main and subordinate clauses in Indo-European. Our results reveal similar rates of change across clause types, with no evidence for any inherent conservatism of subordinate or main clauses. Our models also support evolutionary biases towards SV, VO, and SO orders, consistent with theories of dependency length minimization that favor verb-medial orders and with theories of a subject preference that favor SO orders. Finally, our results show that while the word order in the proto-language cannot be estimated with any reasonable degree of certainty, the early history of the family was dominated by a moderate preference for SVO orders, with substantial uncertainty between VO and OV orders in both main and subordinate clauses.

Keywords: word order evolution, rate of change, phylogenetic inference, Universal Dependencies, Indo-European

1. Introduction

Exploring how and why word orders change under different conditions is a key issue for the evolutionary dynamics and selective forces in syntax. One critical open issue concerns potential differences of word order change between main and subordinate clauses.

It has often been claimed that subordinate clauses are more conservative than main clauses (Givón 1979; Lightfoot 1982; Hock 1991). For example, verb-final order in German is commonly argued to have first changed to verb-second in

main clauses, while subordinate clauses still preserve the old order (Vennemann 1975). Similar claims have been made for the conservatism of subordinate clauses in Old English (Hock 1986), Kru (Hyman 1975), Biblical Hebrew (Givón 1979), and Mon (Jenny 2020). Such conservatism has been motivated by the claim that word order is synchronically more constrained in subordinate than in main clauses, i.e., that “more goes on upstairs than downstairs” according to Ross’ Penthouse Principle (Ross 1973). If this is the case, subordinate clauses show less variation and this in turn makes change less likely. The Penthouse Principle itself is arguably grounded in pragmatic, processing-based and/or learning-related factors: (1) subordinate clauses often contain presupposed background information, and are thus potentially less susceptible to word order variation in response to information structure demands (Hooper & Thompson 1973; Givón 1979); (2) subordinate clauses have been argued to require intermediate storage as entire chunks during processing, which is potentially easier if they follow stricter orderings than main clauses (Bybee 2002); (3) main clauses have been argued to be the prime drivers of innovation because they are arguably the main target of learning mechanisms (“degree-0 learnability”) (Lightfoot 1989, 2006).

Other research challenges the view that subordinate clauses are generally more conservative than main clauses. A case in point is the emergence of SV order (Adverb/Complementizer-S-V) in Middle English subordinate clauses after 1200, while the old verb-second order (Adverb-V-S) persisted in main clauses for several additional centuries up to 1400 (MacLeish 1969; Bean 1983; Stockwell & Minkova 1991). Moreover, in sharp contrast to all other extant Slavic languages, Upper Sorbian exhibits an unmarked (S)OV order (Stone 1993). This innovation most likely started out in subordinate clauses through sustained contact with German and spread to main clauses eventually. Yet another example concerns the emergence of VS(O) in several Palaungic languages. In these languages, verb-initial order is limited to relative clauses, and its traceable source construction suggests that VS(O) is a recent innovation in the Palaungic branch of Austroasiatic (Lee & Jenny 2022).

Finally, some studies propose a generalized rate of change across contexts, suggesting not much difference in the rate at which main and subordinate clauses change word orders (Kroch 1989, 2001; Salaberri 2018). Potential evidence for the generalized rate hypothesis comes from the loss of the verb-second constraint in Middle French, revealing a constant slope in a logistic model (Kroch 1989). This model of change has been generalized to model any binary variant (Yang 2000; Kauhanen & Walkden 2018).

Here we contribute to resolving this conflict of evidence by moving (i) from analyses of diachronic chronologies to estimates of rates of change, (ii) from individual histories of single lineages to analyses of an entire language family, and (iii)

from categorical grammar statements to corpus-based estimates of probabilistic grammars. We motivate these moves as follows.

First, previous research has evaluated conservatism by establishing whether word order has changed first in main clauses or first in subordinate clauses in particular chronologies. However, even if, for example, main clauses have changed before subordinate clauses in a particular case, it might still be the case that the changes in the main clauses have taken more time than the changes in the subordinate clauses and so main clauses really would have been more conservative than subordinate clauses. Such a scenario would not support the inference of an ancestral state based on the resulting order in a subordinate clause. The chronology of a single change cannot provide decisive evidence on conservatism. In response, we assess conservatism directly by estimating rates of change using Bayesian phylogenetic comparative methods (Dunn et al. 2011; Zhou & Bower 2015; Bickel et al. 2015; Widmer et al. 2017).

Second, while individual histories have higher data resolution, they risk missing the larger picture: we have no guarantee that the histories of German, English or French are representative of diachronic trends in Indo-European at large, let alone in the world. To assess a general principle of conservatism we need large-scale quantitative analyses of entire phylogenies. Here we make a first step focusing on Indo-European. We base our analysis on two sets of available phylogenies in order to fully incorporate the uncertainty of tree topologies and branch lengths (Bouckaert et al. 2012; Chang et al. 2015).

Third, the history and distribution of word order has mostly been assessed through categorical grammar data (Gell-Mann & Ruhlen 2011; Dunn et al. 2011; Maurits & Griffiths 2014) such as Dryer's data published in the *World Atlas of Language Structures* (WALS) (Dryer 2013b). In many languages, this categorical coding does not do justice to the orders that are actually used. For instance, though Czech has dominant VO order according to the criteria used in WALS, there are still about 25% OV utterances in text corpora. Thus, the grammar-based approach may underestimate the flexibility of word orders in certain contexts, while overestimating flexibility in other contexts. This is detrimental for assessing rates of change because any language change is ultimately mediated by usage patterns. In response to this, we model word order variation in usage as generated by probabilistic grammars, so that, roughly speaking, 25% OV results from a grammar that generates OV with a .25 probability. We estimate these probabilities in dependency-annotated corpora from 47 Indo-European languages in the Universal Dependencies (UD) scheme, version 2.9 (Zeman et al. 2021) supplemented by our own collection of word order frequencies for Hittite. We then enter these estimates, together with their uncertainties, in our phylogenetic models.

An important challenge in this endeavor is that any rate differences between main and subordinate clauses may be masked by general preferences towards one or the other order. Previous work has shown a distributional preference for subject-first or agent-first ordering (SO or SV) in the world's languages (Greenberg 1963; Hawkins 1983; Tomlin 1986; Dryer 2013b; Napoli & Sutton-Spence 2014; Futrell et al. 2015; Riesberg, Malcher & Himmelmann 2019), or a diachronic trend towards subject initial placement (Givón 1979; Li 1980), although it has remained unclear to what extent this preference is better explained by (neuro-)cognitive preferences of processing and event representation (Comrie 1981; DeLancey 1981; Goldin-Meadow et al. 2008; Kemmerer 2014) or by a topic-first principle akin to the information-structural constraints (Gundel 1975; Givón 1983; Tomlin 1986; Molnár & Hetland 2001) that ultimately also underlie the Penthouse Principle.

When it comes to the order of V and O, there is conflicting evidence, and different explanatory theories have been proposed. When controlling for geographical and phylogenetic sampling biases, (S)OV appears more frequent than (S)VO in the world's languages (Dryer 1989). Yet evolutionary models tend to suggest a higher chance of changes from SOV to SVO than the reverse (Gell-Mann & Ruhlen 2011; Maurits & Griffiths 2014). Diachronic evidence for the loss of OV order has been widely observed in diverse languages or families, especially Indo-European languages (Proto-Germanic and Proto-Romance), Niger-Congo, and arguably Chinese (Vennemann 1974; Hyman 1975; Givón 1979; Li 1980; Trips 2002; Pintzuk & Taylor 2008; Danckaert 2017; Fuß 2018). Other language families, such as Uralic, Afro-Asiatic and Uto-Aztecan, currently encompass more SVO order, though they still preserve some evidence for an earlier SOV order (Givón 1979). In general, most known cases in historical linguistics favor a directional change from SOV to SVO, whereas the opposite change is less commonly observed (Lehmann 1992). Theories for explaining the preferred order of V and O are quite mixed. On the one hand, in conjunction with a subject-first or agent-first hypothesis, theories of dependency length (or phrase domain size) minimization tend to predict a V-medial or VO order (Hawkins 1983; Liu 2008; Hawkins 2014; Futrell, Mahowald & Gibson 2015; Temperley & Gildea 2017). The balanced placement of S and O minimizes the linear distances between the dependents and the V. On the other hand, predictability maximization theory supports a V-final or OV ordering, since placing the dependents early maximizes the predictability of the V head (McDonough et al. 2011; Ferrer i Cancho 2017).

In response to these challenges and conflicting expectations, we model the evolution of word order separately for three pairs of orders (SV vs. VS, VO vs. OV, and SO vs. OS) and compare them individually across main and subordinate clauses. This means that we exclude from our considerations the impact of adver-

bials and other constituents. Our present interest is exclusively in pairwise surface linearization, in line with much research in processing, usage and typology (Greenberg 1963; Vennemann 1975; Hawkins 1983; Dunn et al. 2011; Dryer 2013b; Kemmerer 2014; Ferrer i Cancho 2017). We leave it to future research to assess the evolutionary dynamics of holistic generalizations over grammars, such as verb-second principles or cartographic approaches to syntax.

2. Data and Methods

2.1 Data

The UD database provides a collection of dependency-annotated corpora of diverse languages in the framework of Dependency Grammar (Hudson 1984; Mel'čuk 1988; Liu 2009). Each pair of words in a sentence is linked via directed arcs that indicate their dependency relations and order. We extract the frequency distributions of three pairs of word orders (SV vs. VS, VO vs. OV, and SO vs. OS) from 47 dependency treebanks in Indo-European languages, supplemented by hand-coded data for Hittite. These pairs of word orders guarantee a sufficiently high number of tokens in each corpus, and considerably reduce the complexity of our phylogenetic analysis compared to an approach with a six-way word order typology (Dryer 1997, 2013a). Because the position of pronouns and auxiliaries is often subject to confounding factors from phonology (cliticization) we only consider lexical elements for the main analysis. However, we include a sensitivity analysis in the Supplementary Materials that treats auxiliaries as heads of O and S and another one that excludes all sentences with auxiliaries.

Figure 1 provides an annotated English sentence containing a dependent clause, where the target word orders (SV, VO and SO) are displayed in italics. Here we strictly separate the word orders in main clauses, including only simple clauses, and subordinate clauses, including adjectival (relative) clauses (“acl”), adverbial clauses (“advcl”), clausal subjects (“csubj”), and clausal complements (“ccomp”). Accordingly, when coding word order in the subordinate clause in Figure 1 we exclude the main part (and vice-versa) to avoid any potential confounds from the overall complexity of the sentence.

Two sets of phylogenetic trees of Indo-European are taken from the literature (Bouckaert et al. 2012; Chang et al. 2015). The UD corpora include some ancient languages, like Old French and Old East Slavic, that are not covered in these phylogenies. We graft these languages onto the trees by following conventional assumptions about the relevant branching structure and allowing for uncertainty in the time estimate of the branching event (Widmer et al. 2017: Supplementary

Materials §S1.3). To incorporate the full uncertainty of tree topologies and branch lengths, we randomly sample 1000 posterior phylogenies, and estimate the transition rates on these samples.

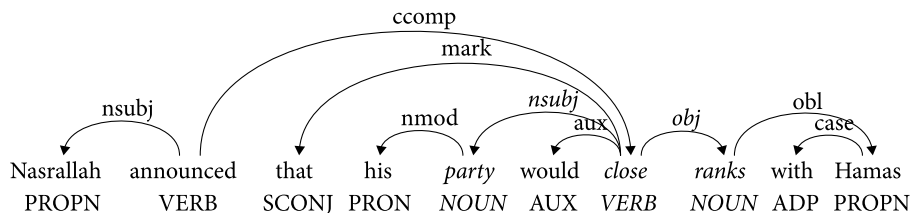


Figure 1. Dependency graph of an English sentence with a complement clause: Each terminal is annotated by its parts of speech tag (e.g., PROPN for “proper noun”), and each arc for its dependency relation (e.g., “nsubj” for “nominal subject”)

2.2 Methods

We use Bayesian phylogenetic inference to test hypotheses on word order evolution in the history of Indo-European. We go beyond earlier approaches where syntactic traits are coded as categorically fixed. Instead, we model word order as the result of probabilistic grammars. For example, there is a certain probability to produce a VO main clause in Czech. Importantly, however, the probabilities cannot be read off from raw proportions because corpus sizes vary substantially across languages in the UD database. This leads to different degrees to which we can be certain about our estimates (i.e., one would trust an estimate of a, say, .25 probability of VO main clauses more when it comes from a corpus of 100,000 sentences than when it comes from a corpus of 100 sentences).

In order to capture these differences in uncertainty, we replace the raw proportions by an estimate from a Bayesian binomial model of the raw frequencies in a given corpus with a flat Beta(1, 1) prior (Baath 2014; see Figure 2 for the procedure and Supplementary Materials §S2.4 for technical details). We then sample the probabilities of each word order from the posterior distribution of this model. This distribution reflects the uncertainty of the estimated word order probabilities, in direct reflection of the corpus size. For example, the posterior probability of a frequent word order will have less uncertainty than that of a rare word order. Occasionally, specific word orders are missing in the corpora because we focus on lexical dependencies. For example, there are no instances of subordinate clauses with lexical S and O instances in the Breton corpus, after excluding all sentences with auxiliaries, and so their order cannot be estimated. In these cases, we randomly draw from the prior distribution to assign a probability for SO/OS order.

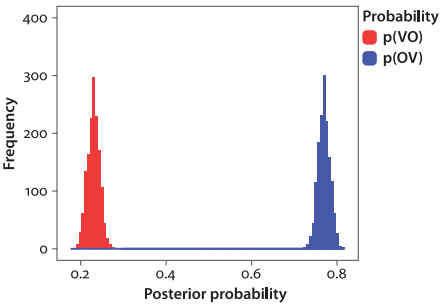
While probabilities can be modeled phylogenetically as continuous traits (e.g., with Brownian motion and Ornstein-Uhlenbeck models; Witzlack-Makarevich et al. 2016), we take an alternative approach that captures the discrete nature of word order differences more directly. To this end, we adopt two-state continuous-time Markov chains (CTMCs) that estimate the rate of change between orders using Felsenstein’s pruning algorithm (Felsenstein 1973; Cathcart 2018; Hoffmann et al. 2021), with one key modification inspired by the *fitMk* implementation of CTMCs in R’s *phytools* package (Revell 2012). While the state probabilities at the tips are usually treated as 0 or 1 in CTMC models of change, the pruning algorithm is not constrained to these values (see Supplementary Materials §S2.4.1), and we let them range over the entire interval between 0 and 1. This provides a match to the probabilistic grammar approach that we take here. For example, a .25 probability that a language is in a state of having VO main clauses corresponds to a grammar with a .25 probability of generating VO in main clauses. In the light of this, we fit a CTMC on a given tree and given probability estimates. We repeat this for 1,000 trees from the posterior tree sample combined with 1,000 samples from the posterior word order probability sample, thus accounting for the uncertainty in both.

The rate parameters in a CTMC denote the instantaneous rates of change between states (i.e., q_{12} and q_{21} denote the changes from state 1 to 2 and from state 2 to 1). Their inverse ($\frac{1}{q_{12}}$ and $\frac{1}{q_{21}}$) correspond to the expected time it takes for a change to happen (also known as “waiting times”; see Supplementary Materials §S2.3 for details). Via matrix exponentiation one can furthermore derive transition probabilities, i.e., the probability that a change happens within a certain time interval ($P(t) = e^{Qt}$). When sufficient time has elapsed, these probabilities converge on what is known as the stationary (or equilibrium) distribution (Maslova 2000; Cysouw 2011). Once stationarity is well approximated, there is, on average, an equal proportion of languages changing in either direction, keeping the relative frequencies steady and independent of the initial frequencies at the root. For example, given stationary probabilities for the transition from state 1 to state 2 ($p_{12} = .2$) and from state 2 to state 1 ($p_{21} = .8$) and a sample of 200 languages, one ends up with 160 languages in state 1 and 40 languages in state 2 regardless of whether one starts with 100 each or with 50 in state 1 and 150 in state 2, i.e., the long-term probabilities of spending time in state 1 are $P(1) = .8$ and those in state 2 are $P(2) = .2$.¹ In our results we report transition rates, stationary probabilities, the expected time for them to be reached, and the observed probabilities in corpora.

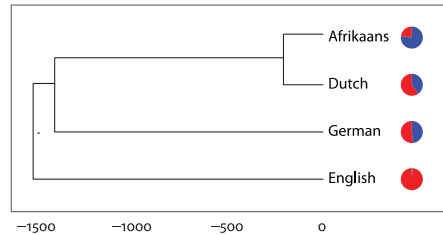
1. Concretely: when starting from 100 in state 1, state 1 gains 80 ($.8 \times 100$) and loses 20 ($.2 \times 100$); when starting from 50 in state 1, state 1 gains 120 ($.8 \times 150$) and loses 10 ($.2 \times 50$).

Language	VO	OV
Afrikaans	196	659
Dutch	1017	679
German	28753	26137
English	1738	11

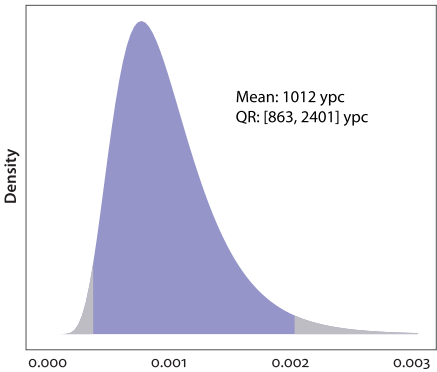
a. Frequencies of the order of V and O in main clauses from Germanic languages



b. Posterior probability estimate for the order of V and O in Afrikaans estimated by a Bayesian Binomial model



c. Germanic subtree with tip states probabilities sampled once from their own posterior distributions



d. Prior distribution of Lognormal(-7, 0.4) for Bayesian phylogenetic inference

Figure 2. A graphic representation of the procedures to estimate the rates of syntactic change in a CTMC model with tip state probability

We estimate the likelihood of transition rate matrices \mathbf{Q} for a given tree and tip state probabilities using an efficient MCMC method (the No-U-Turn sampler; Hoffman & Gelman 2014; Statisticat & LLC. 2021). To get an overall likelihood of the tree with Felsenstein’s pruning algorithm, we need to set a prior root state probability. In the main analysis we adopt an approach proposed by Maddison, Midford and Otto (2007) which weighs each root state according to its conditional likelihood, but we also fit the models under the more common assumption that the priors are the stationary probabilities. (See Supplementary Materials §S2.4 for detailed description of the procedure and the parameters chosen for the MCMC sampler).

Fitting the models also requires prior estimates on rates of change. For this we considered two prior distributions, the lognormal distribution and the gamma distribution, with their parameters set in such a way that the mean values for each rate (q_{12} or q_{21}) is around 1k years per change (ypc), with varying standard deviations (see Figure 2d for one example and Supplementary Materials §S2.2 for the full set). These prior distributions allow us to penalize extremely large and extremely small rate estimates, avoiding changes that are unrealistically fast or slow. For instance, the lognormal distribution with its mean as -7 and standard deviation as 0.4 in Figure 2d will have a mean value for each rate parameter to be around 1012 ypc, and a 95%-quantile range between 863 ypc and 2401 ypc. Much faster or slower change (e.g., within 500 or above 5000 years) is still possible under this prior but less likely. We estimate each rate separately, allowing for different speeds at which word order changes from, say, OV to VO as opposed to VO to OV.

To compare evolutionary biases of word order change, we calculate the ratios of estimated forward (q_{12}) and backward (q_{21}) rates in each model. If the 95% credible interval for these ratios includes 1 (or values near 1), this indicates no decisive bias in one or the other direction. If it does not include 1 (or values near 1), the interval suggests a preferred direction of change towards certain word order.

With the estimated rates of word order change in the phylogenetic models, we can also reconstruct character histories and ancestral state probabilities. To achieve this, we use stochastic character mapping, which infers evolutionary histories along phylogenies given observed tip states probabilities, transition rate matrix Q , and prior root state probabilities (Huelsenbeck, Nielsen & Bollback 2003; Bollback 2006; Revell 2012; Widmer et al. 2017). The mappings are stochastic and so we aggregate multiple mappings ($N=100$) into density estimates, providing the probability distribution of character states at given time intervals of around 30 years (i.e., roughly one generation of speakers). We incorporate the uncertainty of phylogenies into the mappings (sampling 100 trees from the posterior), but visualize the character histories onto a summary tree. For this, we make use of the `make.simmap` function from the `phytools` package in R (Revell 2012).

3. Results

Figure 3 plots the summary tree from Bouckaert et al. (2012), with proportions of observed word orders in the UD data. There are some noticeable differences (around 15%) across clause types in Romance and Germanic languages, especially for the orders of V and S, and V and O. For instance, Romance languages tend to have a higher proportion of SV order in main clauses than subordinate clauses,

whereas Germanic languages show an opposite trend with higher SV order in subordinate clauses than main clauses. In order to assess whether these differences reflect different rates of change we turn to the results of our models.

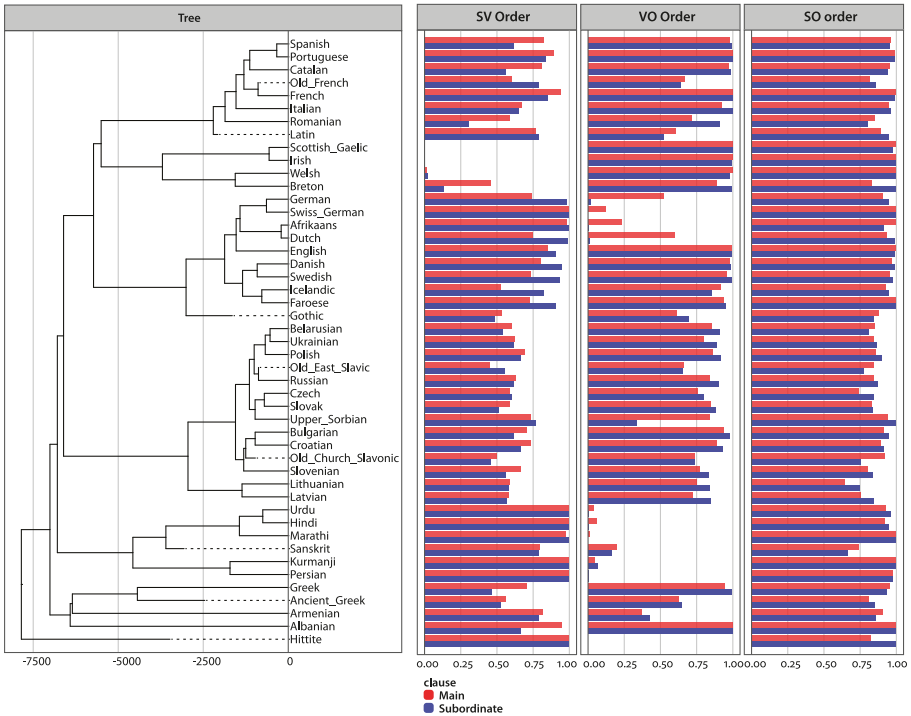


Figure 3. Distribution of word order proportions in main and subordinate clauses mapped onto the summary phylogeny from Bouckaert et al. (2012). The bar charts show the proportions of SV, VO and SO orders in each corpus. See Supplementary Materials Figure S1 for the same plot based on the phylogeny from Chang et al. (2015)

Figure 4 shows the distributions of mean rates for each word order change across clause types in Indo-European. For easier interpretation, we transform the evolutionary rate estimates into the expected number of years for a single change (ypc) by taking the inverse of the rate ($1/q$) (see §S2.3 for details). There are broad overlaps between main and subordinate clauses for all three word orders, suggesting no significant rate differences. Though the rate of change for the order of V and O in subordinate clauses (mean rate: 1509 ypc) seems to suggest slightly higher stability than in main clauses (mean rate: 1317 ypc), the effect is minor given the large credible intervals (i.e., uncertainties in the estimates).

Across individual word orders, we observe a certain amount of rate differences. The mean rate for the order of V and O (1390 ypc) is slower, i.e., patterns

are more conservative, than what we observe in the order of V and S (1104 ypc), followed by the order of S and O (879 ypc) in Figure 4. Similar distributions and rankings can be found with different priors or different phylogenies (Chang et al. 2015) (see Supplementary Materials Figures S3-S5 and S35-S38).

Figure 4 collapses the direction of the changes, i.e., averages the rates of changes from OV to VO with the rates of changes from VO to OV. But since most theories predict systematic differences between directions, we plot the rates separately in Figure 5. While the rate directions differ substantially, the differences are again similar in main and subordinate clauses. In other words, even when conditioning on the directions of change, main and subordinate clauses show very similar rates of change.

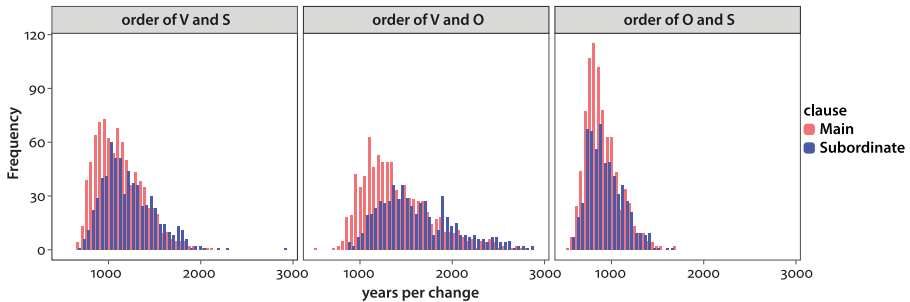


Figure 4. Distributions of mean evolutionary rates for each word order in main and subordinate clauses with the Indo-European phylogenies from Bouckaert et al. (2012), and a prior distribution of Lognormal(-7 , 0.4). For results based under alternative priors, see Figures S3-S5 in the Supplementary Materials. The overall pattern is virtually the same across models

Figure 5 also suggests striking evolutionary biases towards SO and SV orders in both clause types, since the transitions from OS to SO, and from VS to SV, take much less time than the reverse changes. There is also a marked bias towards VO order but the difference is weaker than in the other two orders. To assess these differences further we visualize the ratio between the forward and backward rates in Figure 6. The results suggest that these ratio estimates are far away from equal rates; ratios of 1 (or near 1) are indeed never included in the 95% credible interval. The change from OS to SO is around five times faster than the reverse, and the change from VS to SV is around three times faster than the reverse. Also, the change from OV to VO is around two times faster than the reverse.

A potential confound here is that our results might be artefacts of the fact that several ancient languages (Latin, Ancient Greek, and Gothic) are more V-final than their most closely related contemporary languages. However, fitting the models without any of the ancient languages still shows very similar patterns

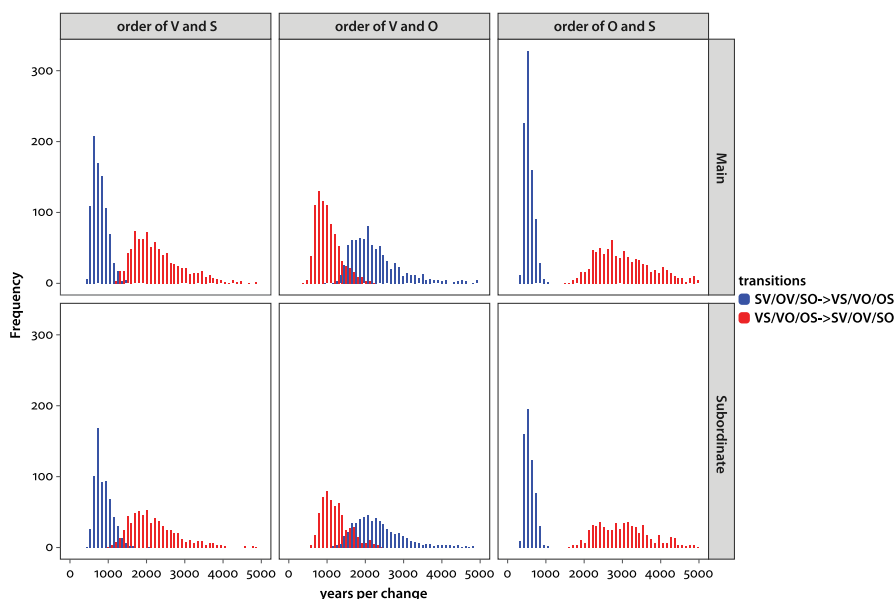


Figure 5. Directions of word order change across clause types with the Indo-European phylogenies from Bouckaert et al. (2012), and a prior distribution of Lognormal($-7, 0.4$). For results based on Chang et al.’s (2015) phylogeny and results based under alternative priors, see Figures S11-S13 and S44-S47 in the Supplementary Materials. The overall pattern is virtually the same across models

(see Figures S22 and S57 in the Supplementary Materials). Another possible confound is that the observed directional changes can be affected by speech act types, such as questions and imperatives. While the available corpus annotations do not allow a fully-fledged study, fitting our models to data where we exclude all sentences with an exclamation or a question mark yields very similar results to those reported here (see Figures S23 and S58). Furthermore, the orders of V and S and of V and O can be affected by the presence and annotation of auxiliaries in some Germanic languages. However, excluding sentences with auxiliaries or treating them (rather than the lexical verb) as heads of O and S shows again similar results (see Figures S24-S25 and S59-S60).

Moreover, the analyses in Figure 6 reveal a small potential difference between the strength of biases in main vs subordinate clauses: Subordinate clauses (blue bars in Figure 6) show rate ratios slightly closer to 1 than main clauses (red bars in Figure 6) in the relative orders of V and O and V and S. In terms of mean transition probability estimates, the probabilities are around 4% to 6% higher in main than in subordinate clauses within 1000 years (see Figure 7). No such effect obtains in the order of S and O, where the difference is less than 1%.

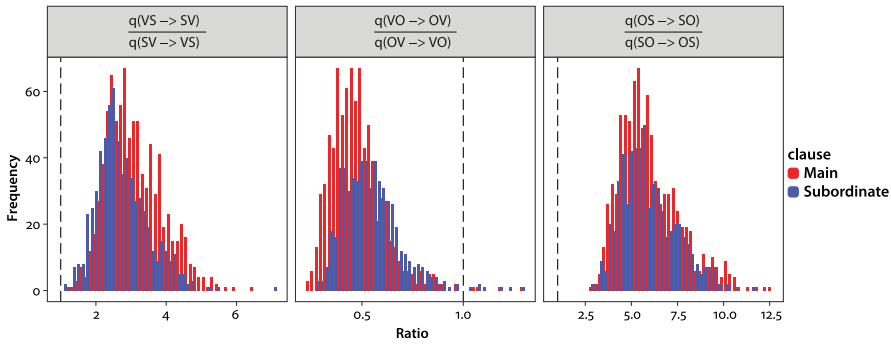


Figure 6. Ratios of forward and backward rates (q_{12}/q_{21}) with the Indo-European phylogenies from Bouckaert et al. (2012), and a prior distribution of $\text{Lognormal}(-7, 0.4)$. The dashed lines indicate a ratio of 1, i.e., equal rates. For results under alternative priors and phylogenies, see Figures S19-S21 and S53-S56 in the Supplementary Materials. The overall pattern is virtually the same across models

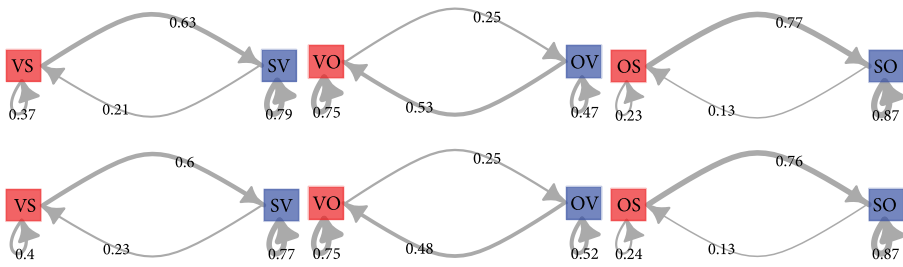


Figure 7. Transition probabilities of each word order over 1000 years with rates estimated with a $\text{Lognormal}(-7, 0.4)$ prior on phylogenies from Bouckaert et al. (2012). The upper panel shows estimated mean transition probabilities for each word order in main clauses, while the lower shows estimated mean transition probabilities for each word order in subordinate clauses. For results based on Chang et al.'s (2015) phylogenies, see Figure S62 in the Supplementary Materials. The overall pattern is virtually the same across models

Our models converge on SO and SV as the dominant order throughout the history of the family (with only recent deviations in Celtic). With regard to the V/O dependency, the models suggest a quick dominance of VO already when Anatolian split off (Figure 8). This pattern directly follows from the relatively fast rates of change that we estimated. This means that no matter what the exact distribution was in the proto-language, it has quickly become established as VO, except in Anatolian. Models based on slower rates (those based on a prior that privileges

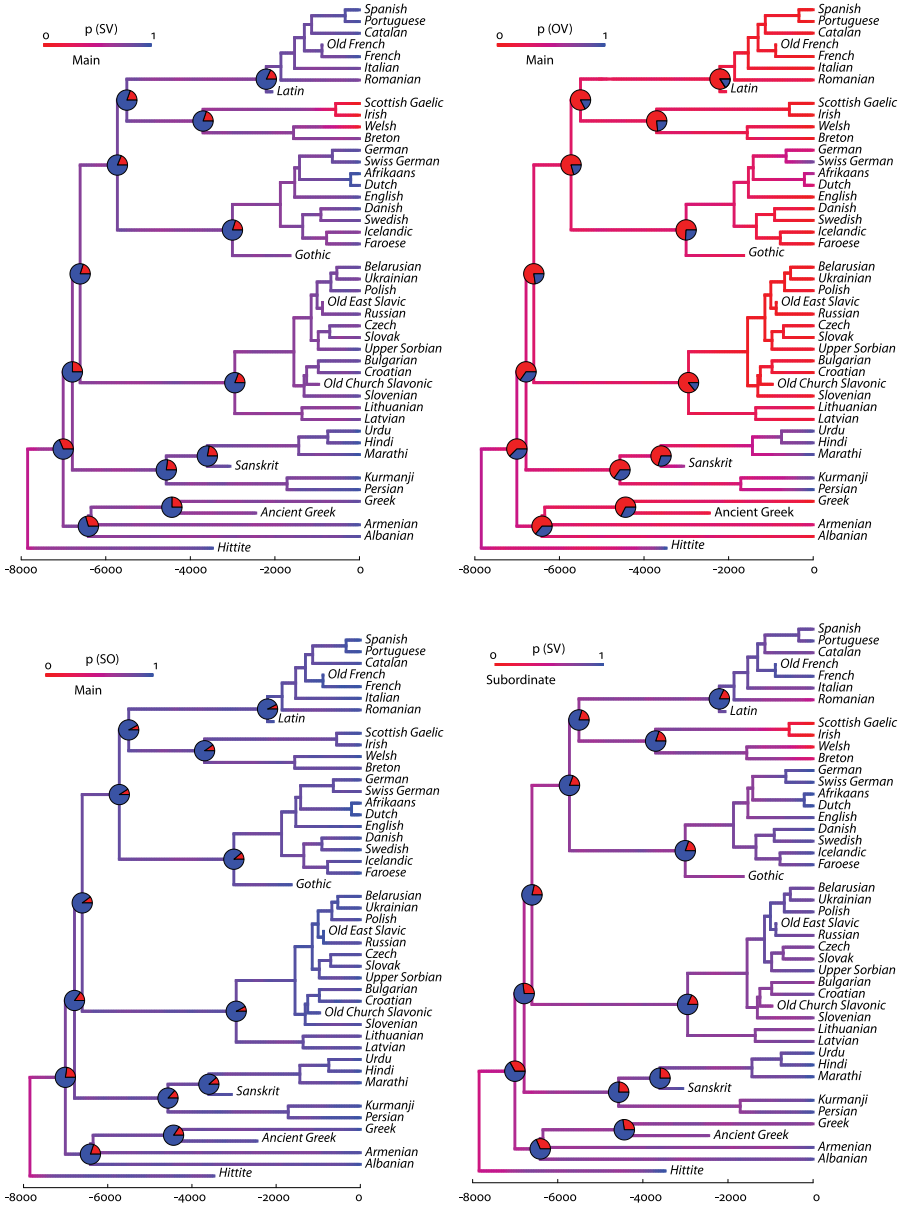
slower rates, i.e., the Lognormal2(−7.8, 0.5) prior; Figures S33 and S70) naturally suggest less change between Anatolian and the proto-language, and they therefore give a bit more weight to OV orders in the early stages of the family. The phylogenies from Chang et al. (2015) show the same overall picture but have slightly more balanced estimates in the early history (under all rate estimates).

4. Discussion and conclusions

Our analysis supports theories that assume equal rates of word order change in main and subordinate clauses (Figure 4), in line with the predictions by Kroch (1989, 2001). While our evidence is obviously limited by the sample of languages we have, it challenges theories that posit subordinate clauses to be universally and/or inherently more conservative than main clauses. This finding is surprising, since our sample is just as biased towards Indo-European languages of Europe as most previous research on word order change across clause types. This suggests that historical observations of word order change in a few well-documented histories like English, French and German, are not representative of the dynamics even in the subtree (clade) in which they are located. The full picture only emerges when estimating change over larger trees and over the entire history of the family. In other words, past research missed the general patterns because they were overly concerned with individual histories and languages; it missed the forest for the trees. If one moves, as we do here, from individual histories and grammar data to rate estimates and corpus data, there is not even good empirical backing for the theory of conservative subordinate clauses in the languages for which it was first developed.

Our analysis also reveals noticeable differences in the rates between different pairs of words/constituents, independent of clause type (Figure 4). We find that S/O changes faster (median 878 ypc) than S/V (median 1115 ypc) than V/O (median 1406 ypc). The higher rates for both S/O and S/V as opposed to V/O is consistent with Hawkins' (1983) Mobility Principle for short constituents since S is typically shorter than O (Temperley 2007). The higher rate of change for S/O as opposed to the other pairs arguably results from the importance of this order for information structure. However, other factors must be involved as well since information structure is also well-known to bear heavily on S/V orders, especially in the Indo-European languages of Europe (Lambrecht 2000; Sasse 2006). Further research is needed to resolve this, factoring in the interaction between S/V and V/O orders and the evolution of case and agreement marking.

A further result of our study is strong differences in the directions of word order changes across clause types: OS changes to SO around five times faster



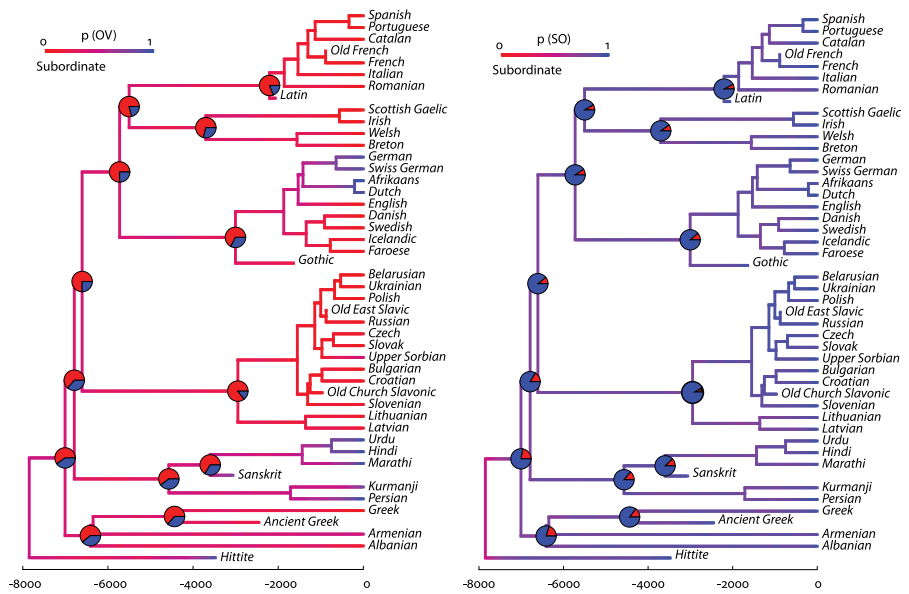


Figure 8. Stochastic character mapping of each word order with the phylogeny from Bouckaert et al. (2012) and rates estimated with a Lognormal(−7, 0.4) prior. To capture the uncertainty of tree topologies and branch lengths, we estimate mappings for a sample of 100 trees with the summary tree as the reference, and aggregate the ancestral states for each internal node. The pie charts in each subplot indicate the estimated proportions of each word order: Red indicates VS, VO, and OS orders, while blue indicates the reverse. The first three mappings represent the probabilities of character states in main clauses, while the second three mappings represent the probabilities of character states in subordinate clauses. For results based on alternative rate estimates and phylogenies, see Figures S31-S33 and S67-S70 in the Supplementary Materials. We also included stochastic character mapping by setting prior root state probabilities at stationarity in Figures S34 and S71. The overall pattern is virtually the same across models except that the slower-rate models allow Anatolian to increase the estimated proportion of OV orders in the early history of the family and that Chang et al.’s phylogeny shows a less pronounced preference of VO throughout

than the reverse; VS to SV around three times faster than the reverse and OV to VO around two times than the reverse (Figure 6). The preference for SO and SV is consistent with earlier claims in the typological literature (Greenberg 1963; Tomlin 1986; Dryer 2013b), based either on a topic-first or agent-first principle. The preference for changes towards VO rather than OV order supports theories of word order based on dependency (or domain) length minimization (Hawkins

1983; Liu 2008; Futrell, Mahowald & Gibson 2015; Temperley & Gildea 2017) over theories based on predictability maximization (McDonough et al. 2011; Ferrer i Cancho 2017). Coupled with the preference for SV, the preference for VO means that languages are more likely to change from SOV to SVO than the reverse. This change significantly shortens the dependency and domain distances between the two NPs and the V, possibly reflecting constraints on working memory.

We also found marginal and tentative evidence for weak differences in the strength of these biases between main and subordinate clauses: while the difference is within the expected random variation, the results raise the possibility that biases towards SV and VO are slightly stronger in main than in subordinate clauses (difference in rate ratio, $\Delta_{\text{main-sub}} \frac{q_{12}}{q_{21}}$, for SV: mean = 0.42 and 95% CI = [-1.66, 2.53]; difference in rate ratio for VO: mean = 0.29 and 95% CI = [-1.19, 1.8]), while the bias towards SO is more even across clause types with difference in rate ratio for SO: mean = 0.16 and 95% CI = [-4.09, 4.43]). If this can be corroborated in further studies with larger samples and in other language families, it might suggest that while there are no general differences in the rates of change, main clauses are slightly more strongly exposed to the factors that drive changes in relative verb position but not, or less so, to the factors that drive the order of S and O. Given our findings on the direction towards SVO, the relevant factors driving verb position are likely to be grounded in working memory constraints on dependency lengths. If this is on the right track, it predicts that main clauses are more strongly subject to such constraints than subordinate clauses, but progress in this question will need to control for interactions between main and subordinate clause word orders (unlike in our study, where we kept them separate). No comparable difference is detectable for the SO bias. This is consistent with a general agent preference in language processing that is remarkably robust against contextual variation of various kinds (Haupt et al. 2008; Demiral, Schlesewsky & Bornkessel-Schlesewsky 2008; Wang et al. 2009; Bickel et al. 2015; Krebs et al. 2018; Riesberg, Malcher & Himmelmann 2019), including variation in topicality (Wang et al. 2010; Sauppe 2017).

Most of our rate estimates are relatively fast, even though our prior distributions allow slower rates of change and one of the posterior estimates (the change from SO to OS) is indeed slow (Figures S13 and S47). The time estimates of our models are in line with what is known from history. For example, the waiting time for a change from VS to SV is around 700 years and this matches with the known history of Breton (Hemon 1975). The waiting time for a change from OV to VO is about 1000 years, and in French, we indeed see a substantial shift towards VO within the last 800 years.

Most word orders appear to evolve so fast as to reach the stationary probability distribution in less time (in around 3300 to 5700 years) than the age of the

Table 1. Summary of the expected amount of time (in years) to reach stationarity, estimated stationary probabilities (p_1/p_2) and observed average probabilities for each word order across languages from phylogeny of Bouckaert et al. (2012) and a Lognormal(−7, 0.4) prior. The expected amount of time to reach stationarity and estimated stationary probabilities are calculated by multiplying the exponentiated rate matrix, Q , with enough time that the resulting probability matrix, P , remains constant within a threshold of 0.001. See Tables S10-S12 and S22–25 in the Supplementary Materials for results with alternative priors and phylogenies

Clause	Order	Time to stationarity	Stationary probabilities	Observed probabilities
Main	V and S	4090	0.25/0.75	0.31/0.69
Main	V and O	5010	0.68/0.32	0.68/0.32
Main	O and S	3330	0.15/0.85	0.11/0.89
Subordinate	V and S	4360	0.28/0.72	0.32/0.68
Subordinate	V and O	5740	0.65/0.35	0.67/0.33
Subordinate	O and S	3370	0.15/0.85	0.12/0.88

family (Table 1). For example, the estimated stationary probabilities for the order of V and O is around .68 being VO and .32 being OV, which is exactly the same as the observed probabilities in corpora. The difference between observed probabilities and stationary probabilities is around .02 to .06 for the orders of V and S and of O and S. These observations are consistent with findings from other areas of syntactic change in Indo-European (Widmer et al. 2017) and with findings from Austronesian that properties of grammar generally evolve faster than the lexicon (Greenhill et al. 2017). The findings are also consistent with the notion that synchronic distributions reflect the stationary probability distribution sufficiently well to allow inference about general patterns (Cysouw 2011; Widmer et al. 2017), i.e., current typological distributions of word order are close to what is expected from their phylogenetic dynamics. In other words, our results suggest that the currently observed data distribution (in Figure 3) is a relatively faithful mirror of the estimated evolution, regardless of what individual documented histories might suggest. We caution, however, that our results only bear on Indo-European and need to be compared to a much larger set of language families before any conclusion can be drawn about the language faculty more generally.

The fast rate estimates have implications for reconstruction. Indeed, the rates of change are fast enough to fall within the root’s uncertainty interval (around 7000 to 10,000 years for Bouckaert et al.’s phylogenies, and 5500 to 7500 years for Chang et al.’s phylogenies), i.e., there will have been major changes already within

that interval, and so it is futile to pick any state as the proto-state. This result challenges previous reconstructions of Proto-Indo-European as favoring a specific word order, such as SOV (Lehmann 1974; Givón 1979; Li 1980; Gell-Mann & Ruhlen 2011; Hock 2013; Maurits & Griffiths 2014), and it lends strong support to reconstructions that acknowledge substantial word order uncertainty (Friedrich 1975; Miller 1975).

The unclear state of the root notwithstanding, stochastic character mapping makes it possible to estimate the most likely evolution of word order probabilities after the break-up of the proto-language. This gives some insight into the early history of the family, at least outside Anatolian. The mapping suggests that the early history was biased towards an SVO order with much variation, similar in kind to what is attested for Ancient Greek.

Funding

This research was funded by the China Scholarship Council (Grant Nr. 201606320224) and Swiss National Science Foundation (Grant Nrs. #P500PH_202882 and #51NF40-180888)



Supplementary materials














Code and data summaries are available at <https://osf.io/x8u2r/>.

Acknowledgments
















We thank the anonymous reviewers and our colleagues, Lena A. Jager and Paola Merlo, for their helpful comments; any remaining errors and misconceptions remain our own.

References












- Baath, Rasmus. 2014. Bayesian First Aid: A package that implements Bayesian alternatives to the classical **test* functions in R. In *UseR! 2014 – the International R User Conference*, 86. https://www.sumsar.net/papers/baath_user14_abstract.pdf
- Bean, Marian C. 1983. *The development of word order patterns in Old English*. London: Croom Helm.
-  Bickel, Balthasar, Alena Witzlack-Makarevich, Kamal K. Choudhary, Matthias Schlesewsky & Ina Bornkessel-Schlesewsky. 2015. The neurophysiology of language processing shapes the evolution of grammar: Evidence from case marking. *PLoS ONE* 10(8). e0132819.
-  Bollback, Jonathan P. 2006. SIMMAP: Stochastic character mapping of discrete traits on phylogenies. *BMC Bioinformatics* 7(1). 88.






-  Bouckaert, Remco, Philippe Lemey, Michael Dunn, Simon J. Greenhill, Alexander V. Alekseyenko, Alexei J. Drummond, Russell D. Gray, Marc A. Suchard & Quentin D. Atkinson. 2012. Mapping the origins and expansion of the Indo-European language family. *Science* 337(6097). 957–960.
-  Bybee, Joan. 2002. Main clauses are innovative, subordinate clauses are conservative: Consequences for the nature of constructions. In Joan Bybee, Michael Noonan & Sandra Thompson (eds.), *Complex sentences in grammar and discourse: Essays in honor of Sandra A. Thompson*, 1–17. Amsterdam: John Benjamins Publishing Company.
-  Cathcart, Chundra Aroor. 2018. Modeling linguistic evolution: A look under the hood. *Linguistics Vanguard* 4(1). 20170043.
-  Chang, Will, Chundra Cathcart, David Hall & Andrew Garrett. 2015. Ancestry-constrained phylogenetic analysis supports the Indo-European steppe hypothesis. *Language* 91(1). 194–244.
- Comrie, Bernard. 1981. *Language universals and linguistic typology*. Oxford: Basil Blackwell.
-  Cysouw, Michael. 2011. Understanding transition probabilities. *Linguistic Typology* 15. 415–431.
-  Danckaert, Lieven. 2017. The loss of Latin OV: Steps towards an analysis. In Enoch Aboh, Eric Haeberli, Genoveva Puskas & Manuela Schonenberger (eds.), *Elements of comparative syntax*. Berlin: De Gruyter.
-  DeLancey, Scott. 1981. An interpretation of split ergativity and related patterns. *Language* 57. 626–657.
-  Demiral, Şükrü, Matthias Schlesewsky & Ina Bornkessel-Schlesewsky. 2008. On the universality of language comprehension strategies: Evidence from Turkish. *Cognition* 106. 484–500.
-  Dryer, Matthew S. 1989. Large linguistic areas and language sampling. *Studies in Language* 13(2). 257–292.
-  Dryer, Matthew S. 1997. On the six-way word order typology. *Studies in Language* 21(1). 69–103.
-  Dryer, Matthew S. 2013a. On the six-way word order typology, again. *Studies in Language* 37(2). 267–301.
- Dryer, Matthew S. 2013b. Order of Subject, Object and Verb. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/>
-  Dunn, Michael, Simon J. Greenhill, Stephen C. Levinson & Russell D. Gray. 2011. Evolved structure of language shows lineage-specific trends in word-order universals. *Nature* 473(7345). 79–82.
-  Felsenstein, Joseph. 1973. Maximum likelihood and minimum-steps methods for estimating evolutionary trees from data on discrete characters. *Systematic Zoology* 22(3). 240–249.
- Ferrer i Cancho, Ramon. 2017. The placement of the head that maximizes predictability. An information theoretic approach. *Glottometrics* (39). 38–71. <http://hdl.handle.net/2117/108830>
- Friedrich, Paul. 1975. *Proto-Indo-European syntax: The order of meaningful elements*. Butte: Montana College of Mineral Sciences.
- Fuß, Eric. 2018. The OV/VO alternation in early German: Diagnostics for basic word order. In Gisella Ferraresi Agnes Jäger & Helmut Weiß (eds.), *Clause structure and word order in the history of German*. Oxford: Oxford University Press.

- doi Futrell, Richard, Tina Hickey, Aldrin Lee, Eunice Lim, Elena Luchkina & Edward Gibson. 2015. Cross-linguistic gestures reflect typological universals: A subject-initial, verb-final bias in speakers of diverse languages. *Cognition* 136. 215–221.
- doi Futrell, Richard, Kyle Mahowald & Edward Gibson. 2015. Large-scale evidence of dependency length minimization in 37 languages. *Proceedings of the National Academy of Sciences* 112(33). 10336–10341.
- doi Gell-Mann, Murray & Merritt Ruhlen. 2011. The origin and evolution of word order. *Proceedings of the National Academy of Sciences* 108(42). 17290–17295.
- Givón, Talmy. 1979. *On understanding grammar*. New York: Academic Press.
- doi Givón, Talmy. 1983. Topic continuity and word order pragmatics in Ute. In Talmy Givón (ed.), *Topic continuity in discourse. A quantitative cross-language study*, 141–214. John Benjamins Publishing Company.
- doi Goldin-Meadow, Susan, Wing Chee So, Asli Özyürek & Carolyn Mylander. 2008. The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences* 105(27). 9163–9168.
- Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of language*, 58–60. MIT Press.
- doi Greenhill, Simon J., Chieh-Hsi Wu, Xia Hua, Michael Dunn, Stephen C. Levinson & Russell D. Gray. 2017. Evolutionary dynamics of language systems. *Proceedings of the National Academy of Sciences* 114(42). E8822–E8829.
- Gundel, Jeanette Marie. 1975. *The role of topic and comment in linguistic theory*. Harwood Academic.
- doi Haupt, Friederike S., Matthias Schlesewsky, Dietmar Roehm, Angela D. Friederici & Ina Bornkessel-Schlesewsky. 2008. The status of subject-object reanalyses in the language comprehension architecture. *Journal of Memory and Language* 59(1). 54–96.
- Hawkins, John A. 1983. *Word order universals*. New York: Academic Press.
- doi Hawkins, John A. 2014. *Cross-linguistic variation and efficiency*. Oxford: Oxford University Press.
- Hemon, Roparz. 1975. *A historical morphology and syntax of Breton*. Dublin: Dublin Institute for Advanced Studies.
- doi Hock, Hans H. 2013. Proto-Indo-European verb-finality: Reconstruction, typology, validation. *Journal of Historical Linguistics* 3(1). 49–76.
- doi Hock, Hans Henrich. 1986. *Principles of historical linguistics*. Berlin: Mouton de Gruyter.
- doi Hock, Hans Henrich. 1991. *Principles of historical linguistics* (2nd edn.). Walter de Gruyter.
- doi Hoffman, Matthew D. & Andrew Gelman. 2014. The No-U-Turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research* 15(1). 1593–1623.
- doi Hoffmann, Konstantin, Remco Bouckaert, Simon J. Greenhill & Denise Kühnert. 2021. Bayesian phylogenetic analysis of linguistic data using BEAST. *Journal of Language Evolution* 6(2). 119–135.
- Hooper, Joan B. & Sandra A. Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4(4). 465–497. <http://www.jstor.org/stable/4177789>
- Hudson, Richard A. 1984. *Word grammar*. Blackwell Oxford.

-  Huelsenbeck, John P., Rasmus Nielsen & Jonathan P. Bollback. 2003. Stochastic mapping of morphological characters. *Systematic Biology* 52(2). 131–158.
- 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*, 267–305. Austin: University of Texas Press.
-  Jenny, Mathias. 2020. Verb-initial structures in Austroasiatic languages. In Mathias Jenny, Paul Sidwell & Mark J. Alves (eds.), *Austroasiatic syntax in areal and diachronic perspective*, 21–45. Boston: Brill.
-  Kauhanen, Henri & George Walkden. 2018. Deriving the constant rate effect. *Natural Language & Linguistic Theory* 36(2). 483–521.
-  Kemmerer, David. 2014. Word classes in the brain: Implications of linguistic typology for cognitive neuroscience. *Cortex* 58. 27–51.
-  Krebs, Julia, Evie Malaia, Ronnie B. Wilbur & Dietmar Roehm. 2018. Subject preference emerges as cross-modal strategy for linguistic processing. *Brain Research* 1691. 105–117.
-  Kroch, Anthony S. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1(3). 199–244.
-  Kroch, Anthony S. 2001. Syntactic change. In Mark R. Baltin & Chris Collins (eds.), *The handbook of contemporary syntactic theory*. Blackwell.
-  Lambrecht, Knud. 2000. When subjects behave like objects: A markedness analysis of sentence focus constructions across languages. *Studies in Language* 24. 611–682.
-  Lee, Wei-Wei & Mathias Jenny. 2022. Syntactic change in Palaungic – Exploring the origins of an atypical Austroasiatic relative construction. *Linguistics of the Tibeto-Burman Area* 45(1). 23–75.
- Lehmann, Winfred P. 1992. *Historical linguistics: An introduction*. Routledge.
- Lehmann, Winfred P. 1974. *Proto-Indo-European syntax*. Austin & London: University of Texas Press.
- Li, Charles N. 1980. *Mechanisms of syntactic change*. University of Texas Press.
- Lightfoot, David. 1982. *The language lottery: Toward a biology of grammars*. Cambridge: MIT Press.
-  Lightfoot, David. 1989. The child's trigger experience: Degree-of learnability. *Behavioral and Brain Sciences* 12(2). 321–334.
-  Lightfoot, David. 2006. *How new languages emerge*. Cambridge: Cambridge University Press.
-  Liu, Haitao. 2008. Dependency distance as a metric of language comprehension difficulty. *Journal of Cognitive Science* 9(2). 159–191.
- Liu, Haitao. 2009. *Dependency grammar from theory to practice*. Beijing: Science Press.
-  MacLeish, Andrew. 1969. *The Middle English subject-verb cluster*. Mouton: The Hague.
-  Maddison, Wayne P., Peter E. Midford & Sarah P. Otto. 2007. Estimating a binary character's effect on speciation and extinction. *Systematic Biology* 56(5). 701–710.
-  Maslova, Elena. 2000. A dynamic approach to the verification of distributional universals. *Linguistic Typology* 4(3). 307–333.
-  Maurits, Luke & Thomas L. Griffiths. 2014. Tracing the roots of syntax with Bayesian phylogenetics. *Proceedings of the National Academy of Sciences* 111(37). 13576–13581.
-  McDonough, Colleen, Lulu Song, Kathy Hirsh-Pasek, Roberta Michnick Golinkoff & Robert Lannon. 2011. An image is worth a thousand words: Why nouns tend to dominate verbs in early word learning. *Developmental Science* 14(2). 181–189.

Mel'čuk, Igor. 1988. *Dependency syntax: Theory and practice*. Albany: State University Press of New York.

-  Miller, D. Gary. 1975. Indo-European: VSO, SOV, SVO or all three. *Lingua* 37. 31–52.
- Molnár, Valéria & Jorunn Hetland. 2001. Informationsstruktur und Reliefgebung. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, and Wolfgang Raible (ed.), *Language typology and language universals: An international handbook*, vol. 2:2, 617–633. De Gruyter.
-  Napoli, Donna Jo & Rachel Sutton-Spence. 2014. Order of the major constituents in sign languages: Implications for all language. *Frontiers in Psychology* 5. 376.
-  Pintzuk, Susan & Ann Taylor. 2008. The loss of OV order in the history of English. In Ans van Kemenade & Los Bettelou (eds.), *The handbook of the history of English*, 249–278. Oxford: Blackwell.
-  Revell, Liam J. 2012. phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* 3(2). 217–223.
-  Riesberg, Sonja, Kurt Malcher & Nikolaus P. Himmelmann. 2019. How universal is agent-first? Evidence from symmetrical voice languages. *Language* 95(3). 523–561.
- Ross, John R. 1973. The penthouse principle and the order of constituents. In Claudia Corum, T Cedric Smith-Stark & Ann Weiser (eds.), *You take the high node and I'll take the low node*, 397–422. Chicago Linguistic Society.
- Salaberri, Iker. 2018. *On the relationship between clause type and syntactic change: A corpus-based cross-linguistic study*. Abstract presented at the 51st Annual Meeting of the Societas Linguistica Europaea. Tallinn.
-  Sasse, Hans-Jürgen. 2006. Theticity. In Giuliano Bernini & Marcia L. Schwartz (eds.), *Pragmatic organization of discourse in the languages of Europe*, 255–308. Berlin: de Gruyter.
-  Sauppe, Sebastian. 2017. Symmetrical and asymmetrical voice systems and processing load: Pupillometric evidence from sentence production in Tagalog and German. *Language* 93(2). 288–313.
- Statisticat & LLC. 2021. *LaplacesDemon: Complete Environment for Bayesian Inference*. R package version 16.1.6. <http://www.bayesian-inference.com/software>
-  Stockwell, Robert & Donka Minkova. 1991. Subordination and word order change in the history of English. In Dieter Kastovsky (ed.), *Historical English syntax*, 367–408. Walter de Gruyter.
- Stone, Gerald. 1993. Sorbian. In Bernard Comrie & Greville G. Corbett (eds.), *The Slavonic languages*, 593–685. London/New York: Routledge.
-  Temperley, David. 2007. Minimization of dependency length in written English. *Cognition* 105(2). 300–333.
-  Temperley, David & Daniel Gildea. 2017. Minimizing syntactic dependency lengths: Typological/cognitive universal? *Annual Review of Linguistics* 4(1). 1–15.
- Tomlin, R. S. 1986. *Basic word order: Functional principles*. Croom Helm.
-  Trips, Carola. 2002. *From OV to VO in Early Middle English*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

- Vennemann, Theo. 1974. Topics, subjects, and word order: from SXV to SVX via TVX. In John Anderson & Charles Jones (eds.), *Historical linguistics*, 339–76. Amsterdam: North Holland.
- Vennemann, Theo. 1975. An explanation of drift. In Charles N. Li (ed.), *Word order and word order change*, 267–305. Austin: University of Texas Press.
- Wang, Luming, Mathias Schlesewsky, Balthasar Bickel & Ina Bornkessel-Schlesewsky. 2010. *The influence of topicality on Chinese word order processing*. Abstract presented at the 23rd Annual CUNY Conference on Human Sentence Processing, New York.
-  Wang, Luming, Matthias Schlesewsky, Balthasar Bickel & Ina Bornkessel-Schlesewsky. 2009. Exploring the nature of the ‘subject’-preference: Evidence from the online comprehension of simple sentences in Mandarin Chinese. *Language and Cognitive Processes* 24. 1180–1226.
-  Widmer, Manuel, Sandra Auderset, Johanna Nichols, Paul Widmer & Balthasar Bickel. 2017. NP recursion over time: Evidence from Indo-European. *Language* 93(4). 799–826.
-  Witzlack-Makarevich, Alena, Taras Zakharko, Lennart Bierkandt, Fernando Zuniga & Balthasar Bickel. 2016. Decomposing hierarchical alignment: Co-arguments as conditions on alignment and the limits of referential hierarchies as explanations in verb agreement. *Linguistics* 54. 531–561.
-  Yang, Charles D. 2000. Internal and external forces in language change. *Language Variation and Change* 12(3). 231–250.
- Zeman, Daniel et al. 2021. *Universal Dependencies 2.9*. LINDAT/CLARIN digital library at the Institute of Formal and Applied Linguistics, Faculty of Mathematics and Physics, Charles University. <http://hdl.handle.net/11234/1-3105>
-  Zhou, Kevin & Claire Bowern. 2015. Quantifying uncertainty in the phylogenetics of Australian numeral systems. *Proceedings of the Royal Society B: Biological Sciences* 282. 20151278.

Résumé

En matière de changement syntaxique, la question de savoir si les phrases enchâssées sont plus conservatrices ou innovatrices que les phrases principales reste en suspens. Nous approchons cette question en utilisant des données provenant de corpora de 47 langues indo-européennes annotées pour l'ordre relatif des principaux constituants verbe (V), sujet (S) et complément d'objet (O). Sur la base de ces données, nous examinons l'évolution des ordres S/V, V/O et S/O dans les phrases principales et enchâssées en appliquant des méthodes d'inférence phylogénétique bayésiennes. Nos résultats révèlent des taux de changement similaires pour les deux types de phrases, sans offrir de preuve de conservatisme particulier ni pour l'un ni pour l'autre. Quant aux trajectoires évolutives, nos modèles soutiennent des biais vers les ordres SV, VO et SO. Ceci est en accord avec l'hypothèse de la minimisation de la longueur des dépendances qui favorise le placement médial du verbe ainsi qu'avec l'hypothèse de la préférence du sujet qui favorise l'ordre SO. Enfin, nos résultats montrent que, bien que l'ordre des mots dans l'indo-européen commun ne puisse être estimé avec certitude, l'histoire précoce de la famille de langues indo-européennes a été dominée par une préférence modérée pour l'ordre SVO, avec une incertitude majeure quant au placement relatif de V et O.


Zusammenfassung

Es ist weiterhin ungeklärt, ob Nebensätze bezüglich Wandel in der Wortfolge konservativer oder innovativer sind als Hauptsätze. Anhand von Daten aus 47 dependenzannotierten Korpora und mit komparativen phylogenetischen Methoden in einem bayesianischen Rahmen modellieren wir hier die Entwicklung der Stellung von S/V, V/O und S/O in Haupt- und Nebensätzen im Indogermanischen. Die Ergebnisse unserer Modelle weisen ähnliche Wechselraten für beide Satztypen aus und geben keine Hinweise auf einen speziellen Konservatismus von Neben- oder Hauptsätzen. Unsere Modelle weisen auf evolutionäre Präferenzen für die Abfolgen SV, VO und SO hin. Diese Beobachtung steht im Einklang mit Theorien, die eine Präferenz für die Minimierung von Dependenzlängen und verbmediale Wortfolgen postulieren, sowie mit Theorien, die von einer Privilegierung von Subjekten ausgehen, was SO-Stellung begünstigt. Des Weiteren geht aus unseren Modellen hervor, dass die Wortstellung des Urindogermanischen nicht mit hinreichender Sicherheit rekonstruiert werden kann, dass aber in der frühen Geschichte der Sprachfamilie SVO moderat präferiert wurde, mit erheblichen Unsicherheiten bezüglich VO- und OV-Stellungen sowohl in Haupt- als auch in Nebensätzen.

Address for correspondence

Yingqi Jing
Department of Linguistics and Philology
Uppsala University
Thunbergsvägen 3H
75126 UPPSALA
Sweden

yingqi.jing@lingfil.uu.se

 <https://orcid.org/0000-0003-2831-5701>

Co-author information

Paul Widmer
Department of Comparative Language
Science University of Zurich
paul.widmer@uzh.ch

Balthasar Bickel
Department of Comparative Language
Science University of Zurich
balthasar.bickel@uzh.ch

Publication history

Date received: 9 July 2020

Date accepted: 10 July 2022

Published online: 17 August 2023