Reconstructing the evolution of Indo-European grammar

Abstract

This study uses phylogenetic methods adopted from computational biology in order to reconstruct features of Proto-Indo-European morphosyntax. We estimate the probability of presence of typological features in Proto-Indo-European on the assumption that these features change according to a stochastic process governed by evolutionary transition rates between them. We compare these probabilities to previous reconstructions of Proto-Indo-European morphosyntax, which use either the comparative-historical method or implicational typology. We find that our reconstruction yields a strong support for a canonical model (synthetic, nominative-accusative, head-final) of the proto-language and a low support for any alternative model. Observing the evolutionary dynamics of features in our data set, we conclude that morphological traits have slower change rates, whereas syntactic traits change faster. Additionally, more frequent, unmarked traits in grammatical hierarchies have slower change rates when compared to less frequent, marked ones, which indicates that universal patterns of economy and frequency impact language change within the family.

1 Introduction

1.1 A century of Indo-European syntactic reconstruction

More than a century has passed since the pioneering work on syntactic reconstruction by the Neogrammarians (Brugmann et al. 1897, 1893; Delbrück 1900; Wackernagel 1920), dealing with core issues in Indo-European grammar, such as case, word order, alignment, agreement, person agreement, position of the verb, and the behavior of clitics. The system reconstructed for Indo-European in these works was fundamentally based on a comparative-historical reconstruction of morphological and syntactic features of ancient Indo-European languages, with a strong focus on a systematic comparison of Old Indo-Aryan, in particular Vedic Sanskrit, with Latin and Greek. This model — which we label 'canonical' — used Sanskrit

as a template for syntactic reconstruction. In Hirt's words: 'Delbrück's point of departure is Sanskrit. If something is not present in Sanskrit, it does not belong to Indo-European' (Hirt 1934:5, our translation). A new era of syntactic reconstruction, which is reflected in Hirt's skeptical view, began with the decipherment of Hittite in 1915 and the discovery of the Anatolian branch of Indo-European. Due to the old age of Anatolian sources, some linguists considered the complex synthetic structure of Old Indo-Aryan and Greek a secondary development, holding the view that Anatolian reflects a more archaic system. Accordingly, the discovery of Anatolian gave rise to alternative theories of Proto-Indo-European grammar, involving ergative (Uhlenbeck 1901; Vaillant 1936) or isolating (Hirt 1934) structure, and resulted in the concept of Indo-Hittite (Sturtevant 1962). Any grammatical reconstruction post-dating Indo-Hittite has to consider the role of Anatolian systems in relation to Proto-Indo-European. Currently, 'Greco-Aryan' and 'Anatolian' models serve as complementary to each other in Indo-European grammar, e.g., for the reconstruction of the verbal system (Clackson 2007:114-42).

Another important approach to syntactic reconstruction emerged during the 1970s, stemming from research on typological implicational universals (Greenberg 1963, 1978), which was adapted to a model for reconstructing syntax (Lehmann 1974). Even though this model was regarded with skepticism from some comparative-historical scholars (Winter 1984), it had an important continuation in the typological diachronic approach of Nichols (1992, 1995, 1998). This approach has grown in importance in the era of computational typology (Bickel 2007; Wichmann 2014), giving birth to several alternative models for explaining typological change (Baker 2011; Croft et al. 2011; Dryer 2011; Dunn et al. 2011; Levy and Daumé 2011; Longobardi and Roberts 2011; Plank 2011; Cathcart et al. 2018).

Since the pioneering work of Greenberg (1963), most syntactic reconstruction has been influenced by implicational typology. By merging typology and comparative-historical syntax, several important contributions to syntactic reconstruction have been published in recent decades, targeting the syntax of Indo-European as well as that of other families (Harris and Campbell 1995). This area of research goes under the name of diachronic typology (Viti 2015). An important approach continues the active-stative reconstruction model for Proto-Indo-European (Bauer 2000; Gamkrelidze and Ivanov 1984; Schmidt 1979). Other targeted domains have been the reconstruction of an active-stative verbal paradigm (Jasanoff 1978), the collective-count plural in the case system (Melchert 2000), dative-subject constructions (Barðdal and Eythórsson 2009), and various aspect of modality, tense, voice, aspect, particles, gender (Meier-Brügger et al. 2010:374-412). Along with these works, there are a number

of excellent overviews on principles of syntactic reconstruction (Barðdal 2014; Ferraresi and Goldbach 2008; Roberts 2007), or monographs compiling recent progress in various areas of syntactic reconstruction (Josephson and Söhrman 2008; Kulikov and Lavidas 2015; Ledgeway and Roberts 2017; Viti 2015). In §§3–5, where we evaluate the results of our reconstruction, we discuss this literature in further detail.

1.2 Outline of the current study

Our study analyzes comparative concepts of Indo-European morphosyntax, including the linguistic categories of ALIGNMENT, VERBAL MORPHOLOGY, NOMINAL MORPHOLOGY, TENSE TYPOLOGY, and WORD ORDER. We analyze data from 125 languages, including ancient, medieval and modern languages from the Indo-European family. Our data set is extracted from the typological subsection of the database DiACL – Diachronic Atlas of Comparative Linguistics (Carling et al. 2018; Carling 2019), a collection of linguistic data from languages of Eurasia and other regions. The original data set of 108 binary features is re-coded to yield 65 categorical (i.e., non-binary) features.

We select a well-known and well-studied family with a long history of scholarship as the basis for our investigation. The aim of the study is twofold: first, we wish to assess the extent to which phylogenetic comparative methods, which can be used to estimate the probability of morphosyntactic features in Proto-Indo-European, agree with the results of previous models of syntactic reconstruction for the Indo-European family. Second, we aim to make inferences regarding the evolutionary dynamics and variability of different morphosyntactic features during the course of the history of the Indo-European languages. In §2, we describe the model, method, and data forming the basis for the current study. In §3, we evaluate the results of the reconstruction for Proto-Indo-European and envisage further research that could emerge from the data and the model. In §3.5, we give the results of a statistical study, where we compare our reconstructed results to three different models to comparative-historical syntax. In §4, we discuss the evolutionary dynamics and variability of the transitions of traits. Finally, in §5, we discuss our results, both in the light of previous reconstructions of Indo-European grammar as well as from the perspective of general theories of grammar evolution. Technical description of the methods used in this paper is found in the Appendix, and full details of the data employed and results are found in online supplementary material. The raw data set is available open access via the DiACL database (`https://diacl.ht.lu.se/`). All code, metadata, and data are available at the following link: `https://zenodo.org/record/4275010`.

Family	Type	Timeframe	Number
Indo-European	Archaic	2000 – 500 BCE	3
	Ancient	500 BCE – 500 CE	5
	Medieval	500 – 1500	29
	Modern	1500 – 2000	79
	Romani	1500 – 2000	9
TOTAL			**125**

Table 1: Number and type of languages in the data set of the current study (see S1)

2 Theory, model, data, coding, method, and analysis

2.1 Comparative-historical, typological and phylogenetic models of reconstruction

The model of morphosyntactic reconstruction introduced by the scholars of Indo-European in the nineteenth century is based primarily on the comparative-historical method, systematizing forms and meanings of morphemes in a manner that sets of paradigms, rules and syntactic patterns can be reconstructed to a proto-language. Even though morphemes can be reconstructed as a result of the comparative-historical method, the reconstruction of their syntactic function is nontrivial, due to the uncertainty of regularity and the problem of establishing directionality of syntactic change (Barðdal 2014). Nevertheless, this method of reconstruction is utilized by a number of scholars, even though there is agreement that the method should not be applied to properties which are unconstrained by morphology, such as word order (Harris and Campbell 1995; Harris 2008). Proponents of the comparative-historical reconstruction model argue that if a specific pattern aided by morphological reconstruction has survived in a majority of languages, then there is reason to reconstruct it to the proto-language (Campbell and Harris 2002:615). Critics of this model point to the directionality problem: we may reconstruct a pattern to an ancestral state of several daughter languages carrying the same pattern, but in case of a disagreement we do not know enough about the directionality of syntactic change to reconstruct one variant over another (Roberts 2007; Walkden 2013:363-67).

The model of reconstruction used by the typologists from the 1960s onwards is based upon a different principle: if language-internal implicational dependencies between typological features, so-called UNIVERSALS, can be identified, then these observations can be used as an argument for reconstructing typological properties to a proto-language. A major obstacle to

the adaptation of this model is how to deal with language-internal conflicts between features with respect to assumed dependencies, both in attested as well as in reconstructed languages. An example is the controversy over Indo-European word order, where reconstruction based on ancient languages does not yield a uniform result with respect to the proto-language (Friedrich 1975; Lehmann 1974; Watkins 1976; Winter 1984).

In phylogenetic comparative methods, the issue of reconstruction is formulated in probabilistic terms, using phylogenetic computational algorithms originally adapted from biology (Calude and Verkerk 2016; Silva and Tehrani 2016; Jäger 2019). These models assume a specific stochastic process underlying CHARACTER EVOLUTION, which usually involves TRANSITION RATES which characterize change between values of a linguistic variable (e.g., different main clause word orders) over a phylogeny. These rates are estimated on the basis of a phylogenetic representation, often a TREE SAMPLE inferred from basic vocabulary patterns or a comparable linguistic feature, and the distribution of the feature among the daughter languages. These rates can be used to reconstruct the probability of a given value at internal nodes of the tree, including the root (i.e., the node ancestral to all others in the tree), as well as infer locations on branches of the tree where change is likely to have taken place (Maurits and Griffiths 2014; Dunn et al. 2017; Widmer et al. 2017; Cathcart et al. 2018; Blasi et al. 2019; Cathcart et al. 2020).

Figure 1 provides a schematic toy diagram of an ancestral state reconstruction problem in a phylogenetic comparative framework. Given a phylogeny with observed data at the tips of the tree, the procedure has two objectives: (1) to infer transition rates between feature values, and (2) to estimate values for unobserved internal nodes that are most likely to have preceded the values displayed by (or inferred for) their descendants. In a Parsimony framework (a model that minimizes the total number of character-state changes), this often involves restricting the number of parallel changes over the phylogeny. In a likelihood-based framework, including its Bayesian extensions, this also involves inferring evolutionary rates that express the probability of changes between different states over various spans of time represented by the branch lengths of the phylogeny. In general, evolutionary rates are inferred while treating internal states as a nuisance factor, which are to be marginalized out for the sake of efficiency. The rates that are inferred, or their posterior distributions under the Bayesian approach, can then be used to estimate the probabilities of different states at different nodes in the tree, starting at the tips and moving toward the root (Felsenstein 2004; Yang 2014).

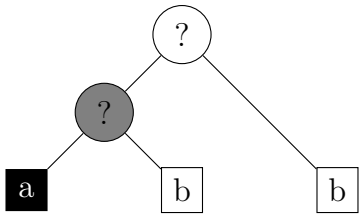


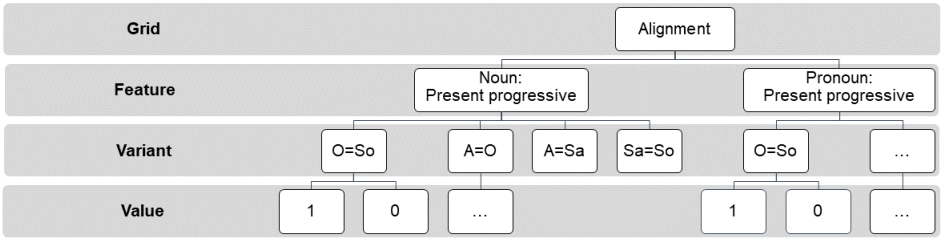
Figure 1: Visualization of ancestral state reconstruction in a phylogenetic comparative model. The probability that a trait is present at a given node is estimated based on the evolutionary rates inferred for the trait, as well as the probability that the trait is present in descendant nodes.

2.2 Data: Original data set and recoding for the current study

We use a data set of Indo-European languages extracted from the DiACL Typology/Eurasia dataset (Carling et al. 2018). The data involve categories of grammar that have been under discussion in both diachronic syntax and general typology for a long time, and are coded according to a hierarchical model designed to represent morphosyntactic features with an adequate level of granularity. Like other similar databases, e.g., AUTOTYP (Bickel and Nichols 2002) or WALS (Dryer and Haspelmath 2013), the data set consists of comparative concepts (Haspelmath 2010), definitions of linguistic features of grammar designed for cross-linguistic comparison. The original binary hierarchical model of DiACL, related to multivariate approaches (Bickel 2007), organizes comparative concepts according to levels of increasing detail. We re-code the binary data so that the data set consists of categorical VARIABLES[1] (e.g., MAIN CLAUSE WORD ORDER) taking multiple values (e.g., SVO, SOV, VSO, V2), organized within larger morphosyntactic CATEGORIES (comprising ALIGNMENT, WORD ORDER, NOMINAL MORPHOLOGY, VERBAL MORPHOLOGY, and TENSE). This results in 65 categorical variables, 64 of which are 'informative' in that they show variation within Indo-European and are thus suitable for phylogenetic analysis.

There are several advantages of using a hierarchical model for typological data, as in the original data set of our study DiACL (Carling et al. 2018) or AUTOTYP (Bickel and Nichols 2002). The most important advantage is the possibility of increasing detail, which enables local adaptations as well as the possibility of testing grammatical relations. Another advantage is the ability to recode the binarized strings of 1 and 0 into new combinations

[1]There are a number of terms for data of this type; in biology, it is common to refer to such data as a a multistate character (e.g., eye color) which can be realized as one of several traits (e.g., green eyes). We use the terms variable and value for the purpose of terminological and conceptual comparability with other work in quantitative linguistics (rather than biology), at times using the term FEATURE interchangeably with value.



a.

Categorical feature Alignment| Noun: Present progressive	**Variant** O=So	**Variant** A=O	**Variant** A=Sa	**Variant** Sa=So	**Trait**
Categorical feature Alignment| Noun: Present progressive	**Value** 0	**Value** 0	**Value** 0	**Value** 1	Noun, Present progressive: Tripartite
	Value 0	**Value** 0	**Value** 1	**Value** 1	Noun, Present progressive: Nominative-accusative
	Value 1	**Value** 1	**Value** 1	**Value** 1	Noun, Present progressive: No marking

b.

Figure 2: Organization of typological data before (a) and after (b) recoding. Figure 2a illustrates the hierarchical principle of organizing linguistic properties into Grids, Features, Variants and Values, which is used in the database DiACL. Figure 2b demonstrates how this hierarchy is mapped into categorical features, which contain blocks of value combinations, defined as traits. Both are illustrated for the categorical feature 'Alignment, Noun: Present progressive' (S32a).

which match a specific research question. A further advantage is the possibility of contrasting features across grammatical categories (cf. §2.3.2).

The procedure for transforming the hierarchically organized original data into our scheme of recoded categorical variables is shown in Figure 2, exemplified using alignment. The complete recoded data is given in S3a. The various coding and recoding strategies are described under the respective sections, where we discuss and evaluate results (§3).

2.3 Additional data sets

2.3.1 Coding of reconstruction models

Our study compares a probabilistic model of reconstruction with insights from comparative-historical approaches to syntactic change. For this purpose, we have selected a number of well-known approaches to the reconstruction of Indo-European syntax against which to compare our results. There is a rich literature on the reconstruction of Indo-European syntax, the full treatment of which is outside of this paper's scope. For the sake of simplicity, we have limited our comparisons to representative publications that address all grammatical cat-

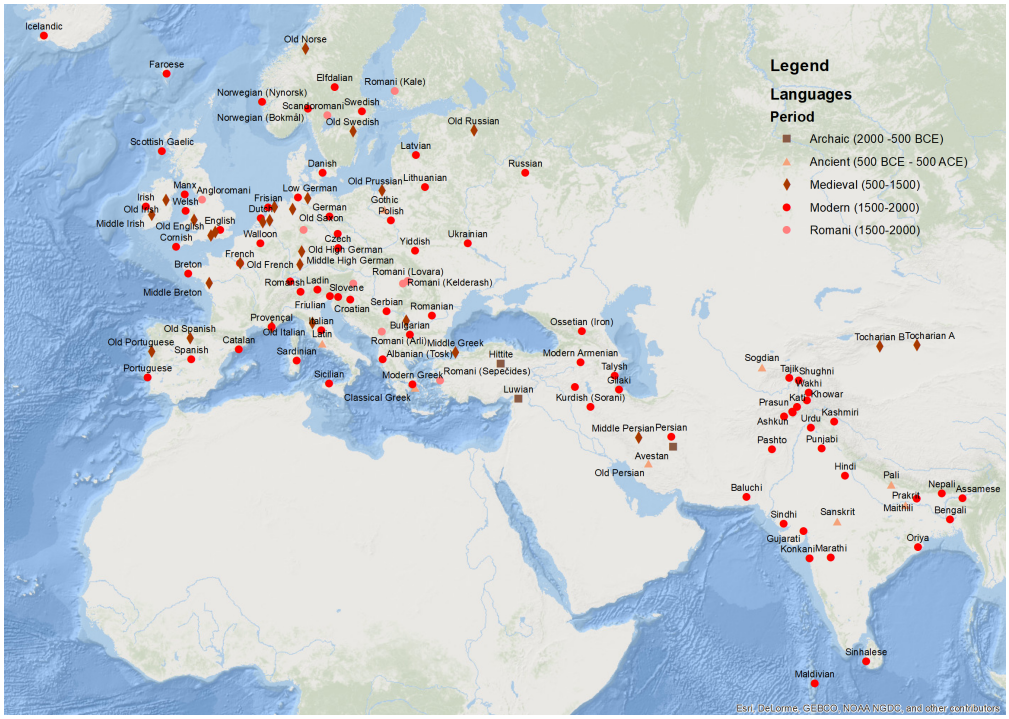


Figure 3: Map of language locations in the data, distinguished by time period

egories present in our data. We found three descriptions that were complete enough that we could treat the models proposed according to a coding scheme against which we could leanly compare our results. These we label CANONICAL (Delbrück 1893, 1897, 1900), ISOLATING (Hirt 1934, 1937), and ACTIVE-STATIVE (Gamkrelidze and Ivanov 1984; Gamkrelidze et al. 1995). It is important to remember that alternative theories (isolating and active-stative) reconstruct a stratified Proto-Indo-European language. At the root, they reconstruct a joint Anatolian and Non-Anatolian stage (Indo-Anatolian), which *later* transforms into a stage which represents the predecessor of the Non-Anatolian languages. Substantial portions of the discussion within alternative theories deal with the process of system change from Indo-Anatolian to Non-Anatolian branches (Pooth et al. 2018). In order to ensure comparability of our results and previous theories of grammar reconstruction, we use an Indo-Anatolian reference phylogeny, which represents the consensus view on branching and time-depth of the Indo-European family. We take the root of our phylogeny, which serves as the joint ancestral stage of the Anatolian and Non-Anatolian sub-branches of the family, to represent the unattested PROTO-INDO-EUROPEAN (PIE) language (see further §§3–5 and S9).

The coding of feature variants of models, including source references, of the raw DiACL data is found in 2b; recoding of feature variants into our categorical features and traits is

found in S3a.

2.3.2 Coding of grammatical hierarchies in the data

Additionally, we implement a coding of grammatical hierarchies between features in our data (S7). The issue of grammatical hierarchies is of key importance to the implicational typology model of Greenberg (1966), Croft (1990, 2003), and Comrie (1981), and is implicitly connected to the frequency of grammatical categories as well as the markedness theory (Haspelmath 2006). During the course of our analyses, we found that our model displayed asymmetric results for different features, not just between basic categories (e.g., word order, nominal morphology, verbal morphology, alignment, tense), but also within categories, between features differing with respect to categories such as tense (present, past) or word class (noun, pronoun) (S2, S3). For this reason, we chose to adopt an additional model of coding in which we identify pairs of features that belong to the same grammatical category but vary with respect to other grammatical categories, which can be defined as in a grammatical hierarchical relation to each other. For the sake of simplicity and comparability, we reduce our grammatical hierarchies to pairs of features, which have been observed in previous literature (where they are often referred to as scales). There is a rich literature on grammatical as well as marking hierarchies in grammar, both from the perspective of individual languages as well as cross-linguistically (Bornkessel-Schlesewsky et al. 2015; Comrie 1981; Croft 2003; Haspelmath 2015; Malchukov 2015). Generally, grammatical hierarchies are based on three different criteria (Croft 1990:92; Croft 2003:156-57):

1. Structural criteria, i.e., marking in grammars,

2. Behavioral criteria, i.e., the inflectional and distributional patterns in languages, and

3. Frequency, i.e., the occurrence in text, both in individual languages and cross-linguistically.

Only a handful of the grammatical hierarchies mentioned in the literature recur in our data, and there is also disagreement about the hierarchical organization of some of the categories in our data. One such example is the relation between future and present. Whereas the original hierarchy of Greenberg (Greenberg and Haspelmath 2005; Greenberg 1966) and Croft (Croft 1990:92-93) puts these traits in the order PRESENT < FUTURE, other scholars (Malchukov 2015; Witzlack-Makarevich and Seržant 2018) place these properties in the order FUTURE < PRESENT < PAST on the basis of existing marking patterns in some languages. The issue is

Category	**Hierarchy of features (unmarked/more frequent < marked/less frequent)**
NP type	pronoun < noun
Tense	present < future
Tense	present < past
Grammatical relation	agent < object
Grammatical relation	agent/object < oblique
Gender	masculine/feminine < neuter

Table 2: Pairwise coded marking hierarchies in our data (see S6 for a complete list of coded relations), based on Croft (1990, 2003); Greenberg (1966); Greenberg and Haspelmath (2005).

complex: we are aware that many languages reverse general hierarchies in their grammatical systems (Bickel 2008; see also Garrett 2008).

For this purpose, we use general grammatical hierarchies (Aissen 2003; Haspelmath 2015) as our point of reference, establishing pairwise hierarchical relations which we then implement for selected features in our data set (Table 2). The reason why we use pairwise relations and not hierarchical scales (e.g., SINGULAR < PLURAL < DUAL) is that we intend to compare grammatical hierarchies and the transition rates inferred by our model in a systematic fashion. Since our data contain features which are defined according to several categories, features may recur in hierarchical pairs. As an example, the features PRONOUN, PRESENT PROGRESSIVE: NOMINATIVE-ACCUSATIVE and NOUN, PRESENT PROGRESSIVE: NOMINATIVE-ACCUSATIVE are in a hierarchical relationship (PRONOUN < NOUN), whereas the features PRONOUN, PRESENT PROGRESSIVE: NOMINATIVE-ACCUSATIVE and PRONOUN, SIMPLE PAST: NOMINATIVE-ACCUSATIVE are also in a hierarchical relationship (PRESENT < PAST). A number of features in our data are not involved in any grammatical hierarchy relation, for a number of reasons. One reason is that they lack a hierarchical grammatical relationship to any other feature in the data. We also choose to consistently mark negative values in a fashion similar to their positive counterparts; for example, NO SYNTHETIC PRESENT PROGRESSIVE and NO SYNTHETIC FUTURE are in a hierarchical relation PRESENT < FUTURE, just as SYNTHETIC PRESENT PROGRESSIVE and SYNTHETIC FUTURE.

We choose a priori not to code any hierarchies for word order. Even though it is evident that head-final traits (OV, RELATIVE-NOUN, POSSESSOR-POSSESSED, etc.) have lower rates of change (S5), we prefer not to enter into a discussion about possible marking hierarchies or general frequencies in word order (Croft 1990:84ff.).

2.4 Methodology: reconstruction with phylogenetic comparative methods

The methodology on which this paper relies assumes that linguistic variables evolve under a CONTINUOUS-TIME MARKOV PROCESS (for an introduction see Liggett 2010), a stochastic model which assumes that there exist rates of change between values of categorical variables which characterize their evolution over time. Accordingly, our model infers rates of change between values of the categorical variables in our data set, using a tree sample representing genetic relationships between languages of the Indo-European family. Once these TRANSITION RATES have been inferred, they can be used to estimate the probability of a value for a given variable at phylogenetic nodes where data are unobserved; these INTERNAL NODES correspond to reconstructible protolanguages, with the ROOT of the tree corresponding to Proto-Indo-European.

Concrete details regarding the generation of the tree sample and the inference procedure can be found in the Appendix. Our tree sample (S9) is generated as follows: we assume a fixed topology that agrees with received philological wisdom, and sample branch lengths from chronologically realistic intervals, yielding a tree with a root age uniformly distributed between 7000 and 6000 years BP. The model is Bayesian; we infer posterior distributions for transition rates, using Felsenstein's PRUNING ALGORITHM (Felsenstein 1981, 2004) to compute the likelihood of these parameters for trees in the tree sample. We estimate the probability of a value for a given variable at the root of the phylogeny (i.e., for Proto-Indo-European) by randomly drawing evolutionary rates from their respective posterior samples, iteratively sampling a value at the root (Nielsen 2002; Huelsenbeck et al. 2003; Bollback 2006), and normalizing the counts for each sampled state to yield probabilities between 0 and 1.[2] In §3, we evaluate these results.

[2]Phylogenetic rate inference and reconstruction requires practitioners to define the PRIOR PROBABILITY of different values of a variable at the root of the phylogeny. A common practice in biology is to use the stationary probability of the CTM process, which gives the probability of the system taking a particular value as time approaches infinity. Felsenstein (2004:252) states that this prior is appropriate, but only if we assume that the model of evolution has been operating for a very long time. An alternative approach is to assume that the equal prior probability of each value at the root, or to treat the root prior as an unknown parameter to be inferred. The issue of how to treat the root prior is not widely discussed in phylogenetic linguists (many studies do not mention the issue at all), with some exceptions (Maurits and Griffiths 2014). In our main analyses, we follow other work (Cathcart et al. 2018; Blasi et al. 2019; Cathcart et al. 2020) in employing the stationary probability as the root prior. At the same time, because use of the stationary probability may bias our reconstructions, we run our models under two additional inference regimes, one using a uniform (i.e., equiprobable) root prior, and one where the root prior is treated as a parameter to be inferred. We find that results critical to the evaluation of our model against different traditional models are not affected by this choice (a full analysis of this issue is found in Appendix E).

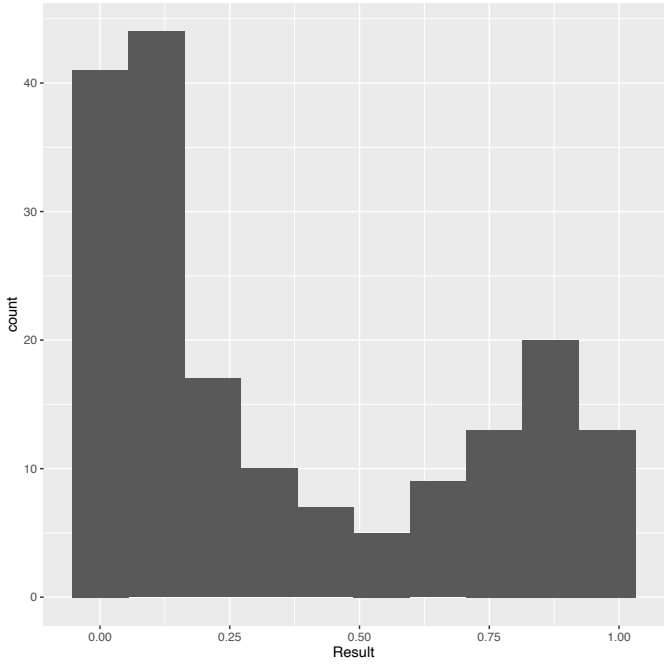


Figure 4: Histogram of results of reconstructions at the proto-language state (see S4). The lowest and (0.0–0.25) and the highest (0.75–1) probability ranges are more frequent than the intermediary ones (0.25–0.75).

In most cases, there is a clear result, where our procedure reconstructs a feature with relative certainty, inferring a high probability for a specific value of a variable and a low probability for the remaining values. This behavior can be seen in a histogram of all reconstruction probabilities (Figure 4); the distribution of these probabilities is U-shaped, in that low (0.0–0.25) and high (0.75–1) probability ranges are more frequent than the intermediate ones (0.25–0.75). While a small number of features are reconstructed with high uncertainty, meaning that we cannot say anything concrete about the value most likely to be present in Proto-Indo-European, the evolutionary dynamics of such features are still of interest, since these features may emerge in later phases of the Indo-European history. Furthermore, the behavior of such variables helps us diagnose the overall behavior of our model, with an eye to why it reconstructs certain patterns with high certainty. We discuss our model's results in detail below.

3 Results: reconstruction

Using the methodology outlined in the previous section, we reconstruct probability distributions across values for each variable in our data set at the root of the phylogeny. These distributions represent probabilities that particular features were present in Proto-Indo-European, under our model. A complete listing of all results, along with figures providing visualizations of the evolutionary history of all variables in our data set, are found in the supplementary material (S4, S8). In the following sections, we provide a detailed assessment of our results, organized thematically according to different domains of morphosyntax that have been discussed at length in the traditional literature on syntactic reconstruction in Indo-European, namely alignment, definiteness, gender, case, verbal typology, verbal morphology, and word order. Finally, we provide a quantitative comparison of our results against received wisdom in the form of models of reconstruction proposed by the different schools or camps of traditional Indo-European syntactic reconstruction described above, which we term the CANONICAL, ACTIVE-STATIVE, and ISOLATING models.

3.1 Alignment

For variables pertaining to alignment, our results support the reconstruction of nominative-accusative alignment in multiple systems (Figure 5). Nominative-accusative alignment is found with nouns as first argument in the present progressive, with nouns as first argument in the simple past, with pronouns as first argument in the present progressive, with pronouns in the simple past, with verbal marking in the present progressive, and in the simple past. At the same time, while nominative-accusative alignment is reconstructed with a higher probability than other alignment types across these systems, the certainty with which it is reconstructed varies. We note that nominative-accusative is more likely in the present progressive than in the simple past, both for nouns and pronouns, and more likely with pronouns as first argument than with nouns. The second most frequent type of alignment is no marking, followed by ergative (both with low probabilities).

This discrepancy is striking. Considering language-internal distributions of the clause and argument types involved, it is clear that nominative-accusative features are reconstructed with higher certainty for grammatical categories of higher cross-linguistic frequency (present, pronoun) as opposed to the more infrequent categories (past, noun). This result is of relevance to discussions of grammatical hierarchies (Croft 2003; Haspelmath 2006) (see §4). Second, we notice that the ergative appears (at a low probability) only in the simple past.

	Masculine/Feminine	Neuter
Nominative	-s/-∅	-∅/-m
Accusative	-m	-∅/-m

Table 3: Markedness in the Proto-Indo-European case paradigm in the active-stative theory (Bauer 2000:45; Szemerényi 1989:169)

	1st person	3rd person M./F.	3rd person N.
Nominative	*ego	*so/*sa	*tod
Accusative	*me	*to	*tod

Table 4: Suppletion in the PIE pronominal paradigm (Bauer 2000:45; Szemerényi 1989:169)

This result is reminiscent of results in the domain of verbal morphology (see §3.3.1), in which the simple past shows different patterns of change from the present progressive.

The reconstruction of patterns of alignment has a long history of discussion in comparative-historical syntax. In the canonical model of Delbrück and Brugmann (Brugmann et al. 1893; Delbrück 1897, 1900), the nominative codes the first argument (S/A), independent of transitivity of the predicate, and the accusative codes the second argument (O) (Meier-Brügger et al. 2010:401-404). However, due to the reconstruction of a case marking of *-s* for AGENT and *-m* for PATIENT, ergative alignment was proposed for Proto-Indo-European at an early date (Uhlenbeck 1901). This theory was later continued by Vaillant (1936) and Soviet scholars of the 1970s (Gamkrelidze and Ivanov 1984; Klimov 1974), who reconstruct an active-stative system, based on the **-os/-*om* distinctions in nominative/accusative and a corresponding **-os/*-om* distinction between genitives of active and inactive noun classes (Gamkrelidze et al. 1995: 233-76). Several scholars continue the active-stative theory (Bauer 2000; Schmidt 1979), reconstructing the relative chronology of the Indo-European paradigm, as well as reconstructing a continuation of change from an active-stative proto-language and into sub-branches, e.g., Italic (Bauer 2000). The source of the active-stative theories is a fundamental marking distinction between animate and inanimate (active-stative), reconstructed from the marking of the cases of nouns and pronouns in Proto-Indo-European (Table 3). The distinction is also reflected by suppletion in the pronominal paradigm (Table 4). The subject case has an unmarked zero-ending, against which the object is marked (Bauer 2000; Martinet 1962:44-46).

Under the active-stative theory, the active alignment is also marked in the two series of verbal endings, the **-mi* (active), and **-h_2e* (inactive) conjugation, supported by the *-mi* and

Probability range	1–0.9	0.9–0.8	0.8–0.7	0.7–0.6	0.6–0.5	0.5–0.4	0.4–0.3	0.3–0.2	0.2–0.1	0.1–0.0
Pronouns										
Present Progressive	nom-acc									neutral
Simple Past			nom-acc						ergative	neutral, tripartite
Nouns										
Present Progressive				nom-acc			neutral			tripartite
Simple Past					nom-acc		neutral		ergative	tripartite

Figure 5: Overview of probability ranges at the proto-language state for alignment

-*ḫi* paradigm setup in Anatolian (Gamkrelidze et al. 1995:254-76). However, as pointed out by other scholars, the formal contrast in Hittite between -*mi* and -*ḫi* conjugation is not reflected by any systematic difference in meaning (Jasanoff 2003:1-40), which is a prerequisite to the active-stative theory. At the same time, active-stative interpretations remain important in many theories of explanation of the Indo-European sets of endings (Jasanoff 1978).

The active-stative theory has no support under our reconstruction, pointing in the direction of nominative-accusative prevalence in Proto-Indo-European, both in the case marking on nouns and pronouns, in verbal conjugation, as well as in the present/past distinction (Figure 5). The active-stative and ergative theories are not generally supported by all Indo-European scholars (Meier-Brügger et al. 2010:412). However, they remain of great interest to us, since they connect to the reconstruction of the Indo-European gender and case systems, which yields interesting results on the basis of our data, consistent with the reconstruction of nominative-accusative alignment.

3.2 Nominal morphology

3.2.1 Case marking in the NP

Our results from the domain of NOMINAL MORPHOLOGY provide information regarding the position of case marking within the noun phrase in Proto-Indo-European. In the data set, we code whether languages mark case on adjectives, articles, the first element of the NP, and the head noun (S3a, 10-13). Our reconstruction provides support for the presence of case marking on head nouns (0.745) and adjectives (0.559), but not on the article, in line with the probable absence of definite articles in Proto-Indo-European (§3.2.2). Our system does not provide support for the presence of a rule that case must be marked on the last member of an NP (0.076). The case marking on the noun is not especially controversial: as

long as we reconstruct a synthetic case system of canonical type (see §3.2.4), we also expect case marking to appear on the nominal head in a noun phrase. However, the relatively lower degree of probability of case marking on the adjective (0.559) is not completely in line with the canonical model, which also reconstructs full case marking, with respect to case and gender, on adjectives (note the higher gender agreement value, see §3.2.3) (Delbrück 1893:402ff.).

3.2.2 Definiteness

On the whole, features pertaining to definiteness are reconstructed with low probabilities, indicating that the presence of definiteness in Proto-Indo-European is unlikely. This is the case for definiteness marked on the adjective, definiteness on the first element of the NP, definiteness on the last element of the NP, definite article, and definiteness suffix (S3a, 14-17). This result is uncontroversial with respect to all models, since it is evident from the historical record that most Indo-European branches developed definiteness marking independently, by means of grammaticalization (Bauer 2007).

3.2.3 Gender

Features targeting noun class and nominal gender (S3a, 18-22) display particularly interesting results. The probabilities of the presence of more than five noun classes (genders) and an animate gender are close to zero. However, the probability for having a masculine/feminine distinction is higher (0.684) than the probability of *not* having a masculine/feminine distinction (0.316). The probability for a special neuter gender is high (0.855). Furthermore, the probability for a predicative adjective to agree with its nominal head in gender is reasonably high (0.673, Figure 6).

These results for gender are noteworthy and somewhat controversial. At an early date, Delbrück (1893:132-133) is hesitant in reconstructing a Proto-Indo-European three-gender system, equivalent to the system found in archaic Indo-European languages, such as Sanskrit or Classical Greek. Considering the formal distribution of endings and the gender syncretism found in later Indo-European branches, he proposes that the three-gender system of Indo-European emerged from a two-gender system, based on an animacy-inanimacy distinction. Hirt (1934:28) reconstructs an Indo-European proto-language with no gender marking at all on nouns. In later literature, there is consensus around an original two-gender model of Proto-Indo-European, where the feminine is secondary (Szemerényi 1989:164-65; Tichy 1993; Gamkrelidze et al. 1995:242-24; Matasović 2004; Luraghi 2011). The issue of gender/noun

Probability range	1–0.9	0.9–0.8	0.8–0.7	0.7–0.6	0.6–0.5	0.5–0.4	0.4–0.3	0.3–0.2	0.2–0.1	0.1–0.0
Gender on predicative adjective				+			−			
Masculine-feminine distinction					+		−			
Neuter			+						−	

Figure 6: Overview of probability ranges at the proto-language state for gender

class is critical to arguments for reconstructing active-stative or ergative systems for Indo-European, and the animacy vs. inanimacy distinction is interpreted as an active vs. inactive, subject vs. non-subject distinction (Meier-Brügger et al. 2010:412). There are discussions of how a three-gender system emerged out of a two-gender system, i.e., how the animacy category split up into a sexus distinction, the possible distinction concrete–abstract and non-collective–collective, and the formation of a feminine gender in *-h_2, originally an abstract suffix, which was extended to the collective (Luraghi 2011; Matasović 2004; Tichy 1993). Although we reconstruct a masculine/feminine distinction with only moderately high probability, this result goes against the mainstream model in that it reconstructs a three-gender system for Proto-Indo-European.

There are several possible reasons for this result. By using comparative concepts, i.e., coded traits with no particular connection to individual pieces of morphological matter, the coding makes no difference between the two-gender system of Hittite (which is assumed to be preserved from Proto-Indo-European) and the two-gender system of, e.g., Dutch or Swedish (which collapsed from a previous three-gender system). Our model assumes that linguistic features evolve under a continuous-time Markov process, which estimates transition rates between values of a linguistic variable over time. The three-gender system is preserved and stable in many branches of Indo-European, as well as occasionally collapsed in some of the branches (e.g., Romance, Germanic), but not in a consistent way (masculine/feminine vs. common/neuter). For that purpose, the model estimates that it is more likely for Anatolian to have collapsed a Proto-Indo-European three-gender system than to have preserved an ancient two-gender system (see Figures 7–8, which display the most probable trajectories of historical development of these features under our model on a maximum clade credibility [MCC] summary tree, and further discussion in section 5).

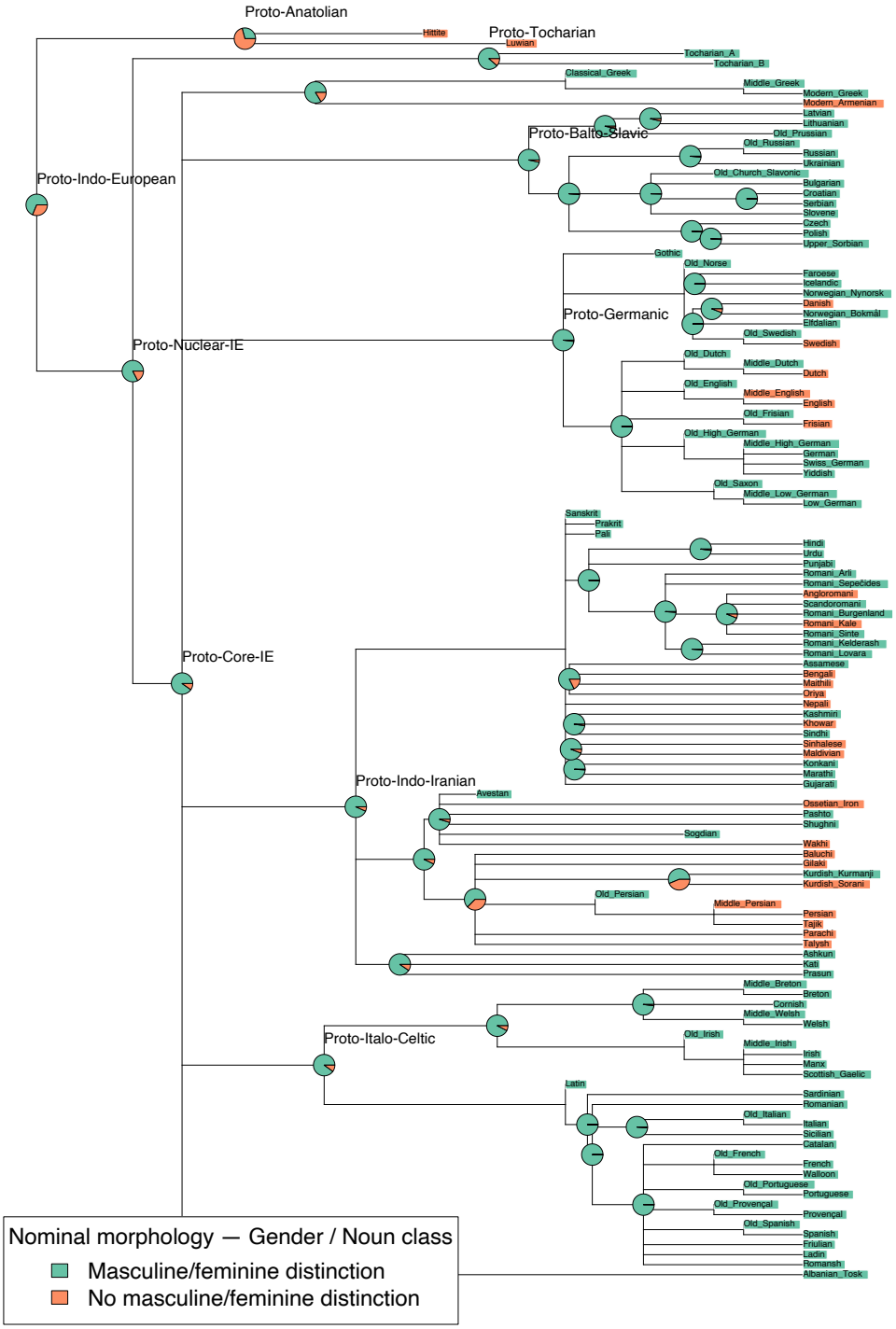


Figure 7: Maximum clade compatibility (MCC) tree with pie charts showing reconstructed probabilities of a masculine and feminine gender distinction at root and internal nodes of tree

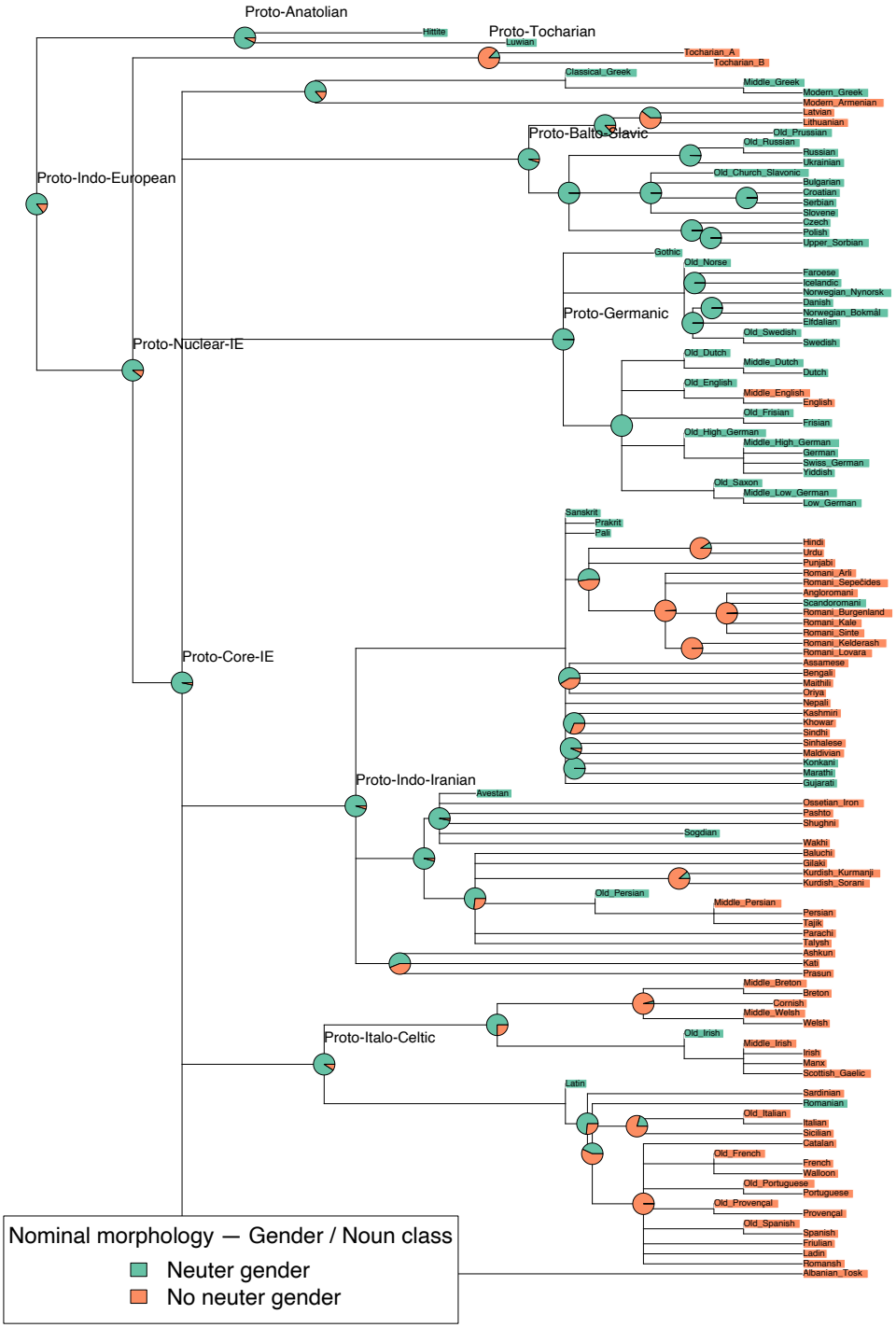


Figure 8: Maximum clade compatibility (MCC) tree with pie charts showing reconstructed probabilities of neuter gender at root and internal nodes of tree

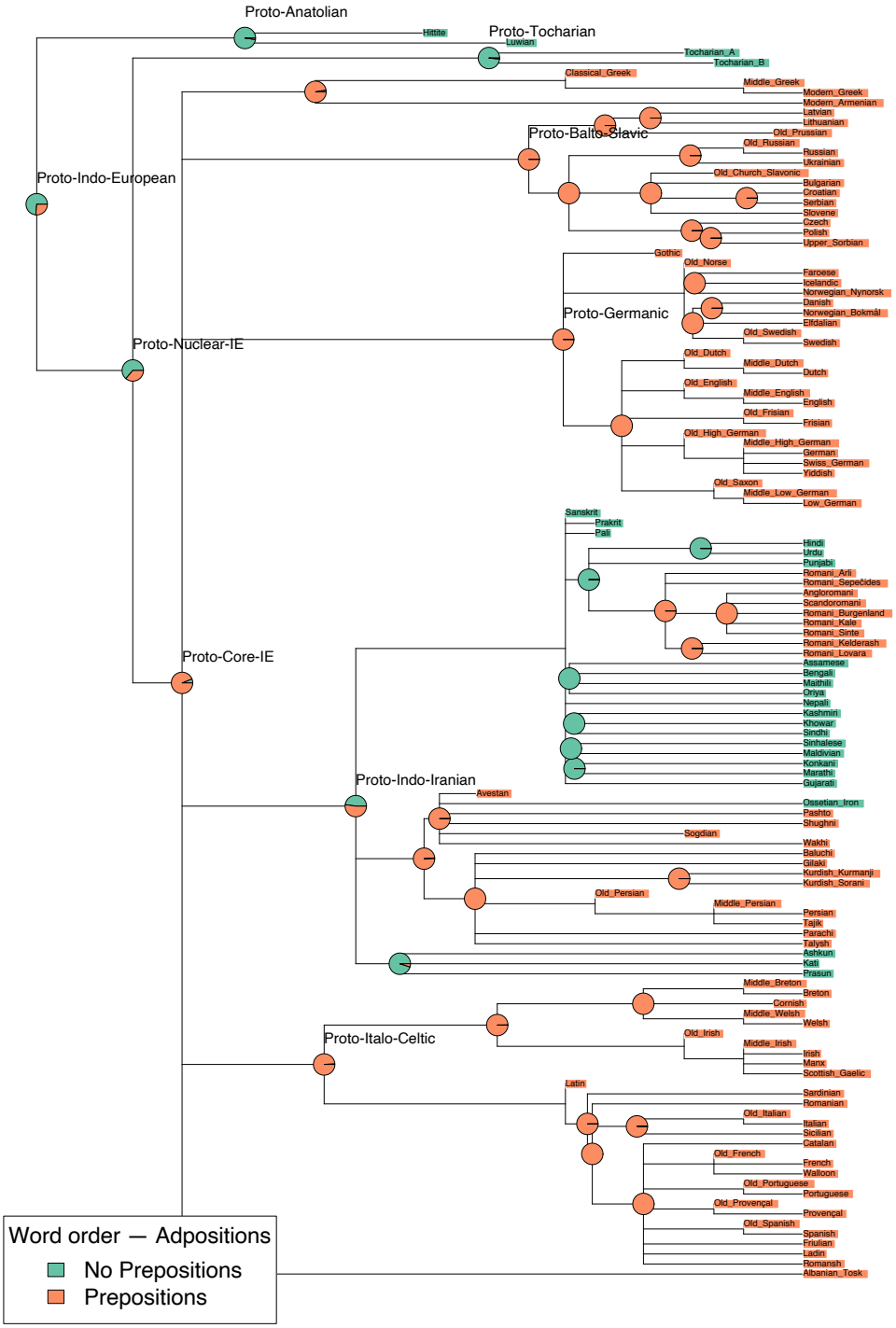


Figure 9: Maximum clade compatibility (MCC) tree with pie charts showing reconstructed probabilities of prepositions at root and internal nodes of tree

3.2.4 Case

Results for features pertaining to case show a degree of agreement with the canonical system of reconstruction similar to that of the features discussed in the foregoing sections. In Indo-European studies, the topics of morphosyntactic reconstruction of nominal morphology, the case system and its functionality, and paths of syncretism in various Indo-European sub-branches have been the subject of much debate, a full discussion of which is outside of this paper's scope. Instead, we use our results as a point of departure in an attempt to assess the extent to which they dovetail with previous theories regarding nominal morphology in comparative-historical syntax.

As far as typological structure is concerned, we code languages for the presence of agglutination for number and for case in nouns and pronouns (S5). All agglutinating traits have low reconstruction probability, somewhat higher for nominal morphology, but close to zero for pronominal morphology (Figure 11). Consequently, the model reconstructs absence of agglutination for Proto-Indo-European, which is more evident for pronouns than for nouns.

In comparative-historical syntax, the discussion of the typological structure of the Indo-European case paradigm relates to the discussions of alignment, described in the previous section. Agglutination is not a key issue in the canonical model, which bases its reconstruction on the synthetic Old Indo-Aryan paradigm. A variant of the active-stative theory by, e.g., Gamkrelidze et al. (1995), is found in Hirt (1934), who reconstructs an uninflected stage of Proto-Indo-European. This stage is preserved in the neuter, which has no marking. In a later stage, the distinction between *-s* and *-m* marks 'grammatical cases', i.e., alignment cases (Table 3). The genitive also represents a variant of the *-s* and *-m* forms. All other cases, 'local cases', are secondarily formed by means of postposing elements. Even though it is not prominent in Hirt's text, it is evident that Hirt presupposed an agglutinating stage of Proto-Indo-European (between an assumed isolating and a synthetic stage), at least for the case paradigm. The proposed pathway from an agglutinating stage to a synthetic stage is reminiscent of the theory of Bopp (1816) regarding the origin of the Proto-Indo-European verbal endings.

The issue connects to the number and type of Proto-Indo-European cases. There is much discussion on this topic, in particular whether Indo-European had a rich case system and was secondarily syncretic, or whether the high number of cases in Old Indo-Aryan and other ancient languages represents an innovation. An argument is that most Indo-European languages show a decay and an increase in case syncretism rather than a growth in case morphology (Szemerényi 1996:158). An exception is Tocharian, which can partly

be explained by the geographic position of Tocharian, surrounded by other agglutinating languages (Schmidt 1982). The evolution of Tocharian is also paralleled in Modern Indo-Aryan languages (Carling 2012). Delbrück (1893:180-91) reconstructs a case system which is identical to Sanskrit, with a nominative/vocative, accusative, genitive, ablative, locative, instrumental, dative and ablative. This is also the paradigm that scholars following the canonical model reconstruct (Meier-Brügger et al. 2010:398-410).

As for number and types of cases (Figure 10), our reconstruction is in line with the canonical model, albeit with a degree of uncertainty. We reconstruct as system with fewer than seven cases in the nominal paradigm (0.584), but with intermediate probability. Beyond that, we find, in the nominal paradigm, an intermediate probability for a dative and a genitive (0.608) but also a low score of (0.184) for having neither a genitive nor a dative. The score for having local cases outside of the core (in our coding the core includes cases for A, S, O, dative, and genitive) is high (0.983); the score for an accusative/objective, i.e., a case for O different from A is also high (0.715), just as the score for a vocative case (0.885).

The pronominal paradigm shows similar tendencies, with some important exceptions: the probability for more than seven cases is close to zero (0.031), the probability for an Agent-Object distinction higher (0.934), the probability for a dative moderate (0.429), and for non-core local cases high (0.922). In addition, the probability of pronominal vocatives is low. As with alignment, we notice that the results of the pronominal paradigm are more distinct, i.e., the difference between the preferred versus non-preferred variant is larger.

In sum, our reconstruction yields medium to high probabilities for a canonical system, with a nominative, accusative, genitive, dative, vocative and one or several local cases, which do not exceed seven in number. The pronominal paradigm is similar, but with the exception that the inference procedure reconstructs a very low probability for a system of more than seven cases, as well as for the vocative.

3.3 Verbal morphology and tense

3.3.1 Verbal morphology

The category VERBAL MORPHOLOGY targets agreement or person concord, i.e., the inflectional morphology of verbs with respect to their syntactic environment (Bickel 2007:169-171). A basic matrix (cf. Baerman and Brown 2013) including the variants *full agreement* (i.e., with reference to person and number) *gender agreement* (with respect to gender), and *no agreement* are matched against the core constituents S/A, O, and the case of the Recipient

Probability range	1–0.9	0.9–0.8	0.8–0.7	0.7–0.6	0.6–0.5	0.5–0.4	0.4–0.3	0.3–0.2	0.2–0.1	0.1–0.0
Pronouns										
Agglutination: case	−									+
Agglutination: number	−									+
Difference A and O	+									−
Difference O/Dative					−	+				
Peripheral cases	+									−
More than 7 cases	−									+
Vocative	−									+
Nouns										
Agglutination: case			−					+		
Agglutination: number			−					+		
Difference A and O			+					−		
Genitive and/or dative						+gen/+dat	−gen/-dat			+gen/-dat, −dat/+gen
Peripheral cases	+									−
More than 7 cases				+			−			
Vocative		+							−	

Figure 10: Overview of probability ranges at the proto-language state for case

Full agreement		1	0	0
No agreement		0	0	1
		Tocharian A *pälk-* 'shine' (Krause and Thomas 1960)	Gothic *bindan* 'bind' (Bammesberger 1986)	Swedish *sitta* 'sit'
Sg.	1	*pälkäm*	*binda*	*sitter*
	2	*pälkät*	*bindis*	*sitter*
	3	*pälkä*	*bindiþ*	*sitter*
Pl.	1	*pälkmäs*	*bindam*	*sitter*
	2	*pälkäc*	*bindiþ*	*sitter*
	3	*pälkiñc*	*bindand*	*sitter*

Table 5: Coding patterns and example paradigms for A-agreement patterns, default paradigm of present progressive, in Tocharian A, Gothic, and Swedish

(dative). The coding captures the typological variation in syncretism between *full* and *no* agreement (see Table 5). Only the A agreement has results that are of interest to us here; the probability for dative and O agreement, which is found in branches of Indo-European, is low for the proto-language (figure 12).

Our results (Figure 11) display two tendencies of interest to us: the higher probability of full A agreement (0.657) against syncretic A agreement (0.212) in the present progressive, and the higher probability of syncretic A agreement (0.207) against full A agreement (0.031) in simple past. Again, we see a pattern in which the more frequent category (present) is reconstructed with higher certainty than the less frequent category (past).

Our data set lacks a more fine-grained distinction between the various categories (e.g., voice, aspect, modality) than are present in our reconstructed Indo-European system and is therefore not fully comparable with the set of endings reconstructed for Indo-European via the comparative method. Much of the system complexity of ancient Indo-European languages is also lost in several branches of the modern languages (Clackson 2007:114ff.), a transformation over time that our data only reflect to a certain degree.

3.3.2 Typological marking of tense

Our data set's category TENSE takes as its focus the typological marking of present progressive and future, which is linguistically relevant from an areal perspective. In the Indo-European family, tense is historically integrated with the category of aspect: the original,

Probability range	1–0.9	0.9–0.8	0.8–0.7	0.7–0.6	0.6–0.5	0.5–0.4	0.4–0.3	0.3–0.2	0.2–0.1	0.1–0.0
Present progressive										
A agreement					full			syncretic		no
O agreement	no									syncretic, full
DAT agreement	no									syncretic, full
Simple past										
A agreement							syncretic	full	no	
O agreement			no							syncretic
DAT agreement		no								syncretic, full

Figure 11: Overview of probability ranges at the proto-language state for verbal agreement

tense-aspect system of Proto-Indo-European is thought to have developed in many branches into a system which is mainly tense-based (Hewson and Bubeník 1997). Aspect features are not included in the original data of DiACL, which spans more families than Indo-European (Carling et al. 2018).

Tense (S3a, 38-44) includes two properties designed to capture the typological profile of the forms used to mark present progressive and future. Looking at the present progressive (S3a, 38-39), the data distinguish whether a language uses a synthetic form or an analytic construction (with an auxiliary). The model reconstructs a high probability for a synthetic form (0.922), which is entirely consistent with the canonical model. There is an ongoing discussion in comparative-historical syntax as to whether some of the synthetic constructions of Indo-European, such as infinitives, might have an analytic origin (Meier-Brügger et al. 2010:320-21), but this is not relevant in connection to the present progressive, which makes our result uncontroversial.

The results pertaining to the future (S3a, 40-44) are also uncontroversial. The data distinguish whether a language uses an analytic construction (with auxiliary), a participle, a particle, a synthetic form, or an aspectual form. Our reconstruction yields a low probability for an analytic future formed by an auxiliary (0.372) but lower probabilities for all other variants. The future in Indo-European is an old issue. Delbrück (1897:242-55) doubted that the future of Sanskrit was derived from a future in Proto-Indo-European, due to its formal similarity with subjunctive and aorist. In further discussions of the verbal system, even within the canonical (Greco-Aryan) model, there is consensus that the future of Greek and Indo-Aryan is a secondary development (Meier-Brügger et al. 2010:295ff.; Rix and Kümmel

2001:10-30; Szemerényi 1989:244-47). Our result cannot contribute to this reconstruction; the probabilities are in general low.

3.4 Word order

Since Greenberg (1966), word order (constituent order, order of meaningful elements) has played a central role in linguistic typological research (Comrie 1981; Dryer 1992; Siewierska 1998). In diachronic syntax, reconstruction of word order remains a controversial issue. At their core, Greenberg's observations targeted implicational relations among word order types, in later literature defined as ORDER OF HEAD AND DEPENDENT (Lehmann 1973), a concept which also includes typological properties beyond word order, following upon the order types (Nichols 1992, 1995, 1998). In Nichols' model, the various dependency types are seen as stable both geographically as well as diachronically, something that is indicated by the fact that the types have regional skewing patterns (Nichols 1995). Word order harmony remains an issue also in computational typology, where the main source of controversy is whether word order patterns are mainly lineage-specific or areal (Baker 2011; Bickel 2011; Croft et al. 2011; Cysouw 2011; Donohue 2011; Dunn et al. 2011). To avoid confusion and not enter into too much detail in the scientific literature on word order, we use the terms 'head-initial' and 'head-final' to refer to the issue of constituent order in sentences and phrases.

In diachronic syntax, the reconstruction of word order is characterized as being beset by methodological difficulties (Roberts 2007:175-98). The source of the uncertainty is the fact that word order in most cases cannot be implicitly connected to any morphosyntactically reconstructable material. Therefore, reconstruction of word order has to be based on, at first, actual observation in attested languages, and second, connections to other properties in language, which may be either reconstructable or not. Consistency and harmony, as well as stability in word order patterns, are central in the model of reconstruction proposed by Lehmann and Nichols (Lehmann 1973, 1974; Nichols 1992), and this is also one of the major sources of criticism against the reconstruction of word order (Lightfoot 2002; Watkins 1976; Winter 1984). Word order consistency is not irrelevant: a majority of the world's languages are consistent. However, some languages are inconsistent, indicating that consistency cannot be used as a sole argument for reconstruction (Campbell and Harris 2002; Harris and Campbell 1995). It is also clear, that a diachronic shift from one type to another, e.g., from head-final to head-initial, is a complex evolution where archaic structures are retained and coexist side-by-side with more recent, changed ones (Bauer 1995).

Word order in Proto-Indo-European is the subject of a century-long debate, beginning

with Delbrück (2010:38-111) and the study on the position of clitics by Wackernagel (1920). The latter remains one of the few unquestioned reconstructed syntactic features of Proto-Indo-European (Clackson 2007:168). In recent decades, two competing positions on Proto-Indo-European word order have emerged, both of which require a consistency approach. Much of the critique of the word order theories revolves around problems of establishing a default order in ancient languages, which form the basis for a proto-language reconstruction (Winter 1984). Other researchers highlight the general problems of word order reconstruction due to the inherent problem of reconstructing variation and change (Lightfoot 2002; Pires and Thomason 2008). The mainstream position, also given by Delbrück, assumes verb-finality (OV) and head-final order for noun phrases (Clackson 2007; Hock 2013; Lehmann 1973, 1974, 1993, 2002; Mallory and Adams 1997:165-71). The competing position, which bases its discussion on problems of Proto-Indo-European relative clauses, assumes VO and head-initial order for Indo-European (Friedrich 1975).

Our data set contains standardized coding of word order in ancient languages and therefore, the results relate to the discussion of word order reconstruction based on evidence from ancient languages. As a rule, the coding policy aims to capture the *dominant* word order in a language, but in uncertain cases, the coding system allows for polymorphic coding, i.e., coding a value 1 for two or several variants (Carling et al. 2018). Word order is also split into a relatively high level of granularity, e.g., distinguishing different clause types (S3a, 51-64).

Considering the results of our reconstruction, we have to bear in mind that our model does not take into account implicational dependencies between variables (e.g., head-final or head-initial; Dunn et al. 2011; Murawaki 2018; Pagel and Meade 2006). The probability of values of a categorical variable is estimated independently of other variables.

Our model reconstructs SOV order with high probability in main clauses (0.905) as well as subordinate clauses (0.899). Furthermore, our model produces reconstructions of high probability for postpositions (0.849), possessor-noun order (0.585), adjective-noun order (0.870), OV order with participles (0.894), as well as OV order with infinitives (0.806; see Figure 12). These results are compatible with the mainstream view on Proto-Indo-European as a head-final language.

Our results for clitic pronouns (S3a, 53-55) are less simple to interpret. For clitic pronouns, we distinguish 2nd position, OV, and VO (if the language does not have clitic pronouns, the variable is not applicable to the language), with finite verb, infinitive, and participle. Our model reconstructs distributions of high uncertainty for all categorical features pertaining to clitics. Here, it is obvious that the situation in languages is too complex, with

Probability range	1–0.9	0.9–0.8	0.8–0.7	0.7–0.6	0.6–0.5	0.5–0.4	0.4–0.3	0.3–0.2	0.2–0.1	0.1–0.0
NP word order										
Adpositional word order		+postp	−prep					+prep	−postp	
Noun-possessor				poss-n			n-poss			
Noun-adjective		adj-n							n-adj	
Clause word order										
Main clause		SOV								VSO, V2, SVO
Subordinate clause		SOV							SVO	VSO, V2
Infinitive WO			OV						VO	irrelevant
Participle WO		O-part								part-O
WH-verb			−					+		
WH-initiality			+					−		
Noun-relative clause				noun-rel				rel-noun		irrelevant

Figure 12: Overview of probability ranges at the proto-language state for word order

too many gains and losses at hand, for a clear picture to emerge. This result is problematic, considering the safe reconstruction of the position of clitics in Indo-European (Krisch 1990).

Finally, we have an interesting result: a reconstructed high probability of noun-relative clause (0.627) over relative clause-noun (0.292; see Figure 12). The position and construction type of the relative clause was a major source of conflict between Lehmann and Friedrich, and has been extensively discussed in Indo-European syntax (Hock 2013; Watkins 1976). In accordance with the consistency theory by Greenberg (Greenberg 1963), continued by Lehmann (1973, 1974), an OV language is more likely to have relative clause-noun order (Harris and Campbell 1995:363-67), which is also the case in several of the archaic languages, such as Sanskrit and Homeric Greek. However, Hittite (paralleled in, e.g., Latin) has the reverse order relative clause-noun, which is inconsistent with OV (Clackson 2007:171-76). This result can only be taken to be provisional; due to the simplified definition of relative clauses in our data (NRel/RelN, which does not distinguish, e.g., correlative relative clauses, type of clause relation, i.e., paratactic or hypotactic, or restrictive/ non-restrictive types), our result does not bear fully on the issue of Proto-Indo-European relative clauses in the degree of detail with which they are treated in the comparative-historical literature (Hock 2013).

3.5 Comparison of results with traditional models of reconstruction

As described above, our coding scheme for different models of traditional Indo-European reconstruction, which we term CANONICAL, ISOLATING, and ACTIVE-STATIVE, is based on three sources: Brugmann-Delbrück (Delbrück 1893, 1897, 1900), Hirt (Hirt 1934), and Gamkrelidze-Ivanov (Gamkrelidze and Ivanov 1984; Gamkrelidze et al. 1995). We identify the values reconstructed to Proto-Indo-European by the different models for the variables in our data set to the extent that information is available, though not all variables in our data set are addressed in these sources. The different models have somewhat differing conceptualizations of the nature of the Proto-Indo-European language. Brugmann-Delbrück do not regard Proto-Indo-European as a diachronically stratified language; rather, they reconstruct a uniform language, based on Old Indo-Aryan, Greek, Latin, and other ancient Indo-European languages. Compared to the others, their model of Proto-Indo-European is simpler: they reconstruct a highly synthetic stage, which in all branches of the family becomes simplified and less synthetic, losing a number of categories. The other models have a different take on this issue. Scholars reconstructing active-stative and isolating systems presuppose that Proto-Indo-European was a language with several diachronic layers, which changed from a hypothetical early active-stative or isolating stage to a later synthetic stage, found in all ancient languages except for Anatolian. The principles and reasons for this change are important in both the isolating and active-stative models (Gamkrelidze et al. 1995:270-71; Hirt 1934:29ff.) as well as in other publications where alternative models are reconstructed for Proto-Indo-European (Bauer 2000; Pooth et al. 2018). The phylogenetic model we employ does not allow us to explicitly stratify the proto-language into layers. It allows us to reconstruct probabilities at the root as well as at ancestral nodes of the tree (see Figures 7–9 and S8); we take the root of the phylogenetic tree to represent the earliest layer of the proto-language in the sources mentioned before, but do not consider any subsequent changes or areal differentiations within Proto-Indo-European. We use a tree sample that is compatible with the Indo-Anatolian hypothesis (S9); the root represents the earliest layer of Proto-Indo-European in all models, before Anatolian split of and subsequent changes began in the Anatolian and Non-Anatolian sub-branches.

The values reconstructed to Proto-Indo-European for each model are found in S3a. We assess the extent to which our results agree with the views of each model on the basis of the likelihood of each model's reconstructed values for each variable in our data set (where applicable), i.e., the probability with which our model reconstructs the value to Proto-Indo-

European for the variable in question. These likelihoods are found in Figure 13; higher values indicate greater agreement. Whereas the active-stative and isolating models differ from our results for some of the domains discussed above, such as nominal and verbal morphology (isolating) or alignment (active-stative), both of these models come close to our reconstruction for other domains such as word order, and in that future tense and definiteness are absent in the reconstructions. At the same time, it is clear that our results show the most agreement with the canonical model of reconstruction (median likelihood = .796), followed by the active-stative (median likelihood = 0.657) and isolating (median likelihood = 0.652) models; agreement with the canonical model is significantly higher than both other models according to a pairwise Wilcoxon signed-rank test for paired samples ($p < 0.01$ with Benjamini-Hochberg correction for multiple comparisons; variables for which a reliable value was not found for all models are excluded). This indicates that our results most clearly resemble the canonical model of Proto-Indo-European, close to the reconstruction outlined by Brugmann and Delbrück in the nineteenth century.

4 Results: evolutionary dynamics

4.1 Reconstructed probabilities, feature distributions, and transition rates

Our reconstructions are estimated from transition rates inferred on the basis of our tree sample and the features in our data set; these rates characterize the behavior of pairwise transitions between all values of each variable in our data set. Specifically, a transition rate represents the average number of times that a change from a value x (e.g., SOV main clause word order) to a value y (e.g., SVO main clause word order) occurs within a 1000-year span. In this section, we assess the extent to which the frequencies of individual features, as well as transition rates pertaining to them transition rates of individual features — or as a proxy, their diachronic stability or instability — influence the reconstructions produced by our phylogenetic model. It may be the case that only highly stable and frequent features have a chance of being reconstructed to the proto-language with high probability, and less frequent features or features in greater flux are unlikely to be reconstructed. If our model essentially carries out a majority rules-style method of reconstruction, then its utility is severely diminished, as a phylogenetic model is not needed to reconstruct the most frequent pattern. If, however, it picks up on more nuanced patterns of change and incorporates these

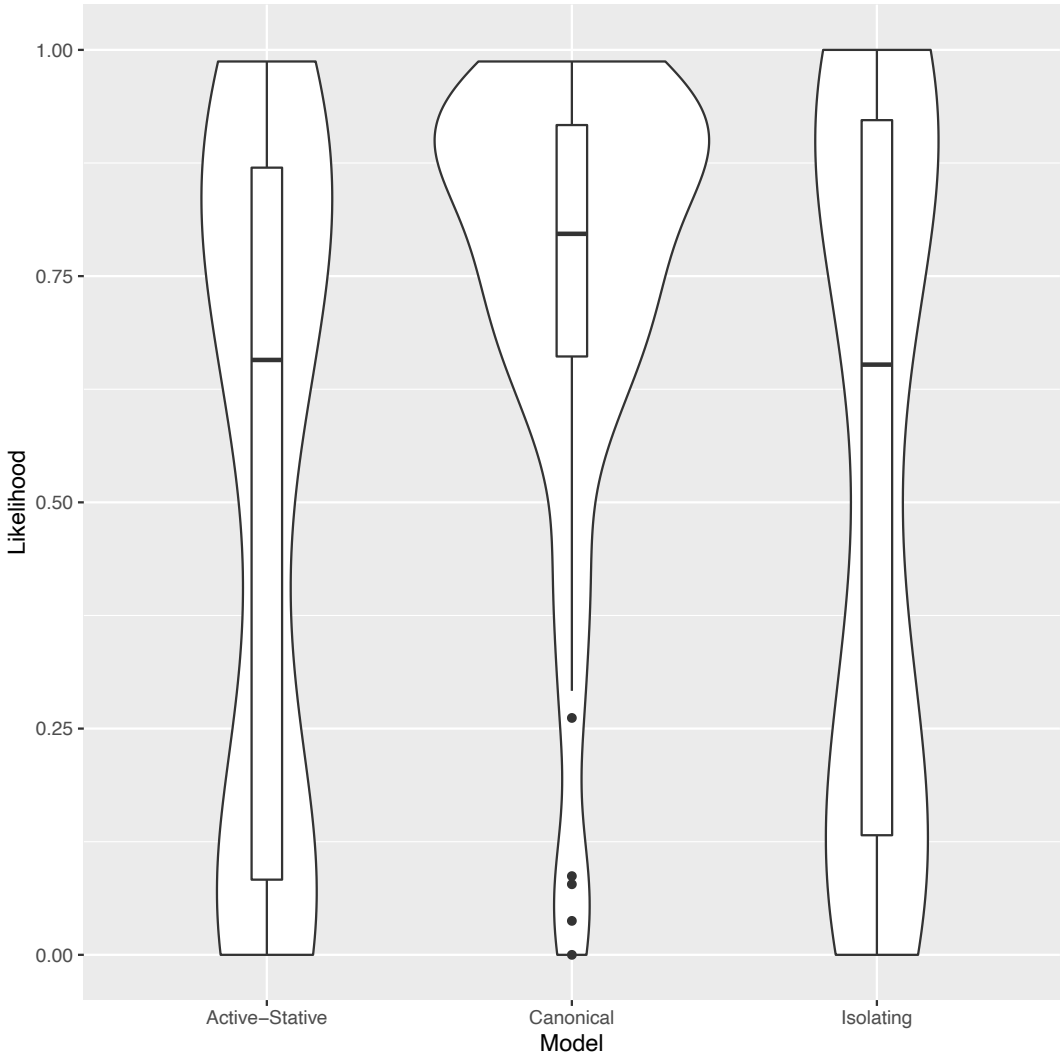


Figure 13: Violin plot indicating the absolute difference between our achieved probability of traits at the root of the phylogenetic tree and the reconstructed estimation of presence (1/0) of the comparative-historical models, labeled 'canonical', 'isolating' and 'active-stative'. Higher values signify greater agreement.

dynamics into the reconstructions it produces, then the value of methods of this sort is evident. Additionally, an analysis of rates of change alongside reconstruction probabilities provides a better understanding of temporal dynamics within Indo-European.

4.1.1 Does our model simply reconstruct the most frequent feature?

If our model simply reconstructs via a majority-rules approach, then there is no real reason to use a phylogenetic model, since our method attends only to the frequency distributions of features across languages, and not patterns of genetic relatedness between said languages. Furthermore, if this is the case, then certain Anatolian features, if they are rare within Indo-European, face a natural disadvantage and will not be reconstructed, which leads to a result in line with the canonical model of Indo-European reconstruction.

In order to assess the degree of sensitivity of our reconstructions to the frequency distributions of features in our data set, we compute the relative frequency in our data set for features reconstructed with highest probability by our model. We carry out this procedure for all languages, as well as only the ancestral languages in our sample. Figure 14 shows the probability with which 'winning' features are reconstructed plotted against their probability in our data set. Visually, it is clear that there is no strong relationship between these quantities, and the correlation between them is not significant (all languages: Spearman's $\rho = 0.117, p = 0.332$; ancestral languages: $\rho = 0.059, p = 0.622$). This indicates that there is no support for the idea that our system's reconstructions are sensitive to the distribution of features within our data set or within more archaic ancestral languages for that matter. Certain highly frequent features are not necessarily reconstructed, if our system infers that the feature is likely to have come about many times in parallel. Additionally, an infrequent feature may be reconstructed if is more likely to have survived into certain languages than come about in parallel.

A few concrete examples serve to exemplify the consequences of this behavior, with respect to the reconstruction of features found in Anatolian to Proto-Indo-European. Figure 7 shows that a masculine/feminine gender distinction is reconstructed by our model to Proto-Indo-European, despite the fact that it is absent in the archaic Anatolian subgroup; because the gender distinction is predominant in Nuclear Indo-European (i.e., all non-Anatolian languages) and because it has been lost several times, our model assigns high probability to a scenario where Anatolian lost the gender distinction during its development (our model also assigns a small amount of probability to a scenario in which the gender distinction was lost independently within Anatolian). In Figure 8, we see that neuter gender is reconstructed to

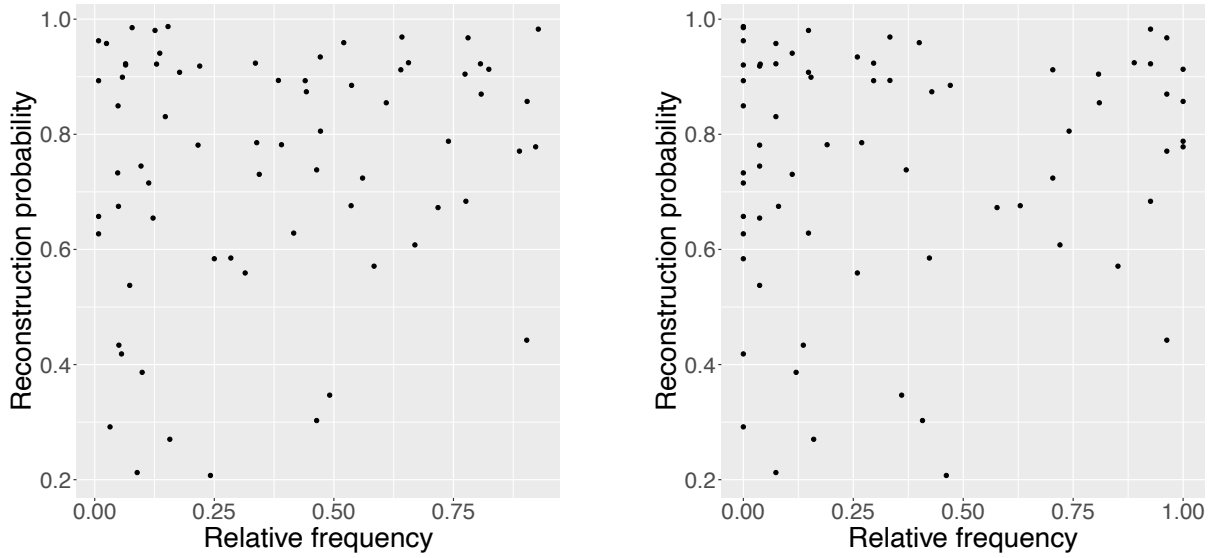


Figure 14: Reconstruction probabilities of features reconstructed with maximum probability to PIE, plotted according to their relative frequency across all languages (left) and ancestral languages (right) in the data set

Proto-Indo-European because it is shared by a substantial number of Core Indo-European (i.e., all of Indo-European to the exclusion of Anatolian and Tocharian) languages and Anatolian, to the exclusion of Tocharian. In Figure 9, we see that our system does not reconstruct prepositions to Proto-Indo-European with high probability; incidentally, although prepositions are reconstructed with high probability for Core Indo-European, our model finds it unlikely that prepositions were lost independently in the history of Anatolian as well as Tocharian (but finds it likely that they were lost on some Indo-Iranian lineages). Hence, we see that Anatolian features have a good chance of being reconstructed to Proto-Indo-European if they are reconstructed to the proto-language of at least one other higher-order or archaic Indo-European branch.

4.1.2 Transition rates

Given the transitions between each pair of values (e.g., ergative → accusative) within a variable, we estimate the overall ENTRY or GAIN rate and the overall EXIT or LOSS rate for individual features according to the formulae given in the Appendix. The full list of inter-feature transition rates as well as the gain and loss rates for each feature are found in S5–6. A scatterplot of gain and loss rates for each feature, organized according to overarching grammatical categories, is found in Figure 15. The size of individual data points indicates

the probability with which a given feature is reconstructed to Proto-Indo-European by our model, with larger size indicating higher probability. The plot is divided according to the median gain and loss rates for our variables; this allows us to divide features into the following four classes of features:

1. High gain rate, high loss rate (upper right quadrant): features of high instability, in frequent flux, gained and lost frequently. These include features pertaining to the presence of case on adjectives, clitics, distinctions between dative and genitive marking, absence of case on nouns, and different alignment systems in the simple past.

2. High gain rate, low loss rate (lower right quadrant): features of high stability to which languages are frequently attracted; gained often and rarely lost. These include features pertaining to the presence of case on nouns, case difference between A and O for nouns as well as pronouns, masculine/feminine distinction, Noun-Relative word order, Possessor-Noun word order, Present progressive by auxiliary, and absence of neuter gender and vocative case.

3. Low gain rate, high loss rate (upper left quadrant): 'recessive' features (cf. Nichols 1993) quickly repulsed by languages when they do occur. These include features pertaining to the presence of future tense by participle, future tense by particle, more than seven cases, more than seven pronominal cases, more than five genders, tripartite alignment, ergative alignment in pronouns, active-stative alignment, double oblique alignment, V2 word order, and VSO word order.

4. Low gain rate, low loss rate (lower left quadrant): highly stable features that arise infrequently. These include features pertaining to the presence of adjective-noun word order, agglutination for case, agreement on prepositions, case on the last member of a NP, definite articles, definite suffixes on adjectives, definite suffixes on nouns, neuter gender, a noun class for animates, and synthetic future tense.

We follow the literature on the diachronic dynamics underlying typological stability and instability (cf. Greenberg 1978; Nichols 2003; Dediu 2010) in our interpretation of these patterns.

For two of the classes, patterns of reconstruction are highly consistent. Stable features that are frequently gained and infrequently lost (lower right quadrant) are virtually always reconstructed with high certainty. Recessive features that are rarely gained and frequently

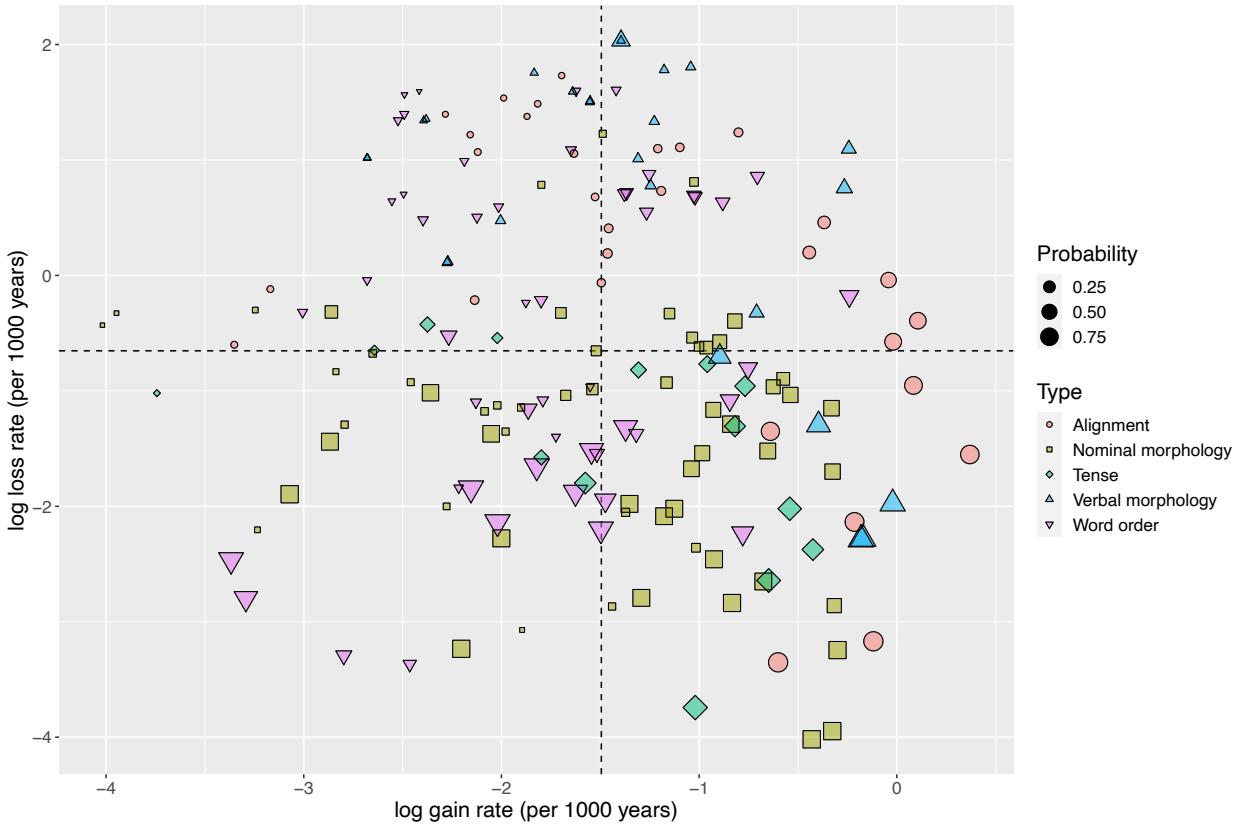


Figure 15: Features in the data set plotted according to their log gain and log loss rates. Shapes represent grammatical categories to which features belong; sizes represent the probability with which features are reconstructed to PIE.

Gain rate	Loss rate	Interpretation	Reconstruction probability
High	High	Unstable trait in flux	Low/High
High	Low	Stable, attractor feature	High
Low	High	Recessive feature	Low
Low	Low	Infrequent but stable feature	Low/High

Table 6: Interpretation of behaviour patterns of traits depending on probability of presence at the proto-language state and gain/loss rate

lost (upper left quadrant) are consistently reconstructed with low probability. For the remaining classes, the behavior of our model is more variable. Infrequently gained but stable features (lower left quadrant) are reconstructed with both high and low probability. For instance, SOV word order (in both main and subordinate clauses), adjective-noun word order, neuter gender, postpositions, and vocative case are reconstructed with high probability, whereas features pertaining to definiteness and agglutination, as well as SVO word order (in both main and subordinate clauses) are reconstructed with low probability. The same variability is found for features that fluctuate (upper right quadrant): features pertaining to nominative-accusative alignment, the genitive/dative distinction, and case on adjectives are reconstructed with probability greater than .5, whereas features pertaining to case mergers, clitic word order and the presence of ergativity in the simple past tense are reconstructed with low probability.

It is noteworthy that features pertaining to nominal morphology and tense (inflectional typology) tend to exhibit slow change and features pertaining to alignment and verbal morphology (agreement) show more rapid patterns of change. Word order is well represented among swift-changing features. These results partly confirm that traits that are immediately bound by morphology, such as nominal morphology and tense, have slower rates of change, in contrast to traits that are not bound by morphology, such as word order and alignment, which have swifter rates of change. Verbal morphology, i.e., agreement patterns, as well as the most stable and frequent word orders, constitute an exception to these tendencies.

All in all, these results show that the reconstructions produced by our model are not simply artifacts of feature distributions across our data set; on the contrary, they reflect multifaceted patterns of change; for certain types of stability or instability, features are reconstructed with either high or low probability, while for other patterns, there is more variability among reconstruction probabilities. These patterns are summarized in Table 6.

4.2 Transition rates and grammatical hierarchies

Earlier, we mentioned the existence of asymmetries in the certainty with which features are reconstructed to Proto-Indo-European that correspond to differences in grammatical hierarchies. This asymmetry can be found also in transition rates pertaining to the features in questions. As described in §2.3.2, we organize our features into hierarchical pairs which belong to the same grammatical category but which vary with respect to other features of the grammar. The categories we identify in our data are restricted to the following features (UNMARKED/MORE FREQUENT < MARKED/LESS FREQUENT; see Table 2):

PRONOUN < NOUN

PRESENT < PAST

AGENT < OBJECT

AGENT/OBJECT < OBLIQUE

MASCULINE/FEMININE < NEUTER

Since loss rates are an inverse measure of the longevity of a given feature (i.e., shorter-lived features are lost at a higher rate), we measure whether the loss rates differ significantly across the unmarked/marked feature pairs that we identify in our data set. We find that the loss rates of marked traits are higher than those of unmarked traits ($V = 851, p < 0.001$, according to a one-sided Wilcoxon signed-rank test), indicating that marked traits are lost more frequently than unmarked, more frequent traits.

This is an interesting result, but it is not entirely unexpected. The idea of grammatical or marking hierarchies in the traditional sense (Comrie 1981; Croft 1990; Greenberg 1966) is based on the notion that higher-ranking categories as a rule are more frequent in languages. The idea that more frequent and basic categories, both in grammar and lexicon, are more conservative and archaic, due to their everyday use, has a long history in Indo-European linguistics (Meillet 1948:135). Many lexemes, which are typically part of Swadesh lists, such as kinship words, body parts, numerals, fire, water, and so forth, as a rule preserve more archaic paradigms, including change in stem consonants (e.g., *-r-/-n-, -l-/-n-*) or ablauting patterns (qualitative, quantitative) (Meier-Brügger et al. 2010:336-48). The reflexes in daughter languages of the most frequent verbs, such as PIE *$h_1$es-* 'to be' or PIE *$h_1$ey-* 'to go', are typically irregular, preserving archaic inflection patterns and categories (Rix and Kümmel 2001:232-33, 41-42). On the other hand, analogy and other types of changes that

harmonize and simplify language structures, making them easier to memorize, are more frequently found among words and categories of lower frequency. By means of phylogenetic methods, we know that there is, at least in the lexicon of basic vocabulary, a correlation between frequency and substitution rates: the most frequent meanings are concepts with generally lower substitution rates (Pagel et al. 2007). Transferred to a scenario of grammatical hierarchies, we expect the unmarked categories, representing the more frequently used categories, to have lower loss rates and longer periods between transitions, whereas we expect the marked categories, representing less frequent categories, to have higher loss rates and shorter periods between transitions.

5 Discussion

In the preceding sections, we presented the results of Proto-Indo-European reconstruction using phylogenetic comparative methods, and provided a careful analysis of our model's behavior in order to better understand the mechanisms underlying the results that it produces. We found broad support for the canonical model of Indo-European syntactic reconstruction, largely because the features reconstructed under alternative models undergo evolutionary dynamics that make them unlikely to survive into the languages that attest them; they are more likely to have been innovated in parallel. The behavior of the model we use rests on the assumption that we can make inferences about the behavior of linguistic features in prehistory on the basis of their behavior during attested history; calibrating these dynamics according to attested patterns of change is made possible by the use of ancestry constraints. The methodology that we use relies on a number of simplifying assumptions regarding the nature of change. One of these is the assumption of RATE UNIFORMITY, namely that rates of change between values of a linguistic variable are the same on all lineages of Indo-European. There are a number of methods which relax this assumption in order to incorporate RATE VARIATION or HETEROTACHY (Tuffley and Steel 1998; Heath et al. 2011; Beaulieu and O'Meara 2014), but the utility of these methods may be limited relative to their increased computational complexity; phylogenetic linguistic analyses assuming rate homogeneity dovetail well with independent evidence (Blasi et al. 2019), and incorporating heterotachy produces no or little improvement (Chang et al. 2015; Blasi et al. 2020). Our data consist of typological variables rooted in comparative concepts that can be easily operationalized. Other approaches to syntactic reconstruction, rooted in particular syntactic theories, have different assumptions regarding the levels of representation that should be

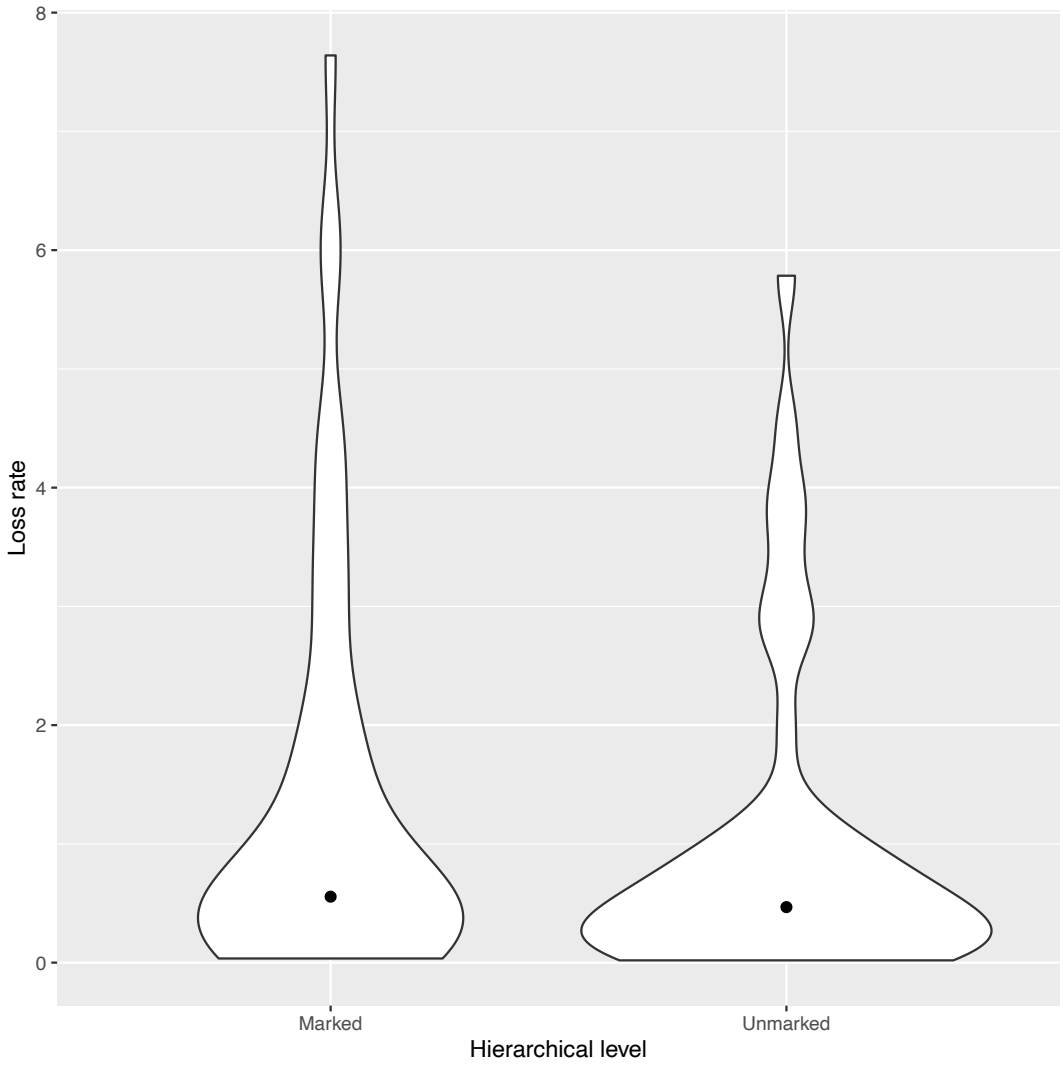


Figure 16: Distributions of loss rates for marked (more frequent) and unmarked (less frequent) features. Median values are indicated by dots. Unmarked (more frequent) traits are lost on average less frequently than marked (less frequent) ones.

reconstructed to the proto-language (Hale and Kissock 2015); at the same time, similar work makes use of bioinformatic algorithms to address questions of linguistic prehistory (Longobardi et al. 2013). Furthermore, certain syntactic theories make strong predictions regarding syntactic change over many generations of first language acquisition (Berwick and Niyogi 1996; Yang 2000) which, if correct, can potentially find support in phylogenetic models. While some of the assumptions we make may be challenged by other scholars, both in terms of methodology and the nature of the data we employ, we believe that our contribution is of great utility to Indo-European studies, as the assumptions of our probabilistic model are explicit, the results we present are replicable, and we provide a means of explicit evaluation against the hypotheses explored in this paper.

Our results shed light on interesting cross-linguistic diachronic tendencies. At the same time, our data are confined to one family, Indo-European, and the problem of embracing uniformitarianism within one family leads to the problem of sample diversity pointed out by Levy and Daumé (2011). When investigating differences in diachronic dynamics across hierarchies, we build upon insights from a cross-linguistic sample while assessing results derived from one family only, which is not possible without adapting a uniformity-of-state framework to language evolution: rules that govern language structure are similar in the present and in the past, and all languages reflect some basic universal principles (Croft 2003:233; Roberts 2007:174; Walkden 2019). We are forthcoming in the admission that restricting data to one family gives a limited picture. Some of the observed patterns are obviously uniquely Indo-European, such as the loss of synthetic structure. At the same time, the markedness rubric we employ is derived from cross-linguistic typological observations beyond the scope of Indo-European, allowing us to avoid circular reasoning. On the basis of trends within Indo-European, we confirm broader hypotheses regarding cross-linguistically universal tendencies, namely that factors such as economy and frequency interfere in the processes of language evolution (Croft 2003; Haspelmath 2015). Furthermore, our finding that features pertaining to unmarked, more frequent categories are lost more frequently than those of marked, less frequent categories dovetails nicely with the common finding that frequently used items are resistant to various types of change (e.g., Diessel 2007).

It is remarkable how close our comparative phylogenetic reconstruction approaches to a canonical reconstruction model of Indo-European syntax (Delbrück 1893, 1897, 1900; Krahe et al. 1972). Despite all the variation and change in the grammar in the Indo-European family, our model reconstructs a highly synthetic, mainly head-final language, with nominative-accusative alignment, independent of tense and animacy degree of the first argument, case

marking on nouns, no definite article, three genders (masculine, feminine, neuter), predicative gender agreement, a non-agglutinating case system with less than seven cases but with a nominative, accusative, dative, genitive, and vocative, also in pronouns, a synthetic present, no future, full agreement in present tense of verbs, but not in the past tense, postpositions, OV infinitive word order, SOV in main and subordinate clauses, possessor-noun, adjective-noun, noun-relative clause, OV participle word order, and WH-initial word order. The outcome is striking. We see the structure of a grammatical system which has been retained to a high degree through many branches of Indo-European, and which is remarkably constant, despite several millennia of language contact and change, loss of categories, emergence of new categories by grammaticalization, and substantial typological changes in, e.g., word order patterns.

Alternative models (see 1.1, 2.3.1) assume far-reaching typological changes between Proto-Indo-European and the Non-Anatolian sub-branches of the tree. These changes are supported by internal reconstruction based on Proto-Indo-European paradigmatic correlations in combination with a comparison with other unrelated language families (see §3). Given a nuanced understanding of Indo-European chronology (Bouckaert et al. 2012; Chang et al. 2015; Meid 1975; Schlerath 1981), as well as both attested as well as estimated information regarding the timespans characterizing change between typological features of the type that we investigate here (Croft 2003:252; Haspelmath 2018; Hock and Joseph 1996:183-184), it is increasingly clear why there is limited support for the alternative theories. On the basis of what can be inferred regarding change between alignment systems between languages in our sample, for instance, a relatively rapid development form ergative or active-stative alignment (as assumed by the active-stative model) or via grammaticalization from an isolating system with as yet undeveloped agreement marking (as assumed by the isolating model) to a nominative-accusative system is less likely than the retention of nominative-accusative alignment. All these models take into account a large amount of data (morphological, typological) that may be connected directly or indirectly to the typological traits investigated in our study.

In the realm of alignment, nominative-accusative alignment is dominant in many contemporary and most historical states of the family (Carling 2019:31-50), and also reconstructed to the proto-language (Figure 5). Absence of agreement marking (features involving NO MARKING) which implies isolating structure, is reconstructed to the proto-language with low probability. In conclusion, nominative-accusative alignment is stable and the only noteworthy trend over time is a higher rate of transition from nominative-accusative to no marking

(SYNTHETIC > ISOLATING). Development of other systems (ergative, active-stative, tripartite) from nominative-accusative are marginal and have low transition rates. Similar patterns can be seen across other data types: in general, within Indo-European, there is greater evidence for developments in the directions SYNTHETIC > ISOLATING and SYNTHETIC > AGGLUTINATING. Developments in the opposite direction are not impossible, but are unlikely to have taken place between Proto-Indo-European and its descendants. An exception is word order, where features exhibit varying degrees of stability and instability.

Our results are of key relevance to larger discussions regarding typological stability, as well as the suitability of typological data for language classification (Dediu and Levinson 2012; Dediu and Cysouw 2013; Dunn et al. 2011; Plank 2011). We observe two overarching feature classes characterized by different patterns of change. Either the change is slow and overwhelmingly unidirectional, moving from synthetic to isolating, which is found mainly in the paradigmatic categories, i.e., nominal morphology and parts of verbal morphology, where a synthetic system is broken down in the direction of an isolating system. This type of change conforms to a model of a unidirectional, cyclic typological change, occurring at a slow change rate (Croft 2003:227ff.). Alternatively, the evolutionary trend is oscillating, with high amounts of gains and losses. This occurs in the syntactic (non-paradigmatic) categories, mainly by word order, alignment, and partly by verbal morphology traits. This type generally conforms to a theory of a punctuated and non-directional change, which may take any direction depending on a combination of internal pressure and areality-induced change (Dixon 1997). Given this result, the search for a phylogenetic signal in the evolution of non-homologous, structural linguistic features is difficult without considering areality and ancient language data.

Finally, we employed a relatively uncontroversial and neutral model of change that has been used in state-of-the art work in phylogenetic linguistics. Simple models like the Continuous-Time Markov Process assume that a feature can be born and die (or that change between traits of a multistate character can occur) with a given rate, and there are no a priori restrictions on the values taken by these rates, as long as they are positive. These models are appropriate for grammatical and lexical data, and are suitable in situations when the system may leave and return to states, as in the case of variants of word order, agreement, case, or different lexical meanings. More complex models like the stochastic dollo character (SDC; Nicholls and Gray 2006), which assume that features are born only once in the history of a language family, are useful for features which cannot return to identical states, such as morphological traits bound to specific forms, irreversible outcomes of sound change (e.g.,

mergers), or features which may undergo grammaticalization. Using an SDC model would likely yield different results for our data, at least for gender. At the same time, many scholars agree that SDC models are unsuitable for not only comparative concepts, which aim to capture features cross-linguistically, independent of linguistic matter, but other types of linguistic characters (Chang et al. 2015), though modifications thought to be more appropriate for linguistic data have been proposed (Bouckaert and Robbeets 2017). However, we are confident that a Bayesian approach has the capacity beyond a comparative-historical model to contribute in a meaningful way to the theoretical discussions about trends in diachrony, directionality of syntax, rates of gains and losses, stability of features and categories, as well as correlations to important aspects of typology such as frequency, economy, hierarchies, and general trends in grammar change. Future work can potentially contrast the results of different evolutionary models in applications like the one undertaken in this paper; researchers wishing to argue for a specific evolutionary model over others (along with its concomitant result) may employ posterior predictive checks (Box 1980; see also Appendix D) to demonstrate that their model is a better fit to the data than others.

6 Conclusion

The current paper had several foci. We reconstructed the evolutionary history of selected aspects of Indo-European morphosyntax by means of a model which infers patterns of diachronic development of linguistic features over a phylogeny. This allowed us to infer the most probable value of a given linguistic variable in the unattested Proto-Indo-European language. We used a dataset of binary coded comparative concepts, recoded as categorical features, which also contained data from extinct and historical Indo-European languages. We focused on five categories of grammar: alignment, nominal morphology, verbal morphology, tense, and word order. The result at the proto-language state we compared to previous reconstructions of Proto-Indo-European grammar by means of the comparative-historical method and diachronic typology. The methodology that we used allowed us to compare ideas from the traditional comparitive-historical linguistic literature with our model's output. We found that phylogenetic reconstruction produced a consistent and coherent system, which corresponded to a highly synthetic, mainly head-final language, with nominative-accusative alignment, independent of tense and animacy degree of the first argument, case marking on nouns, no definite article, three genders (masculine, feminine, neuter), predicative gender agreement, a non-agglutinating case system with less than seven cases but with a nomina-

tive, accusative, dative, genitive, and vocative, also in pronouns, a synthetic present, no future, full agreement in present tense of verbs, but not in the past tense, postpositions, OV infinitive word order, SOV in main and subordinate clauses, possessor-noun, adjective-noun, noun-relative clause, OV participle word order, and WH-initial word order. This reconstruction matched a canonical model of Proto-Indo-European grammar, as is was described by the Neogrammarians already in the nineteenth century.

We also analyzed the inferred inter-feature transition rates on which our reconstructions are based. Our analysis shed light on different tendencies of change across features. In general, traits that were reconstructed to the proto-language had relatively low loss and gain rates, which implies that the reconstructed typological system is consistent and stable in the family. The most noteworthy tendency was a change from synthetic to isolating structure. In addition, a general tendency was for morphological (paradigmatic) categories (nominal morphology and tense) to have low change rates, whereas syntactic categories (alignment and word order) to have higher change rates. Verbal morphology opposed this tendency with high change rates. Finally, we divided our grammatical traits (excluding word order) into hierarchical pairs by different members of categories available in our data, such as tense (present, past), word class (noun, pronoun), or gender. We found that the unmarked, more frequent traits are lost less frequently than marked, less frequent traits; this difference was significant.

In sum, our result supported the theory that grammar evolution is both divergent, down to the level of highest granularity, as well as following general, universal principles. Over the 6000–7000 year cycle represented in our data, morphological traits tended to show a unidirectional path of change, fundamentally moving from synthetic to isolating, whereas word order, alignment and person agreement properties showed more non-directional and unpredictable paths of change, with higher rates of gains and losses. The results were in line both with a cyclic and a punctuated model of change. Our results also indicated that the variability of change in grammar over time is governed by general, 'universal' tendencies, such as grammatical hierarchies and frequency.

References

Aissen, J. (2003). Differential object marking: Iconicity vs. economy. *Natural Language Linguistic Theory 21*(3), 435–483.

Baerman, M. and D. Brown (2013). Syncretism in verbal person/number marking. In M. S.

Dryer and M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Baker, M. C. (2011). The interplay between universal grammar, universals, and lineage specificity. *Linguistic Typology 15*(2), 473–482.

Bammesberger, A. (1986). *Untersuchungen zur vergleichenden Grammatik der germanischen Sprachen. Bd 1, Der Aufbau des germanischen Verbalsystems*. Heidelberg: Winter.

Barðdal, J. (2014). Syntax and syntactic reconstruction. In C. Bowern and B. Evans (Eds.), *The Routledge Handbook of Historical Linguistics*, pp. 343–373. London - New York: Routledge.

Barðdal, J. and T. Eythórsson (2009). The origin of the oblique subject construction: an Indo-European comparison. In V. Bubenik, J. Hewson, and S. Rose (Eds.), *Grammatical change in Indo-European languages*, pp. 179–193. Amsterdam-Philadelphia: John Benjamins.

Bates, D., M. Mächler, B. Bolker, and S. Walker (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software 67*(1), 1–48.

Bauer, B. (2000). *Archaic syntax in Indo-European : the spread of transitivity in Latin and French*. Trends in linguistics. Studies and monographs, 99-0115958-X ; 125. Berlin: Mouton de Gruyter.

Bauer, B. (2007). The definite article in Indo-European. Emergence of a new grammatical category? In E. Stark, E. Leiss, and W. Abraham (Eds.), *Nominal Determination. Typology, context constratis, and historical emergence*, pp. 103–139. Amsterdam-Philadelphia: John Benjamins.

Bauer, B. L. M. (1995). *The Emergence and Development of Svo Patterning in Latin and French: Diachronic and Psycholinguistic Perspectives*. Cary, UNITED STATES: Oxford University Press, Incorporated.

Beaulieu, J. M. and B. C. O'Meara (2014). Hidden Markov models for studying the evolution of binary morphological characters. In L. Z. Garamszegi (Ed.), *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology: Concepts and Practice*, pp. 395–408. Heidelberg, New York, Dordrecht, London: Springer.

Berwick, R. C. and P. Niyogi (1996). Learning from triggers. *Linguistic Inquiry*, 605–622.

Bickel, B. (2007). Typology in the 21st century: Major current developments. *Linguistic Typology 11*(1), 239–251.

Bickel, B. (2008). On the scope of the referential hierarchy in the typology of grammatical relations. In G. Corbett Greville and M. Noonan (Eds.), *Case and Grammatical Relations. Studies in honor of Bernard Comrie*, pp. 191–210. Amsterdam - Philadelphia: John Benjamins.

Bickel, B. (2011). Statistical modeling of language universals. *Linguistic Typology 15*(2), 401–413.

Bickel, B. and J. Nichols (2002). Autotypologizing databases and their use in fieldwork. In P. Austin, H. Dry, and P. Witternburg (Eds.), *Proceedings of the International LREC Workshop on Resources and Tools in Field Linguistics, Las Palmas, 26 - 27 May 2002*. Nijmegen: ISLE and DOBES.

Blasi, D. E., S. Moran, S. Moisik, P. Widmer, D. Dediu, and B. Bickel (2020). Languages, evolution and statistics: human sound systems were shaped by changes in bite configuration. Response to Tarasov & Uyeda (2020). *BioRxiv*.

Blasi, D. E., S. Moran, S. R. Moisik, P. Widmer, D. Dediu, and B. Bickel (2019). Human sound systems are shaped by post-Neolithic changes in bite configuration. *Science 363*(6432).

Bollback, J. P. (2006). SIMMAP: Stochastic character mapping of discrete traits on phylogenies. *BMC Bioinformatics 7*, 88.

Bopp, F. (1816). *Über das Conjugationssystem der Sanskritsprache*. Frankfurt am Main: Andreas.

Bornkessel-Schlesewsky, I., A. L. Mal čukov, and M. Richards (2015). *Scales and hierarchies : a cross-disciplinary perspective*. Berlin: De Gruyter Mouton.

Bouckaert, R., P. Lemey, M. Dunn, S. J. Greenhill, A. V. Alekseyenko, A. J. Drummond, R. D. Gray, M. A. Suchard, and Q. D. Atkinson (2012). Mapping the origins and expansion of the indo-european language family. *Science 337*(6097), 957–960.

Bouckaert, R. R. and M. Robbeets (2017). Pseudo Dollo models for the evolution of binary characters along a tree. *BioRxiv*, 207571.

Box, G. E. P. (1980). Sampling and Bayes' inference in scientific modelling and robustness. *Journal of the Royal Statistical Society. Series A (General) 143*, 383–430.

Brugmann, K., B. Delbrück, and B. Delbrück (1893). *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen : kurzgefasste Darstellung der Geschichte des Altindischen, Altiranischen (Avestischen u. Altpersischen), Altarmenischen, Altgriechischen, Albanesischen, Lateinischen, Oskisch-Umbrischen, Altirischen, Gotischen, Althochdeutschen, Litauischen und Altkirchenslavischen. Bd 3, Vergleichende Syntax der indogermanischen Sprachen, T. 1.* Strassburg: Trübner.

Brugmann, K., B. Delbrück, and B. Delbrück (1897). *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen : kurzgefasste Darstellung der Geschichte des Altindischen, Altiranischen (Avestischen u. Altpersischen), Altarmenischen, Altgriechischen, Albanesischen, Lateinischen, Oskisch-Umbrischen, Altirischen, Gotischen, Althochdeutschen, Litauischen und Altkirchenslavischen. Bd 4, Vergleichende Syntax der indogermanischen Sprachen, T. 2.* Strassburg: Trübner.

Calude, A. S. and A. Verkerk (2016). The typology and diachrony of higher numerals in indo-european: a phylogenetic comparative study. *Journal of Language Evolution 1*(2), 91.

Campbell, L. and A. C. Harris (2002). Syntactic reconstruction and demythologizing 'myths and the prehistory of grammars'. *Journal of Linguistics 38*(3), 599.

Carling, G. (2012). Development of form and function in a case system with layers: Tocharian and Romani compared. *Tocharian and Indo-European Studies 13*, 57–76.

Carling, G. (2019). *Mouton Atlas of Languages and Cultures. Vol. 1: Europe and West, Central and South Asia.* Berlin - Boston: Mouton de Gruyter.

Carling, G., F. Larsson, C. Cathcart, N. Johansson, A. Holmer, E. R. Round, and R. Verhoeven (2018). Diachronic Atlas of Comparative Linguistics (DiACL) – A Database for Ancient Language Typology. *PLOS ONE 13(10).*

Carpenter, B., A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, and A. Riddell (2017). Stan: A probabilistic programming language. *Journal of statistical software 76*(1).

Cathcart, C., G. Carling, F. Larsson, N. Johansson, and E. R. Round (2018). Areal pressure in grammatical evolution. *Diachronica 35*(1), 1–34.

Cathcart, C. A. (2018). Modeling linguistic evolution: a look under the hood. *Linguistics Vanguard 1*.

Cathcart, C. A., A. Hölzl, G. Jäger, P. Widmer, and B. Bickel (2020). Numeral classifiers and number marking in indo-iranian: a phylogenetic approach. *Language Dynamics and Change*, 1 – 53.

Chang, W., C. Cathcart, D. Hall, and A. Garrett (2015). Ancestry-constrained phylogenetic analysis supports the Indo-European steppe hypothesis. *Language 91*(1), 194–244.

Clackson, J. (2007). *Indo-European linguistics : an introduction*. Cambridge textbooks in linguistics, 99-0104661-0. Cambridge, UK ;: Cambridge University Press.

Comrie, B. (1981). *Language universals and linguistic typology : syntax and morphology*. Oxford: Blackwell.

Croft, W. (1990). *Typology and universals*. Cambridge textbooks in linguistics, 99-0104661-0. Cambridge: Cambridge University Press.

Croft, W. (2003). *Typology and universals*. Cambridge textbooks in linguistics, 99-0104661-0. Cambridge: Cambridge Univ. Press.

Croft, W., T. Bhattacharya, D. Kleinschmidt, D. E. Smith, and T. F. Jaeger (2011). Greenbergian universals, diachrony, and statistical analyses. *Linguistic Typology 15*(2), 433–453.

Cysouw, M. (2011). Understanding transition probabilities. *Linguistic Typology 15*(2), 415–431.

Dediu, D. (2010). A Bayesian phylogenetic approach to estimating the stability of linguistic features and the genetic biasing of tone. *Proceedings of the Royal Society of London B 278*(1704), 474–9.

Dediu, D. and M. Cysouw (2013). Some structural aspects of language are more stable than others: A comparison of seven methods. *PLOS ONE 8*(1), e55009.

Dediu, D. and S. C. Levinson (2012). Abstract profiles of structural stability point to universal tendencies, family-specific factors, and ancient connections between languages. *PLoS ONE 7*(9), e45198.

Delbrück, B. (1893). *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen : kurzgefasste Darstellung der Geschichte des Altindischen, Altiranischen (Avestischen u. Altpersischen), Altarmenischen, Altgriechischen, Albanesischen, Lateinischen, Oskisch-Umbrischen, Altirischen, Gotischen, Althochdeutschen, Litauischen und Altkirchenslavischen. Bd 3, Vergleichende Syntax der indogermanischen Sprachen, T. 1.* Strassburg: Trübner.

Delbrück, B. (1897). *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen : kurzgefasste Darstellung der Geschichte des Altindischen, Altiranischen (Avestischen u. Altpersischen), Altarmenischen, Altgriechischen, Albanesischen, Lateinischen, Oskisch-Umbrischen, Altirischen, Gotischen, Althochdeutschen, Litauischen und Altkirchenslavischen. Bd 4, Vergleichende Syntax der indogermanischen Sprachen, T. 2.* Strassburg: Trübner.

Delbrück, B. (1900). *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen : kurzgefasste Darstellung der Geschichte des Altindischen, Altiranischen (Avestischen u. Altpersischen), Altarmenischen, Altgriechischen, Albanesischen, Lateinischen, Oskisch-Umbrischen, Altirischen, Gotischen, Althochdeutschen, Litauischen und Altkirchenslavischen. Bd 5, Vergleichende Syntax der indogermanischen Sprachen, T. 3.* Strassburg: Trübner.

Delbrück, B. (2010). *Vergleichende Syntax der indogermanischen Sprachen*, Volume 2 of *Cambridge Library Collection - Linguistics*. Cambridge: Cambridge University Press.

Diessel, H. (2007). Frequency effects in language acquisition, language use, and diachronic change. *New ideas in psychology 25*(2), 108–127.

Dixon, R. M. W. (1997). *The Rise and Fall of Languages [Elektronisk resurs]*.

Donohue, M. (2011). Stability of word order: Even simple questions need careful answers. *Linguistic Typology 15*(2), 381–391.

Dryer, M. and M. Haspelmath (Eds.) (2013). *World Atlas of Linguistic Structures*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Dryer, M. S. (1992). Greenbergian word order correlations. *Language 68*(1), 81–138.

Dryer, M. S. (2011). The evidence for word order correlations. *Linguistic Typology 15*(2), 335–380.

Dunn, M., T. K. Dewey, C. Arnett, T. Eythórsson, and J. Barðdal (2017). Dative sickness: A phylogenetic analysis of argument structure evolution in Germanic. *Language: Journal of the Linguistic Society of America 93*(1), e1–e22.

Dunn, M., S. J. Greenhill, S. C. Levinson, and R. D. Gray (2011). Evolved structure of language shows lineage-specific trends in word-order universals. *Nature 473*(7345), 79–82.

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution 17*(6), 368.

Felsenstein, J. (2004). *Inferring phylogenies*. Sunderland, Mass.: Sinauer.

Ferraresi, G. and M. Goldbach (2008). *Principles of syntactic reconstruction*. Current issues in linguistic theory, 0304-0763 ; 302. Amsterdam ;: John Benjamins Pub. Company.

Friedrich, P. (1975). *Proto-Indo-European syntax : the order of meaningful elements*. Journal of Indo-European studies. Monograph, 99-0123417-4 ; 1. Butte, Mont.

Gamkrelidze, T. V. and V. V. Ivanov (1984). *Indoevropejskij jazyk i indoevropejcy : rekonstrukcija i istoriko-tipologičeskij analiz prajazyka i protokul tury = Indo-European and the Indo-Europeans : a reconstruction and historical typological analysis of a protolanguage and a proto-culture*. Tbilisi: Izd. Tbilisskogo univ.

Gamkrelidze, T. V., V. V. Ivanov, and W. Winter (1995). *Indo-European and the Indo-Europeans : a reconstruction and historical analysis of a proto-language and a proto-culture*. Trends in linguistics. Studies and monographs, 99-0115958-X ; 80. Berlin: Mouton de Gruyter.

Garrett, A. (2008). Paradigmatic uniformity and markedness. In J. Good (Ed.), *Explaining linguistic universals: Historical convergence and universal grammar*, pp. 124–143. Oxford: Oxford University Press.

Greenberg, J. H. (1963). *Universals of language : report of a conference held at Dobbs Ferry, New York, April 13-15, 1961*. Cambridge, Mass.: MIT Press.

Greenberg, J. H. (1966). *Language universals : with special reference to feature hierarchies*. Janua linguarum: Series minor 59. The Hague : Mouton, 1966.

Greenberg, J. H. (1978). *Universals of human language*. Stanford: Stanford Univ. Press.

Greenberg, J. H. and M. Haspelmath (2005). *Language Universals [Elektronisk resurs] : With Special Reference to Feature Hierarchies*. Berlin ;New York: De Gruyter.

Hale, M. and M. Kissock (2015). Syntactic reconstruction: The Correspondence Problem revisited. Paper presented at the 17th Diachronic Generative Syntax Conference (DIGS 17), Reykjavik, May 31.

Harris, A. C. (2008). Reconstruction in syntax. reconstruciton of patterns. In G. Ferraresi and M. Goldbach (Eds.), *Principles of Syntactic Reconstruction*. Amsterdam - Philadelphia: John Benjamins.

Harris, A. C. and L. Campbell (1995). *Historical syntax in cross-linguistic perspective*. Cambridge studies in linguistics, 0068-676X ; 74. Cambridge: Cambridge Univ. Press.

Haspelmath, M. (2006). Against markedness (and what to replace it with). *Journal of Linguistics 42*(1), 25–70.

Haspelmath, M. (2010). Comparative concepts and descriptive categories in crosslinguistic studies. *Language 86*(3), 663–687.

Haspelmath, M. (2015). Descriptive scales versus comparative scales. In I. Bornkessel-Schlesewsky, A. Malchukov, and M. D. Richards (Eds.), *Scales and Hierarchies. A Cross-Disciplinary Perspective*, pp. 45–58. Berlin - Boston: De Gruyter.

Haspelmath, M. (2018). Revisiting the anasynthetic spiral. In H. Narrog and B. Heine (Eds.), *Grammaticalization from a Typological Perspective*. Oxford: Oxford University Press.

Heath, T., M. Holder, and J. Huelsenbeck (2011, 11). A Dirichlet Process Prior for estimating lineage-specific substitution rates. *Molecular Biology and Evolution 29*, 939–55.

Hewson, J. and V. Bubeník (1997). *Tense and aspect in Indo-European languages : theory, typology, diachrony*. Current issues in linguistic theory, 0304-0763 ; 145. Amsterdam ;: Benjamins.

Hirt, H. A. (1934). Das Nomen. Kasuslehre. In *Indogermanische Grammatik:*, Volume 6, pp. 29–75. Cambridge: Cambridge University Press.

Hirt, H. A. (1937). *Indogermanische Grammatik. T. 7, Syntax, 2 : die Lehre vom Einfachen und zusammengesetzten Satz*. Heidelberg: Carl Winter.

Hock, H. H. (2013). Proto-indo-european verb-finality reconstruction, typology, validation. *Journal of Historical Linguistics 3*(1), 49–76.

Hock, H. H. and B. D. Joseph (1996). *Language history, language change, and language relationship : an introduction to historical and comparative linguistics*. Trends in linguistics. Studies and monographs, 99-0115958-X ; 93. Berlin: Mouton de Gruyter.

Huelsenbeck, J. P., R. Nielsen, and J. P. Bollback (2003). Stochastic mapping of morphological characters. *Systematic Biology 52*(2), 131.

Jasanoff, J. H. (1978). *Stative and middle in Indo-European*. Innsbrucker Beiträge zur Sprachwissenschaft, 99-0115455-3 ; 23. Innsbruck: Institut für Sprachwissenschaft der Universität.

Jasanoff, J. H. (2003). *Hittite and the Indo-European Verb. [Elektronisk resurs]*. Oxford scholarship online. Oxford : Oxford University Press, 2003.

Josephson, F. and I. Söhrman (2008). *Interdependence of diachronic and synchronic analyses*. Studies in language companion series, 0165-7763 ; 103. Amsterdam ;: John Benjamins.

Jäger, G. (2019). Computational historical linguistics. *Theoretical Linguistics 45*(3/4), 151–182.

Klimov, G. A. (1974). On the character of languages of active typology. *Linguistics 11*.

Krahe, H., H. Schmeja, and W. Meid (1972). *Grundzüge der vergleichenden Syntax der indogermanischen Sprachen, publisher = Inst. f. vergleichende Sprachwiss.* Innsbrucker Beiträge zur Sprachwissenschaft, 99-0115455-3 ; 8. Innsbruck.

Krause, W. and W. Thomas (1960). *Tocharisches Elementarbuch. B. 1, Grammatik*. Heidelberg.

Krisch, T. (1990). Das Wackernagelsche Gesetz aus heutiger Sicht. In H. Eichner and H. Rix (Eds.), *Sprachwissenschaft und Philologie. Jacob Wackernagel und die Indogermanistik heute*, pp. 64–81. Wiesbaden: Harrassowitz.

Kulikov, L. I. and N. Lavidas (2015). *Proto-Indo-European syntax and its development*. Amsterdam ;: John Benjamins Publishing Company.

Ledgeway, A. and I. G. Roberts (2017). *The Cambridge Handbook of historical syntax*.

Lehmann, W. P. (1973). A structural principle of language and its implications. *Language* (1), 47–66.

Lehmann, W. P. (1974). *Proto-Indo-European syntax*. Austin: Univ. of Texas P.

Lehmann, W. P. (1993). *Theoretical bases of Indo-European linguistics*. London: Routledge.

Lehmann, W. P. (2002). *Pre-Indo-European*. Journal of Indo-European Studies. Monograph, 0895-7258 ; 41. Washington, D.C.: Institute for the Study of Man.

Levy, R. and H. Daumé (2011). Computational methods are invaluable for typology, but the models must match the questions. *Linguistic Typology 15*(2), 393–399.

Liggett, T. M. (2010). *Continuous Time Markov Processes: An Introduction*, Volume 113 of *Graduate Studies in Mathematics*. Providence, RI: American Mathematical Society.

Lightfoot, D. W. (2002). Myths and the prehistory of grammars. (1), 113.

Longobardi, G., C. Guardiano, G. Silvestri, A. Boattini, and A. Ceolin (2013). Toward a syntactic phylogeny of modern indo-european languages. *Journal of Historical Linguistics 3*(1), 122–152.

Longobardi, G. and I. Roberts (2011). Non-arguments about non-universals. *Linguistic Typology 15*(2), 483–495.

Luraghi, S. (2011). The origin of the proto-indo-european gender system: Typological considerations. *Folia Linguistics 45*(2), 435–464.

Malchukov, A. L. (2015). Towards a typology of split ergativity: A TAM-hierarchy for alignment splits. In I. Bornkessel-Schlesewsky, A. Malchukov, and M. D. Richards (Eds.), *Scales and Hierarchies: A Cross-Disciplinary Perspective*, pp. 275 – 296. Berlin - New York: Mouton de Gruyter.

Mallory, J. P. and D. Q. Adams (1997). *Encyclopedia of Indo-European culture*. London: Fitzroy Dearborn.

Martinet, A. (1962). *A functional view of language : being the Waynflete lectures delivered in the College of St. Mary Magdalen, Oxford 1961*. Oxford: Clarendon Press.

Matasović, R. (2004). *Gender in Indo-European*. Heidelberg: Winter.

Maurits, L. and T. L. Griffiths (2014). Tracing the roots of syntax with bayesian phylogenetics. *Proceedings of the National Academy of Sciences 111(37)*, 13576–13581.

Meid, W. (1975). Probleme der räumlichen und zeitlichen gliederung des indogermanischen. In H. Rix (Ed.), *Flexion und Wortbildung*, pp. 204–219. Wiesbaden: Ludwig Reichert.

Meier-Brügger, M., M. Fritz, and M. Mayrhofer (2010). *Indogermanische Sprachwissenschaft*. Berlin :: de Gruyter.

Meillet, A. (1948). *Linguistique historique et linguistique générale*. Collection linguistique (Paris), 0560-5202 ; 8. Paris: Champion.

Melchert, C. (2000). Tocharian plurals in -nt and related phenomena. *Tocharian and Indo-European Studies 9*, 53–75.

Murawaki, Y. (2018). *Analyzing Correlated Evolution of Multiple Features Using Latent Representations*.

Nicholls, G. K. and R. D. Gray (2006). Quantifying uncertainty in a stochastic Dollo model of vocabulary evolution. In P. Forster and C. Renfrew (Eds.), *Phylogenetic methods and the prehistory of languages*, pp. 161–71. Cambridge: McDonald Institute for Archaeological Research.

Nichols, J. (1992). *Linguistic diversity in space and time : Linguistic diversity in space and time*. Chicago: Univ. of Chicago Press.

Nichols, J. (1993). Ergativity and linguistic geography. *Australian Journal of Linguistics 13*, 39–89.

Nichols, J. (1995). Diachronically stable structural features. In H. Andersen (Ed.), *Historical Linguistics 1993. Selected Papers from the 11th International Conference on Historical Linguistics. Los Angeles 16-20 August 1993.*, pp. 337–355. Amsterdam - Philadelphia: John Benjamins.

Nichols, J. (1998). The eurasian spread zone and the indo-european dispersal. In R. Blench and M. Spriggs (Eds.), *Archaeology and Language II. Archaeological Data and Linguistic Hypotheses*, pp. 220–266. New York: Routledge.

Nichols, J. (2003). Diversity and stability in languages. In B. D. Joseph and R. D. Janda (Eds.), *The Oxford Handbook of Historical Linguistics*, pp. 283–310. Oxford: Blackwell.

Nielsen, R. (2002). Mapping mutations on phylogenies. *Systematic Biology* (5), 729.

Pagel, M., Q. D. Atkinson, and A. Meade (2007). Frequency of word-use predicts rates of lexical evolution throughout indo-european history. *Nature 449*, 717.

Pagel, M. and A. Meade (2006). Bayesian analysis of correlated evolution of discrete characters by reversible-jump markov chain monte carlo. *The American Naturalist* (6), 808.

Pires, A. and S. G. Thomason (2008). How much syntactic reconstruction is possible? In G. Ferraresi and M. Goldbach (Eds.), *Principles of Syntactic Reconstruction*, pp. 27–72. Amsterdam: Benjamins.

Plank, F. (2011). Call for debate re word-order universals. *Linguistic Typology 15*(2), 333–334.

Pooth, R., P. Kerkhof, L. Kulikov, and J. Barðdal (2018). *The Origin of Non-Canonical Case Marking of Subjects in Proto-Indo-European: Accusative, Ergative, or Semantic Alignment*.

Rix, H. and M. Kümmel (2001). *LIV : Lexikon der indogermanischen Verben : die Wurzeln und ihre Primärstammbildungen*. Wiesbaden: Reichert.

Roberts, I. G. (2007). *Diachronic syntax*. Oxford textbooks in linguistics, 99-2380132-2. Oxford: Oxford University Press.

Schlerath, B. (1981). Ist ein Raum/Zeit-Modell für eine rekonstruierte Sprache möglich? *Zeitschrift für vergleichende Sprachforschung 95*(2), 175–202.

Schmidt, K. H. (1979). Reconstructing active and ergative stages of Pre-Indo-European. In F. Plank (Ed.), *Ergativity: Towards a Theory of Grammatical Relations.*, pp. 333–345. London: Academic Press.

Schmidt, K. H. (1982). Typusrelevanter Sprachwandel flektierend zu agglutinierend und seine Korrelationen. *Études Finno-Ougriennes XV*, 335–346.

Siewierska, A. (1998). *Constituent order in the languages of Europe*. Eurotyp, 99-2389693-5 ; 1. Berlin: Mouton de Gruyter.

Silva, S. G. d. and J. J. Tehrani (2016). Comparative phylogenetic analyses uncover the ancient roots of indo-european folktales. *Royal Society Open Science*.

Sturtevant, E. H. (1962). The indo-hittite hypothesis. *Language 38*(2), 105–110.

Szemerényi, O. (1989). *Einführung in die vergleichende Sprachwissenschaft*. Die Sprachwissenschaft, 0724-5009. Darmstadt: Wissenschaftl. Buchgesellschaft.

Szemerényi, O. (1996). *Introduction to Indo-European linguistics*. Oxford: Clarendon Press.

Tichy, E. (1993). Kollektiva, Genus femininum und relative Chronologie im Indogermanischen. *Historische Sprachforschung / Historical Linguistics 106*(1), 1–19.

Tuffley, C. and M. Steel (1998). Modeling the covarion hypothesis of nucleotide substitution. *Mathematical Biosciences 147*(1), 63–91.

Uhlenbeck, C. C. (1901). Agens und Patiens im Kasussystem der indogermanischen Sprachen. *Indogermanische Forschungen 12*(170-171).

Vaillant, A. (1936). *L' Ergatif indo-européen*. C. Klincksieck.

Vehtari, A., A. Gelman, and J. Gabry (2017). Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and computing 27*(5), 1413–1432.

Viti, C. (2015). Historical syntax: Problems, materials, methods, hypotheses. In C. Viti (Ed.), *Perspectives on Historical Syntax*, pp. 3–34. Amsterdam - Philadelphia: John Benjamins.

Wackernagel, J. (1920). *Vorlesungen über Syntax mit besonderer Berücksichtigung von Griechisch, Lateinisch und Deutsch*. Basel: in Kommissionsvlg von E. Birkhäuser.

Walkden, G. (2013). The correspondence problem in syntactic reconstruction. *Diachronica 30*(1), 95–122.

Walkden, G. (2019). The many faces of uniformitarianism in linguistics. *Glossa: A Journal of General Linguistics 4*(1), 1.

Watkins, C. (1976). Towards proto-indo-european syntax: problems and pseudoproblems. In S. Steever, C. A. Walker, and S. S. Mufwene (Eds.), *Papers from the Parasession on Diachronic Syntax*, pp. 305–326. Chicago: Chicago Linguistic Society.

Wichmann, S. (2014). Diachronic stability and typology. In C. Bowern and B. Evans (Eds.), *The Routledge Handbook of Historical Linguistics*, Volume 212-224. London - New York: Routledge.

Widmer, M., S. Auderset, J. Nichols, P. Widmer, and B. Bickel (2017). NP recursion over time: Evidence from Indo-European. *Language 93*(4), 799–826.

Winter, W. (1984). Reconstructional comparative linguistics and the reconstruction of undocumented stages in the development of languages and language families. In J. Fisiak (Ed.), *Historical Syntax*, pp. 613–625. Berlin - New York: Mouton de Gruyter.

Witzlack-Makarevich, A. and I. A. Seržant (2018). Differential argument marking: Patterns of variation. In I. A. Seržant and A. Witzlack-Makarevich (Eds.), *Diachrony of differential argument marking*, pp. 1–40. Berlin: Language Science Press.

Yang, C. D. (2000). Internal and external forces in language change. *Language variation and change 12*(3), 231–250.

Yang, Z. (2014). *Molecular Evolution: A Statistical Approach*. Oxford: Oxford University Press.

Appendices

Here, we describe the process used to generate the tree sample used for inference in this paper, as well as the details of the inference process used to infer evolutionary transition rates and estimate the probabilities of typological states at unobserved nodes in our tree such as the root (corresponding to Proto-Indo-European).

A Tree Sample

We assume a fixed topology (i.e., branching structure) for the Indo-European languages, informed by the comparative-historical *communis opinio*, as well as recent work on computational phylogenetic methods (cf. Chang et al. 2015), but incorporate uncertainty over branch lengths in the tree in the manner described below. For each tip (i.e., a vertex of the tree where data are observed, or an attested language) or node (i.e., a vertex of the tree where data are not observed, or a proto-language) in tree topology, we establish an upper and lower bound for the date of attestation each language or proto-language, in years before present; in the case of proto-languages, these dates are informed by existing computational

phylogenetic work (e.g., Bouckaert et al. 2012; Chang et al. 2015) and carefully selected termini post quem and termini ante quem based on the historical and archaeological record. The following process is carried out in order to generate a tree with stochastically sampled branch lengths:

Starting at the root, we sample a value $\text{date(root)} \sim \text{Uniform(lower(root), upper(root))}$, where lower and upper refer to the lower and upper bounds established for possible dates of attestation for the root. Then subsequently, moving in pre-traversal order (i.e., from the root of the tree to the tips), we sample a date for each remaining node from the uniform distribution, as follows:

$$\text{date(node)} \sim \text{Uniform}(\max(\text{date(parent)}, \text{lower(node)}), \text{upper(node)})$$

The above sampling statement ensures that all branch lengths are positive. From these sampled dates, we compute the length of the branch between a node n and its daughter d, $b(n, d)$, as follows:

$$b(nd) = \text{date}(n) - \text{date}(d)$$

This process is carried out 10000 times, yielding a tree sample. We scale our tree sample's branch lengths by dividing by 1000.

B Inference of Evolutionary Rates

For a given categorical feature with S states, there are $S(S-1)$ possible transition types between different states. We assume that transitions between two different states follow a Continuous Time Markov process, parameterized by $S(S-1)$ transition rates, which we write as $\mathbf{R}$. We place a Gamma$(1, 1)$ prior over each transition rate in $\mathbf{R}$; the mean of this distribution is 1, corresponding to roughly one change per millennium. The posterior probability of the rates $\mathbf{R}$ can be approximated as follows over an entire tree sample:

$$P(\mathbf{R}|D, \mathbf{T}) \propto P(D, \mathbf{R}|\mathbf{T}) = \sum_{T \in \mathbf{T}} P(D, \mathbf{R}|T) P(T|\boldsymbol{T}) \approx \frac{1}{|\boldsymbol{T}|} \sum_{T \in \boldsymbol{T}} P(D|T, \mathbf{R}) P(\mathbf{R})$$

D refers to the observed data; $\boldsymbol{T}$ denotes the tree sample, and T a single tree in the sample. The likelihood of a given tree and set of rates under the observed data, $P(D, \mathbf{R}|\mathbf{T})$,

can be efficiently computed according to Felsenstein's Pruning Algorithm (Felsenstein 1981). Moving in post-traversal order (i.e., from the tips of the tree toward the root), the Pruning Algorithm calculates the likelihood of a given state s at an internal (i.e., non-tip) node n as follows:

$$\mathcal{L}(n = s) = \prod_{d \in d(n)} \left(\sum_{r \in S} \mathcal{L}(d = r) P_{sr}(b(n, d); \mathbf{R}) \right)$$

Above, $d(n)$ denotes the daughters of node n; $b(a, b)$ denotes the length of the branch connecting nodes a and b (i.e., the displacement in time between two languages); $P_{sr}(t; \mathbf{R})$ denotes the probability of a transition from state s to state r over a time period of length t, given the rates $\boldsymbol{R}$. If data are attested at n, $\mathcal{L}(n = s)$ is either 1 or 0, depending on whether the state attested is s. If data are missing, $\mathcal{L}(n = s)$ is set to 1 for $s \in S$. The overall likelihood of the tree is equal to the following:

$$P(D|\mathbf{R}, T) = \sum_{s \in S} \pi(s) \mathcal{L}(\text{root} = s)$$

Above, $\pi(s)$ denotes the prior probability of state s, which we take to be the stationary probability of s under the continuous-time Markov process, following a number of works from the biological literature (Huelsenbeck et al. 2003; Nielsen 2002; Felsenstein 2004); see Appendix E.

We use the No U-Turn Sampler of RStan (Carpenter et al. 2017) to infer the posterior distributions over transition rates between different states for each categorical feature in our data set. Inference is run over 4 chains for 2000 iterations, with the first half of samples discarded as burn-in. This process is carried out for each tree in the tree sample, and posterior samples are combined for each feature.

C Ancestral State Reconstruction

For each categorical feature, we estimate the distribution over possible states or values at the root of the tree (i.e., for Proto-Indo-European) using samples from the posterior distribution of $\mathbf{R}$. For a sampled vector of rates $\hat{\mathbf{R}}$, the probability of state s at the root is equal to the following:

$$P(\text{root} = s|\hat{\mathbf{R}}) = \frac{\pi(s)\mathcal{L}(\text{root} = s)}{\sum_{r \in S} \pi(r)\mathcal{L}(\text{root} = r)}$$

At each iteration of the inference algorithm, we sample a state at the root as follows:

$$z(\text{root}) \sim \text{Categorical}(P(\text{root}|\hat{\mathbf{R}}))$$

We discard the first half of samples and average the remaining draws of z, yielding a probability between 0 and 1 for each state at the root of the tree.

D Validation

We employ a validation technique inspired by Leave-One-Out cross-validation (LOO-CV; see Vehtari et al. 2017 for a review) in order to see how well the phylogenetic model can reconstruct values observed at tips of the tree when they are held out during the inference process (in general, LOO-CV is used to facilitate direct comparison of goodness-of-fit between competing statistical models; the values we report here serve as a baseline measure against which to compare future work). In particular, we are interested in seeing how accurately state values for observed languages in our sample that are treated as ancestral to other languages (e.g., Latin, Sanskrit, Ancient Greek) relative to non-ancestral languages (e.g., Spanish, Hindi, Modern Greek); high accuracy for held-out ancestral languages indicates that our model carries out ancestral state reconstruction with high accuracy for attested ancestral languages (implying that the model has the potential to generalize to unattested ancestral languages such as Proto-Indo-European). For each feature and each language in our data set, we sample a tree at random, set the likelihood of each state in the language under consideration to 1 (as recommended for missing data; Felsenstein 2004:255), then carry out the evolutionary rate inference described above. Subsequently, we use rates from the posterior distribution of $\mathbf{R}$ to reconstruct the state of the held-out language t as follows. First, we sample a vector of rates $\hat{\mathbf{R}}$ from the posterior and use it to sample a state for the direct ancestor of t, which we denote as n:

$$z(n) \sim \text{Categorical}(\mathcal{L}(n|\hat{\mathbf{R}}))$$

We can use this sampled ancestral state to estimate the probability that t is in state r under a CTM process parameterized by $\hat{\mathbf{R}}$ and sample from this probability distribution:

$$P(t = r) = P_{sr}(b(n, t); \mathbf{R});$$

$$z \sim \text{Categorical}(P(t = r))$$

We carry out this estimation procedure on a randomly sampled tree from our tree sample for each feature and language, and compute the percentage of held-out values generated on the basis of the posterior sample that match the attested value. Figure 17 shows these accuracy scores for each language-feature pair, divided according to whether the language is ancestral or non-ancestral. The overall median accuracy score, pooled across ancestral and non-ancestral languages, is 0.873. Accuracy scores are significantly higher for ancestral languages (median=0.977) than for non-ancestral languages (median=0.830) according to a Mann-Whitney U test ($W = 7147700, p < 2.2e^{-16}$), indicating that the use of ancestry constraints increases posterior predictive accuracy, and that our model can re-capture held-out ancestral values with considerable success. Additionally, we fit a mixed-effects model with log accuracy as a response to whether or not a language is ancestral, with random intercepts by linguistic variable and language,[3] finding via the likelihood ratio test that whether not a language is ancestral is a highly significant predictor of log accuracy ($\chi^2_{\text{LR}}(1) = 18.685, p < 0.0001$).

Additionally, we investigate the degree to which held-out accuracy is dependent on artifacts of the data. For non-ancestral languages, we assess the correlation between the probability of with which our model accurately generates a held-out value and the probability with which this value occurs among non-held-out languages. For ancestral languages, we measure the correlation between the probability of an accurate held-out value and the probability with which this value occurs among the held-out ancestral node's descendants. We find that both correlations are significant (non-ancestral: Spearman's $\rho = 0.221, p < 0.001$, ancestral: $\rho = 0.676, p < 0.001$), indicating that LOO-CV accuracy is sensitive to patterns found in the data. At the same time, these correlations are weak to moderate, indicating that other factors — possibly including the ability of our model to learn meaningful diachronic patterns — influence held-out accuracy as well.

E Choice of prior probability at root

We investigate the extent to which our reconstructions are sensitive to choices of ROOT PRIOR, a key ingredient in the PRUNING ALGORITHM (Felsenstein 1981, 2004). A standard practice in phylogenetics is to set the root prior to be equal to the stationary distribution of the variable of interest under the rates of the CTM process thought to characterize its evolution.

[3] The function call in `lme4` (Bates et al. 2015) is the following: `log(accuracy) ~ ancestral + (1|variable) + (1|language)`

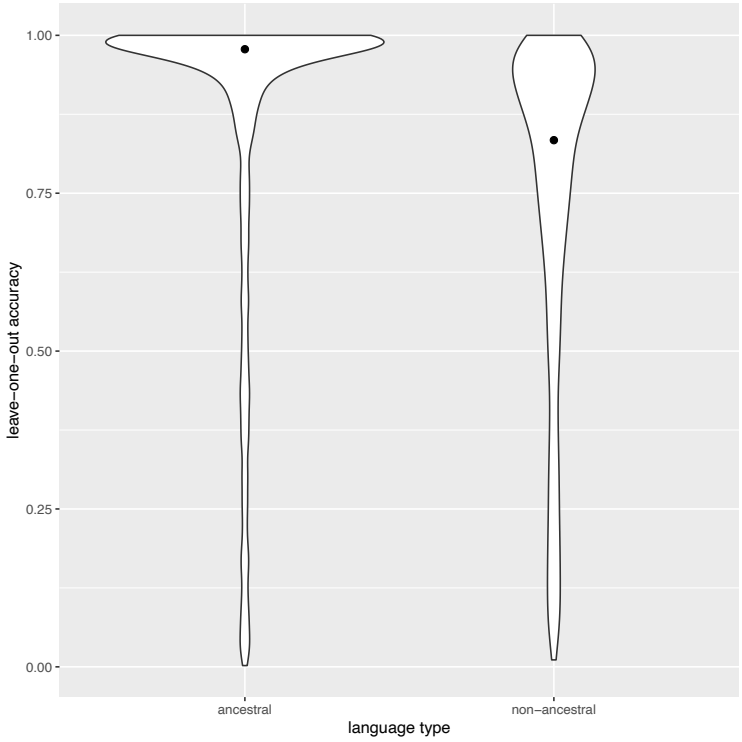


Figure 17: Held-out accuracy values for each language and feature, organized according to whether the feature is ancestral or not. Median values are indicated by dots.

For a binary character with a gain rate α and loss rate β, the stationary probability is equal to $\frac{\alpha}{\alpha+\beta}$.[4] For non-binary data, the stationary distribution $\boldsymbol{\pi}$ is a vector of probabilities satisfying the equations $\boldsymbol{\pi}Q$ and $\sum_{i=1}^{|\boldsymbol{\pi}|} \pi_i = 1$, where Q is the CTM rate matrix. This practice presupposes that the character in question has been evolving for a 'very long time according the particular model of [character evolution] we are using' (Felsenstein 2004:255; cf. Cathcart 2018).

Though a sensible and established choice, this prior has the potential to influence the results of a phylogenetic reconstruction, particularly if the stationary probability is highly skewed toward a particular character value. To explore this issue, we carry out rate inference and reconstruction for the features in our data set using two alternative root priors, a UNIFORM prior which places equal prior probability of each value, and a DIRICHLET prior which infers the distribution of values at the root as a free parameter to be inferred. For a variable with D values, we set the prior probability of each value to $\frac{1}{D}$ under the uniform regime; under the Dirichlet regime, we place a D-length symmetric Dirichlet prior with a concentration parameter of 1 over the feature distribution at the root.

The different prior regimes reconstruct the same features with highest probability for the majority of variables, but differ according to a fraction of variables. These are presented below, along with the probabilities with which they are reconstructed, for stationary probability (SP), uniform prior (UP) and Dirichlet prior (DP); cells for which an alternative regime agrees with the SP regime are blank.

[4] An alternative parameterization for binary data is to draw a stationary probability $p \in [0, 1]$ and a change rate s, from which gain and loss rates can be derived as ps and $(1-p)s$, respectively.

SP	UP	DP
Clitic finite V: Category irrelevant (0.442)	Clitic finite V: V2 (0.270)	Clitic finite V: V2 (0.269)
Clitic Infinitive V: Category irrelevant (0.433)	Clitic Infinitive V: V2 (0.254)	Clitic Infinitive V: V2 (0.249)
Clitic Participle V: Category irrelevant (0.418)	Clitic Participle V: V2 (0.236)	Clitic Participle V: V2 (0.240)
No case on article (0.732)	Case on article (0.621)	Case on article (0.618)
No difference O and Dative (pronouns) (0.570)	Difference, O and Dative (pronouns) (0.519)	Difference, O and Dative (pronouns) (0.529)
No Future by participle (0.724)	Future by participle (0.767)	Future by participle (0.762)
Not more than 7 cases (0.583)	More than 7 cases (0.948)	More than 7 cases (0.950)
NRel: Noun - Relative (0.627)	NRel: Relative - Noun (0.315)	NRel: Relative - Noun (0.316)
Possessor - Noun (0.585)	Possessor - Noun and Noun - Possessor (0.374)	Possessor - Noun and Noun - Possessor (0.377)
Reflexive with Agent (0.959)	Reflexive not with Agent (0.505)	Reflexive not with Agent (0.505)
Simple past: Full A Agreement (0.386)	—	Simple past: Syncretic A Agreement (0.206)

It is worth noting that with a handful of exceptions, the features reconstructed differently by the alternative models are reconstructed with high uncertainty, and as a whole, the differing features are not of key importance to the canonical model of reconstruction supported by the main results presented in this paper (see §3).

F Entry/Gain and Exit/Loss Rates

For each multistate character, we compute the mean rate at which each state is gained or lost in the following manner. We compute the gain rate for state i, or rate at which state i is entered, as follows, where $R(j \to i)$ denotes the rate from state j to state i, and $p(s)$ denotes the equilibrium or stationary probability of state s:

$$\frac{\sum_{j \neq i} p(j) R(j \to i)}{\sum_{j \neq i} p(j)}$$

The loss rate for state i, or rate at which state i is exited, is computed as follows:

$$\sum_{j \neq i} R(i \to j)$$

Supplementary materials: data and results

S1

Languages, including latitude and longitude, used in the current study.

Language	Longitude	Latitude
Albanian (Tosk)	19.988251	40.446947
Angloromani	-2.355272	53.779521
Ashkun	70.791057	35.255909
Assamese	91.79022	26.143538
Avestan	55.9499	31.70708
Baluchi	65.03622	26.27794
Bengali	88.36119	22.566866
Breton	-4.102864	47.995619
Bulgarian	23.3535	42.698586
Catalan	2.179642	41.379225
Classical Greek	23.999205	37.696861
Cornish	-5.567517	50.191816
Croatian	15.992661	45.813486
Czech	14.468136	50.071244
Danish	12.583182	55.679885
Dutch	4.898126	52.374367
Elfdalian	14.040485	61.225938
English	-0.100059	51.497728
Faroese	-6.773322	62.009754
French	2.347857	48.854451
Frisian	5.795597	53.20394
Friulian	13.485001	46.127028
German	13.409092	52.520822
Gilaki	49.283999	37.526056
Gothic	20.199324	53.042874
Gujarati	72.139869	22.44456
Hindi	77.224996	28.625252
Hittite	34.621158	40.013951
Icelandic	-21.939672	64.144237
Irish	-9.301101	53.244432
Italian	12.385145	43.115956
Kashmiri	75.939182	34.322996
Kati	71.314571	35.799992
Khowar	72.51658	36.457958
Konkani	72.871066	19.262081
Kurdish (Kurmanji)	43.822746	37.788081

Kurdish (Sorani)	45.728512	35.821161
Ladin	11.923755	46.657384
Latin	12.4923	41.890189
Latvian	24.113274	56.915
Lithuanian	25.255852	54.680183
Low German	9.488653	53.931994
Luwian	36.788273	36.620598
Maithili	86.155454	26.075397
Maldivian	73.511939	4.17191
Manx	-4.484657	54.151988
Marathi	75.494714	19.38701
Middle Breton	-1.532472	47.207854
Middle Dutch	5.959192	52.213073
Middle English	-0.758622	51.313607
Middle Greek	28.997555	41.058644
Middle High German	9.717403	48.742717
Middle Irish	-8.634407	52.658918
Middle Low German	10.68504	53.867783
Middle Persian	51.675883	32.647027
Middle Welsh	-3.380017	52.240359
Modern Armenian	44.527159	40.147389
Modern Greek	23.694334	38.028622
Nepali	89.243625	26.659702
Norwegian (Bokmål)	10.740802	59.912785
Norwegian (Nynorsk)	5.075361	60.330984
Old Church Slavonic	30.453243	50.415519
Old Dutch	5.123259	52.09095
Old English	-1.314163	51.060696
Old French	2.347857	48.854451
Old Frisian	6.567738	53.215696
Old High German	9.962616	49.795477
Old Irish	-6.873827	53.726794
Old Italian	11.249249	43.776281
Old Norse	10.397522	63.426975
Old Persian	52.891196	29.935501
Old Portuguese	-8.629141	41.15818
Old Provençal	6.222519	43.426825
Old Prussian	19.900799	54.436292
Old Russian	31.275345	58.521428
Old Saxon	8.806043	53.07585
Old Spanish	-4.023666	39.857094
Old Swedish	16.321936	58.481083

Oriya	86.278412	20.608549
Ossetian (Iron)	44.675689	43.043802
Pali	83.000379	27.617895
Parachi	69.702595	34.832889
Pashto	68.188804	31.432929
Persian	55.748043	32.735314
Polish	20.994015	52.237892
Portuguese	-9.135132	38.708043
Prakrit	84.727979	25.379098
Prasun	70.741619	35.33661
Provençal	6.222519	43.426825
Punjabi	74.349411	31.536604
Romani (Arli)	19.881168	42.970349
Romani (Bugurdži)	21.222651	42.118106
Romani (Burgenland)	16.377534	47.193164
Romani (Kale)	21.470967	60.972087
Romani (Kelderash)	22.495093	47.187683
Romani (Lovara)	23.090199	47.401544
Romani (Sepečides)	27.131666	38.392804
Romani (Sinte)	10.203424	51.4992
Romanian	26.086373	44.427511
Romansh	9.838341	46.496755
Russian	37.64843	55.677584
Sanskrit	76.28612	23.511415
Sardinian	9.32737	40.31746
Scandoromani	13.158622	59.481307
Scottish Gaelic	-7.368974	57.16276
Serbian	20.472164	44.801327
Shughni	71.503041	38.389635
Sicilian	13.368345	38.115141
Sindhi	69.674802	23.253479
Sinhalese	79.85635	6.935115
Slovene	14.50823	46.050593
Sogdian	66.962682	39.646576
Spanish	-3.703482	40.416881
Swedish	17.438718	59.73823
Swiss German	8.538354	47.36888
Tajik	70.342617	38.663458
Talysh	48.730017	38.854174
Tocharian A	86.154785	41.760336
Tocharian B	81.514435	41.645502=
Ukrainian	30.525414	50.449474

Upper Sorbian	14.425442	51.179914
Urdu	73.018423	33.681135
Wakhi	72.760057	37.219477
Walloon	4.865382	50.464478
Welsh	-4.38518	52.926407
Yiddish	24.031167	49.853028

S2a

List of typological features (from DiACL database https://diacl.ht.lu.se) used in the current study.

Grid = topmost organizational unit in database, corresponding to linguistic domain; Feature = second organizational unit in database; Feature description = Description of Feature in database; Variant = lowest organizational unit in database; Variant description = description of variant in database; ID = unique database ID of Variant.

Grid	Feature	Feature description	Variant	Variant description	ID
Word order	Adpositions	Do adpositions normally occur before or after the noun?	Prep	Does the language have a substantial set of prepositions? E.g. English in the house	213
Word order	Adpositions	Do adpositions normally occur before or after the noun?	Post	Does the language have a substantial set of postpositions?	214
Word order	Noun-adjective	Do adjectives normally occur before or after the noun?	NA	Do most adjectives occur after the noun?	215
Word order	Noun-adjective	Do adjectives normally occur before or after the noun?	AN	Do most adjectives occur before the noun?	216
Word order	Noun-relative clause	Do relative clauses normally occur before or after the noun?	NRel	Do most relative clauses occur after the noun?	217
Word order	Noun-relative clause	Do relative clauses normally occur before or after the noun?	RelN	Do most relative clauses occur before the noun?	218
Word order	Noun-possessor	Do possessors normally occur before or after the noun?	N-Poss	Do most possessors occur after the noun they possess? The possessor should be an animate noun, and neither a proper name nor a pronoun!	219

Word order	Noun-possessor	Do possessors normally occur before or after the noun?	Poss-N	Do most possessors occur before the noun they possess? The possessor should be an animate noun, and neither a proper name nor a pronoun!	220
Word order	WH-element	What is the position of the WH-question word?	WH-initial	Is the WH-question word always obligatorily the first element in a question (e.g., it does not trigger inversion)?	221
Word order	WH-element	What is the position of the WH-question word?	WH-V	Does the WH-question word always immediately precede the verb (i.e., stand directly before the verb, either initially or non-initially)?	222
Word order	Main clause	What is the canonical (neutral) word order in a main clause?	SVO	What is the canonical (neutral) word order in a main clause? NB: V2 languages like Swe and Ger do NOT count as SVO even though SVO is most frequent.	223
Word order	Main clause	What is the canonical (neutral) word order in a main clause?	V2	V2 implies that initial adverb triggers V-SUBJ word order (Swedish, German etc.).	224
Word order	Main clause	What is the canonical (neutral) word order in a main clause?	VSO	What is the canonical (neutral) word order in a main clause?	225
Word order	Main clause	What is the canonical (neutral) word order in a main clause?	SOV	What is the canonical (neutral) word order in a main clause?	226
Word order	Subordinate clause	What is the canonical (neutral) word order in a subordinate clause?	SVO	What is the canonical (neutral) word order in a subordinate clause? NB: V2 languages like Swe and Ger do NOT count as SVO even though SVO is most frequent.	227
Word order	Subordinate clause	What is the canonical (neutral) word order in a subordinate clause?	V2	V2 implies that initial adverb triggers V-SUBJ word order (Swedish, German etc.).	228
Word order	Subordinate clause	What is the canonical (neutral) word order in a subordinate clause?	VSO	What is the canonical (neutral) word order in a subordinate clause?	229

Word order	Subordinate clause	What is the canonical (neutral) word order in a subordinate clause?	SOV	What is the canonical (neutral) word order in a subordinate clause?	230
Word order	Infinitive	Does the object normally occur before or after an infinitive? E.g.: to make pancakes (VO) Pfannkuchen machen (OV)	VO		231
Word order	Infinitive	Does the object normally occur before or after an infinitive? E.g.: to make pancakes (VO) Pfannkuchen machen (OV)	OV		232
Word order	Participle	Does the object normally occur before or after a participle? E.g.: making pancakes (VO) Pfannkuchen machend (OV)	VO		233
Word order	Participle	Does the object normally occur before or after a participle? E.g.: making pancakes (VO) Pfannkuchen machend (OV)	OV		234
Word order	Clitic pronouns finite verb	Does the clitic object pronoun normally occur before or after a finite verb? E.g.: Je les fais. (OV) If a language does not have clitic object pronouns, it would be 0 in both OV and VO.	VO		235

Word order	Clitic pronouns finite verb	Does the clitic object pronoun normally occur before or after a finite verb? E.g.: Je les fais. (OV) If a language does not have clitic object pronouns, it would be 0 in both OV and VO.	OV		236
Word order	Clitic pronouns finite verb	Does the clitic object pronoun normally occur before or after a finite verb? E.g.: Je les fais. (OV) If a language does not have clitic object pronouns, it would be 0 in both OV and VO.	2nd position	Does the clitic pronoun always occur in 2nd position, not specifically before or after the verb? (Wackernagel position)	237
Word order	Clitic pronouns infinitive	Does the clitic object pronoun normally occur before or after an infinitive? E.g.: Je veux les faire. (OV) If a language does not have clitic object pronouns, it would be 0 in both OV and VO.	VO		238
Word order	Clitic pronouns infinitive	Does the clitic object pronoun normally occur before or after an infinitive? E.g.: Je veux les faire. (OV) If a language does not have clitic object pronouns, it would be 0 in both OV and VO.	OV		239

Word order	Clitic pronouns infinitive	Does the clitic object pronoun normally occur before or after an infinitive? E.g.: Je veux les faire. (OV) If a language does not have clitic object pronouns, it would be 0 in both OV and VO.	2nd position	Does the clitic pronoun always occur in 2nd position, not specifically before or after the verb?	240
Word order	Clitic pronouns participle	Does the clitic object pronoun normally occur before or after a participle? E.g.: En les faisant... (OV) If a language does not have clitic object pronouns, it would be 0 in both OV and VO.	VO		241
Word order	Clitic pronouns participle	Does the clitic object pronoun normally occur before or after a participle? E.g.: En les faisant... (OV) If a language does not have clitic object pronouns, it would be 0 in both OV and VO.	OV		242

Word order	Clitic pronouns participle	Does the clitic object pronoun normally occur before or after a participle? E.g.: En les faisant... (OV) If a language does not have clitic object pronouns, it would be 0 in both OV and VO.	2nd position	Does the clitic pronoun always occur in 2nd position, not specifically before or after the verb?	243
Nominal morphology	Nominal case	What is the realization of case at nouns (nominal heads)?	O-case	Are there different noun forms for agent and object case? (English: 0 (no cases) Russian: 1 (different noun forms for accusative and nominative) Basque: 1 (different noun forms for ergative and absolutive)	244
Nominal morphology	Nominal case	What is the realization of case at nouns (nominal heads)?	DAT	Is there a specific case form for the recipient, which is different from the case form of, e.g., the object? (E.g. The man gives a book (O) to the child (DAT))	245
Nominal morphology	Nominal case	What is the realization of case at nouns (nominal heads)?	GEN	Is there a special case form to express genitive, which is different from the agent/object case?	246
Nominal morphology	Nominal case	What is the realization of case at nouns (nominal heads)?	GEN/DAT	Is there a special noun form to express genitive, which is not the same as dative (recipient) case?	247
Nominal morphology	Nominal case	What is the realization of case at nouns (nominal heads)?	VOC	Is there a special noun form to express vocative which is not the same as agent or object case?	248
Nominal morphology	Nominal case	What is the realization of case at nouns (nominal heads)?	OBL-Cases	Are there any cases besides agent, object, genitive, dative, and vocative? (E.g., local cases)	249
Nominal morphology	Nominal case	What is the realization of case at nouns (nominal heads)?	>7 Cases	Are there more than 7 cases?	250
Nominal morphology	Nominal case	What is the realization of case at nouns (nominal heads)?	AGGL.CASE	Are there cases which are visibly agglutinative, i.e., built up by several distinct, segmentable affixes?	251

Nominal morphology	Nominal case	What is the realization of case at nouns (nominal heads)?	AGGL.CASE.NR	Are plural cases formed by combining an (infixed) plural affix and a case affix in an agglutinative manner?	252
Nominal morphology	Pronominal case	What is the realization of case at pronouns (pronominal heads)? Mainly 1st and 2nd person pronouns, ignoring 3rd person pronouns (which often come from demonstratives).	A $\neq$ O	In pronouns, is the marking different for the case of the agent and object?	253
Nominal morphology	Pronominal case	What is the realization of case at pronouns (pronominal heads)? Mainly 1st and 2nd person pronouns, ignoring 3rd person pronouns (which often come from demonstratives).	DAT $\neq$ O	In pronouns, is the marking different for the case of the recipient and the object?	254
Nominal morphology	Pronominal case	What is the realization of case at pronouns (pronominal heads)? Mainly 1st and 2nd person pronouns, ignoring 3rd person pronouns (which often come from demonstratives).	VOC	See Nominal case	255

Nominal morphology	Pronominal case	What is the realization of case at pronouns (pronominal heads)? Mainly 1st and 2nd person pronouns, ignoring 3rd person pronouns (which often come from demonstratives).	OBL-Cases	See Nominal case	256
Nominal morphology	Pronominal case	What is the realization of case at pronouns (pronominal heads)? Mainly 1st and 2nd person pronouns, ignoring 3rd person pronouns (which often come from demonstratives).	>7 Cases	See Nominal case	257
Nominal morphology	Pronominal case	What is the realization of case at pronouns (pronominal heads)? Mainly 1st and 2nd person pronouns, ignoring 3rd person pronouns (which often come from demonstratives).	AGGL.CASE	See Nominal case	258
Nominal morphology	Pronominal case	What is the realization of case at pronouns (pronominal heads)? Mainly 1st and 2nd person pronouns, ignoring 3rd person pronouns (which often come from demonstratives).	AGGL.CASE.NB	See Nominal case	259

Nominal morphology	Case marking	On which elements of the NP is the case marking obligatory?	CASE-LAST	Is the case marking obligatory on the last element of the NP (i.e., it is only realized once in the NP, even if it consists of several elements)?	260
Nominal morphology	Case marking	On which elements of the NP is the case marking obligatory?	CASE-FIRST	Is the case marking obligatory on the first element of the NP (i.e., it is only realized once in the NP, even if it consists of several elements)?	261
Nominal morphology	Case marking	On which elements of the NP is the case marking obligatory?	CASE-N	Is the case marking obligatory realized on the noun?	262
Nominal morphology	Case marking	On which elements of the NP is the case marking obligatory?	CASE-ADJ	Is the case marking obligatory on the adjective?	263
Nominal morphology	Case marking	On which elements of the NP is the case marking obligatory?	CASE-ART	Is the case marking realized on the article?	264
Nominal morphology	Gender / noun class	How is gender / noun class realized in the language?	M/F	Is there an obligatory gender distinction between masculine and feminine realized on an agreeing article or adjective? Can be either on the adjective (Russian) or on both the article and the adjective (German), or even on a verb (as in some NE Caucasian languages).	265
Nominal morphology	Gender / noun class	How is gender / noun class realized in the language?	NEUTR	Is there a special neutral gender for nouns realized on an agreeing article, adjective or verb?	266
Nominal morphology	Gender / noun class	How is gender / noun class realized in the language?	ANIM	Is there a special noun class for non-human animates realized on an agreeing article, adjective or verb?	267
Nominal morphology	Gender / noun class	How is gender / noun class realized in the language?	<5 GENDER	Are there more than 5 noun classes (or genders)?	268
Nominal morphology	Definiteness marking	How is definiteness marking realized in the language?	DEF-ART	Is there a special word class of definite articles which occur in NPs without adjectives? (E.g.: German, English but not Swedish)	269

Nominal morphology	Definiteness marking	How is definiteness marking realized in the language?	N-DEF	Is there a suffix for definiteness on the noun? (E.g.: Swedish but not English!)	270
Nominal morphology	Definiteness marking	How is definiteness marking realized in the language?	ADJ-DEF	Is there a suffix for definiteness on the adjective? This includes cases when the ADJ has a different form in definite and indefinite NPs (Swe "det stora huset", Ger "das grosse Haus").	271
Nominal morphology	Definiteness marking	How is definiteness marking realized in the language?	DEF-LAST	Is the definiteness marking obligatory on the last element of the NP (so it is only realized once in the NP, even if it consists of several elements)? If there is no definiteness marking at all, it will be 0!	272
Nominal morphology	Definiteness marking	How is definiteness marking realized in the language?	DEF-FIRST	Is the definiteness marking obligatory on the first element of the NP (so it is only realized once in the NP, even if it consists of several elements)? If there is no definiteness marking at all, it will be 0!	273
Nominal morphology	Gender agreement	How is gender agreement realized in the language?	PRED-ADJ	Does a predicative adjective agree with the subject of the clause in gender?	274
Nominal morphology	Preposition agreement	How is preposition agreement realized in the language?	PREP-PRON-AGR	Can a preposition agree in person with its object?	275
Verbal morphology	Simple PAST, A	In simple past, how is verbal agreement realized with respect to the agent?	PST:A-AGR-FULL	In simple past: does the verb crossreference the agent in all persons /numbers?	276
Verbal morphology	Simple PAST, A	In simple past, how is verbal agreement realized with respect to the agent?	PST:NO-A-AGR	In simple past: does the verb not crossreference the agent on the verb at all (e.g., Swedish)?	277
Verbal morphology	Simple PAST, A	In simple past, how is verbal agreement realized with respect to the agent?	PST:A-Gender-AGR	In simple past, does the verb agree in gender with the subject of a transitive verb? (e.g., Russian, Polish).	278

Verbal morphology	Simple PAST, O	In simple past, how is verbal agreement realized with respect to the object?	PST:O-AGR-FULL	In simple past: does the verb crossreference the object in all persons /numbers?	279
Verbal morphology	Simple PAST, O	In simple past, how is verbal agreement realized with respect to the object?	PST:NO-O-AGR	In simple past: does the verb not crossreference the object on the verb at all (e.g., Swedish, English, Russian)?	280
Verbal morphology	Simple PAST, O	In simple past, how is verbal agreement realized with respect to the object?	PST:O-Gender-AGR	In simple past, does the verb agree in gender with the object?	281
Verbal morphology	Simple PAST, DAT	In simple past, how is verbal agreement realized with respect to the indirect object of ditransitive verbs?	PST:DAT-AGR-FULL	In simple past: does the verb crossreference the dative in all persons /numbers?	282
Verbal morphology	Simple PAST, DAT	In simple past, how is verbal agreement realized with respect to the indirect object of ditransitive verbs?	PST:NO-DAT-AGR	In simple past: does the verb not crossreference the dative on the verb at all (e.g., Swedish, English, Russian)	283
Verbal morphology	Simple PAST, DAT	In simple past, how is verbal agreement realized with respect to the indirect object of ditransitive verbs?	PST:DAT-Gender-AGR	In simple past, does the verb agree in gender with the indirect object of a ditransitive verb?	284
Verbal morphology	Present progressive, A	In present progressive: how is verbal agreement realized with respect to the agent?	PROG:A-AGR-FULL	In present progressive: does the verb crossreference the agent in all persons /numbers?	285
Verbal morphology	Present progressive, A	In present progressive: how is verbal agreement realized with respect to the agent?	PROG:NO-A-AGR	In present progressive: does the verb not crossreference the agent on the verb at all (e.g., Swedish)	286

Verbal morphology	Present progressive, A	In present progressive: how is verbal agreement realized with respect to the agent?	PROG:A-Gender-AGR	In present progressive, does the verb agree in gender with the subject of a transitive verb?	287
Verbal morphology	Present progressive, O	In present progressive, how is verbal agreement organized with respect to the object?	PROG:O-AGR-FULL	In present progressive: does the verb crossreference the object in all persons /numbers?	288
Verbal morphology	Present progressive, O	In present progressive, how is verbal agreement organized with respect to the object?	PROG:NO-O-AGR	In present progressive: does the verb not crossreference the object on the verb at all (e.g. Swedish, Russian)?	289
Verbal morphology	Present progressive, O	In present progressive, how is verbal agreement organized with respect to the object?	PROG:O-Gender-AGR	In present progressive, does the verb agree in gender with the object of a transitive verb?	290
Verbal morphology	Present progressive, DAT	In present progressive, how is verbal agreement organized with respect to the indirect object of ditransitive verbs?	PROG:DAT-AGR-FULL	Iin present progressive: does the verb crossreference the indirect object of a ditransitive verb in all persons /numbers?	291
Verbal morphology	Present progressive, DAT	In present progressive, how is verbal agreement organized with respect to the indirect object of ditransitive verbs?	PROG:NO-DAT-AGR	In present progressive: does the verb not crossreference the indirect object of ditransitive verbs at all (e.g. Swedish, English, Russian)?	292
Verbal morphology	Present progressive, DAT	In present progressive, how is verbal agreement organized with respect to the indirect object of ditransitive verbs?	PROG:DAT-Gender-AGR	In the present progressive, does the verb agree in gender with the indirect object of a ditransitive verb?	293

Verbal morphology	Allocutive agreement	Does the verb agree with the receiver (the person one is speaking to) without the speaker being an argument in the sentence (allocutive agreement, probably no for all languages but Basque!)	ALLOC	Does the verb agree with the receiver (the person one is speaking to) without the speaker being an argument in the sentence (allocutive agreement, probably no for all languages but Basque!)	294
Tense	Future	How is future realized in the language?	FUT.AUX	Is there a future formed by an auxiliary? (E.g., will in English?	295
Tense	Future	How is future realized in the language?	PERF.FUT	Is there a future formed by using the perfective aspect? (0 if the language does not have verbal aspects! E.g., Russian, Georgian)	296
Tense	Future	How is future realized in the language?	FUT.Participle	Is there a future formed by a participle? (E.g., Armenian, Basque)	297
Tense	Future	How is future realized in the language?	FUT.Particle	Is there a future formed by a particle preceding a finite verb? (E.g., Albanian)	298
Tense	Future	How is future realized in the language?	FUT.Synth	Is there a synthetical future? (E.g., French, Spanish)	299
Tense	Continous present	How is present progressive realized in the language?	Present	Is there a synthetic present in progressive function?	300
Tense	Continous present	How is present progressive realized in the language?	Progressive present	Is there a progressive present form constructed by combining a present participle with a finite auxiliary verb?	301
Alignment	Noun: Simple Past	In simple past: how is the marking of subject and object of nouns realized?	N:PST:A=O?	In simple past: Is the noun form for A the same as for O? Ie: Does the noun look the same when it is subject of a transitive clause than when it is object of a transitive clause?	302

Alignment	Noun: Simple Past	In simple past: how is the marking of subject and object of nouns realized?	N:PST:A=Sa?	In simple past:Is the noun form for A the same as for Sa? Ie: Does the noun look the same when it is subject of a transitive clause as when it is subject of an agentive intransitive verb such as "work" or "dance"?	303
Alignment	Noun: Simple Past	In simple past: how is the marking of subject and object of nouns realized?	N:PST: O=So?	In simple past: Is the noun form for O the same as for So? Ie: Does the noun look the same when it is object of a transitive clause as when it is subject of an unaccusative verb such as "fall" or "die"?	304
Alignment	Noun: Simple Past	In simple past: how is the marking of subject and object of nouns realized?	N:PST: Sa=So?	In simple past: Does a noun bear the same case form when it is Sa (subject of e.g. work) or So (subject of e.g. fall or die)? Ie: There does not exist a split into stative and active intransitive verbs.	305
Alignment	Noun: Present Progressive	In present progressive: how is the marking of subject and object of nouns realized?	N:PROG: A=O?	In present progressive: Is the noun form for A the same as for O? I.e.: Does the noun look the same when it is subject of a transitive clause and when it is object of a transitive clause?	306
Alignment	Noun: Present Progressive	In present progressive: how is the marking of subject and object of nouns realized?	N:PROG: A=Sa?	In present progressive:Is the noun form for A the same as for Sa? I.e.: Does the noun look the same when it is subject of a transitive clause and when it is subject of an agentive intransitive verb such as "work" or "dance"?	307
Alignment	Noun: Present Progressive	In present progressive: how is the marking of subject and object of nouns realized?	N:PROG: O=So?	In present progressive: Is the noun form for O the same as for So? I.e.: Does the noun look the same when it is object of a transitive clause and when it is subject of an unaccusative verb such as "fall" or "die"?	308

Alignment	Noun: Present Progressive	In present progressive: how is the marking of subject and object of nouns realized?	N:PROG: Sa=So?	In present progressive: does a noun bear the same case form when it is Sa (subject of e.g., “work”) or So (subject of e.g., “fall” or “die”)? I.e.: The language does not have a split between stative and active intransitive verbs.	309
Alignment	Pronoun: Simple Past	In present progressive: how is the marking of subject and object of pronouns realized?	P:PST: A=O?	In simple past: Is the pronoun form for A the same as for O? I.e.: Does the pronoun look the same when it is subject of a transitive clause than when it is object of a transitive clause?	310
Alignment	Pronoun: Simple Past	In present progressive: how is the marking of subject and object of pronouns realized?	P:PST: A=Sa?	In simple past: Is the pronoun form for A the same as for Sa? I.e.: Does the pronoun look the same when it is subject of a transitive clause than when it is subject of an agentive intransitive verb such as “work” or “dance”?	311
Alignment	Pronoun: Simple Past	In present progressive: how is the marking of subject and object of pronouns realized?	P:PST: O=So?	In simple past: Is the pronoun form for O the same as for So? I.e.: Does the pronoun look the same when it is object of a transitive clause than when it is subject of an unaccusative verb such as “fall” or “die”?	312
Alignment	Pronoun: Simple Past	In present progressive: how is the marking of subject and object of pronouns realized?	P:PST: Sa=So?	In simple past: Does the pronoun bear the same case form when it is Sa (subject of e.g. work) or So (subject of e.g. fall or die)? I.e.: There does not exist a split into stative and active intransitive verbs.	313
Alignment	Pronoun: Present Progressive	In present progressive: how is the marking of subject and object of pronouns realized?	P:PROG: A=O?	In present progressive: Is the pronoun form for A the same as for O? I.e.: Does the pronoun look the same when it is subject of a transitive clause than when it is object of a transitive clause?	314

Alignment	Pronoun: Present Progressive	In present progressive: how is the marking of subject and object of pronouns realized?	P:PROG: A=Sa?	In present progressive:Is the pronoun form for A the same as for Sa? I.e.: Does the pronoun look the same when it is subject of a transitive clause as when it is subject of an agentive intransitive verb such as “work” or “dance”?	315
Alignment	Pronoun: Present Progressive	In present progressive: how is the marking of subject and object of pronouns realized?	P:PROG: O=So?	In present progressive: Is the pronoun form for O the same as for So? I.e.: Does the pronoun look the same when it is object of a transitive clause than when it is subject of an unaccusative verb such as “die” or “fall”?	316
Alignment	Pronoun: Present Progressive	In present progressive: how is the marking of subject and object of pronouns realized?	P:PROG: Sa=So?	In present progressive: Does a pronoun bear the same case form when it is Sa (subject of e.g. work) or So (subject of e.g. fall or die)? Ie: There does not exist a split into stative and active intransitive verbs.	317
Alignment	Verb: Simple Past	In simple past, how is alignment realized on the verb?	V:PST:A=O?	In simple past: Is the verb affix for A the same as for O? I.e.: Does the verb look the same when it refers to the subject of a transitive clause than when it refers to the object of a transitive clause? If there is no O-marking on the verb, but there is an S-marking, the answer would be no, they do not look the same. (e.g., German, Russian) If there is neither an O, nor an A marking, like in Swedish, the answer would be yes, they look the same!	318
Alignment	Verb: Simple Past	In simple past, how is alignment realized on the verb?	V:PST:A=Sa?	In simple past: Is the verb affix for A the same as for Sa? I.e.: Does the verb look the same when it refers to subject of a transitive clause than when it refers to subject of an agentive intransitive verb like “work” or “dance”?	319

Alignment	Verb: Simple Past	In simple past, how is alignment realized on the verb?	V:PST:O=So?	In simple past: Is the verb affix for O the same as for So? I.e.: Does the verb look the same when it refers to the object of a transitive clause as when it refers to the subject of an unaccusative verb (such as "fall" or "die")?	320
Alignment	Verb: Simple Past	In simple past, how is alignment realized on the verb?	V:PST:Sa=So?	In simple past: Is the verb affix the same for Sa (subject of e.g. work) as or So (subject of e.g. fall or die)? I.e., does the verb agreement affix look the same regardless of whether the verb is "work" or "die" (as in German: "arbeitete-st", "starb-st"). I.e.: There does not exist a split into unaccusative and agentive intransitive verbs.	321
Alignment	Verb: Present Progressive	In present progressive, how is alignment realized on the verb?	V:PROG: A=O?	In present progressive: Is the verb affix for A the same as for O? Ie: Does the verb look the same when it refers to the subject of a transitive clause than when it refers to the object of a transitive clause? If there is no O-marking on the verb, but there is an S-marking, the answer would be no, they do not look the same. (e.g. German, Russian) If there is neither an O, nor an A marking like in Swedish, the answer would be yes, they look the same!	322
Alignment	Verb: Present Progressive	In present progressive, how is alignment realized on the verb?	V:PROG: A=Sa?	In present progressive:Is the verb affix for A the same as for Sa? Ie: Does the verb look the same when it refers to subject of a transitive clause as when it refers to subject of an agentive intransitive verb such as "work"?	323

Alignment	Verb: Present Progressive	In present progressive, how is alignment realized on the verb?	V:PROG: O=So?	In present progressive: Is the verb affix the same for O as for So? Ie: Does the verb look the same when it refers to the object of a transitive clause as when it refers to the subject of an unaccusative verb (such as "fall" or "die")?	324
Alignment	Verb: Present Progressive	In present progressive, how is alignment realized on the verb?	V:PROG: Sa=So?	In present progressive: Is the verb affix the same for Sa (subject of e.g. "work") as or So (subject of e.g. "fall" or "die")? I.e. does the verb agreement affix look the same regardless of whether the verb is "work" or "die" (as in German: arbeite-t, stirb-t). I.e.: There does not exist a split into unaccusative and agentive intransitive verbs.	325
Alignment	Compare PROG-PAST	What is the marking relation between subject and object in present progressive and simple past?	PROG_So= PAST_So	Does the subject of e.g. die or fall bear the same case in both progressive present and simple past? (the answer for e.g., Megrelian would be no)	326
Alignment	Compare PROG-PAST	What is the marking relation between subject and object in present progressive and simple past?	PROG_A= PAST_O	Does the subject of a transitive verb in the present progressive bear the same case form as the object of a transitive verb in the simple past? (e.g. as in Georgian)	327
Alignment	Compare PROG-PAST	What is the marking relation between subject and object in present progressive and simple past?	PAST_A= PROG_O	Does the subject of a transitive verb in the simple past bear the same case form as the object of a verb in the present progressive? (e.g., Kurdish)	328
Alignment	Reflexive pronoun in transitive clause	What is the alignment of reflexive pronouns?	REFL-ref-A	In a transitive clause, can O be a reflexive which refers back to A (as in English "herself", Swedish "sig")?	329

Alignment	Reflexive pronoun in transitive clause	What is the alignment of reflexive pronouns?	REFL-ref-O	In a transitive clause, can A be a reflexive which refers back to O (as appears to be the case in some Caucasian languages)?	330

S2b

Matrix of coding of schools with sources of data. ID = Variant ID of database (S2a).

ID	Canonical	Brugmann & Delbrück	Isolating	Hirt	Active-stative	Gamkrelidze & Ivanov
213	1	1900:104-109	0	1937:135f.	0	1995:313
214	1	1900:104-109	1	1937:135f.	1	1995:313
215	0		1	1937:243	0	
216	1	1900:94-100	1	1937:243	1	
217	0		1	1937:234	0	
218	1	1900:405-406	1	1937:235	1	1995:307
219	0		1	1937:234	0	
220	1	1900:102-103	1		1	1995:304
221	1	1900:259-271	NA		NA	
222	0		NA		NA	
223	0		1	1937:227ff	0	
224	0		0	1937:263f.	0	
225	0		1	1937:227ff	0	
226	1	1900:80-83	1	1937:227ff	1	1995:313f
227	0		1	1937:227ff	0	
228	0		0	1937:263f	0	
229	0		1	1937:227ff	0	
230	1	1900:83-85	1	1937:227ff	1	1995:313f
231	0		1	1937:267	0	
232	1	1900:80-83	1	1937:267	1	1995:313f
233	0		1	1937	0	
234	1	1900:80-83	1		1	1995:313f
235	0		0	1937:228ff	0	
236	0		0	1937:228ff	1	1995:313f
237	1	1900:416-417. 1893:475	1	1937:228ff	0	
238	NA		NA		NA	
239	NA		NA		NA	
240	NA		NA		NA	
241	NA		NA		NA	
242	NA		NA		NA	
243	NA		NA		NA	
244	1	1893:189-190	0	1934:29ff	1	1995:244f.
245	1	1893:189-190	0	1934:29ff	0	1995:249
246	1	1893:185-187	0	1934:29ff	0	1995:245
247	1	1893:184-187	0	1934:29ff	1	1995:246
248	1	1893:188	0	1934:29ff	NA	
249	1	1893:180-191	0	1934:29ff	0	1995:245ff
250	0	1893:180-191	0	1934:29ff	0	1995:245
251	0		0	1934:29ff	1	1995:245
252	0		0	1934:29ff	1	1995:245
253	1	1911: 412-427	0	1937:237	1	1995:253f
254	1		0	1937:237	0	1995:253f
255	0		0	1937:237	NA	
256	1		0	1937:237	0	1995:253f
257	0		0	1937:237	0	1995:253f
258	0		0	1937:237	NA	
259	0		0	1937:237	NA	
260	0		0	1934:29ff	1	1995:242

261	0		0	1934:29ff	0	1995:242
262	1	1893:172-400	0	1934:29ff	1	1995:242
263	1	1893:172-400	0	1934:29ff	1	1995:242
264	0		0	1934:29ff	0	1995:242
265	1	1893:132-133	0	1934:27ff	0	1995:243
266	1	1893:132-133	1	1934:27ff	1	
267	1	1893:132-133	1	1934:27ff	1	1995:238
268	0		0	1934:27ff	0	
269	0		1	1937:236f.	0	
270	0		0		0	
271	0		0		0	
272	0		1	1937:236f.	0	
273	0		0		0	
274	1				1	1995:242
275	0		NA		0	
276	1	1916:670-671	0	1928:357ff	0	1995:258
277	0		0		0	1995:258
278	0		0		0	
279	0		NA		0	
280	1		NA		0	
281	0		NA		0	
282	0		NA		0	
283	1		NA		1	
284	0		NA		0	
285	1	1916:595-596	0	1928:357ff	0	1995:258
286	0		0	1928:357ff	0	
287	0		0	1928:357ff	0	
288	0		NA		1	1995:258
289	1		NA		0	
290	0		NA		0	
291	0		NA		NA	
292	1		NA		NA	
293	0		NA		NA	
294	0		NA		0	
295	0	1897:242-255	0	1928:357ff	0	
296	0		0	1928:357ff	0	
297	0		0	1928:357ff	0	
298	0		0	1928:357ff	0	
299	0		0	1928:357ff	0	
300	1	1916:668-669	0	1928:357ff	1	1995:258
301	0		0	1928:357ff	0	
302	0		1		0	
303	1	1893:187-188	1		1	1995:267
304	0		1		1	1995:267f
305	1		1		0	
306	0		1		0	
307	1		1		1	1995:267f
308	0		1		1	1995:267f
309	1		1		0	
310	0		NA		0	
311	1		NA		1	1995:267f
312	0		NA		1	1995:267f
313	1		NA		0	
314	0		NA		0	
315	1		NA		1	1995:267f
316	0		NA		1	1995:267f
317	1		NA		0	
318	0		1	1934:76ff.	0	
319	1		1	1934:76ff.	1	1995:267f
320	0		1	1934:76ff.	1	1995:267f
321	1		1	1934:76ff.	0	
322	0		1	1934:76ff.	0	
323	1		1	1934:76ff.	1	1995:267f
324	0		1	1934:76ff.	1	1995:267f
325	1		1	1934:76ff.	0	

326	1		1	1934:76ff.	1
327	0		0		0
328	0		0		0
329	1	1893:497-498	NA		NA
330	0		NA		NA

S3a

Categorical features and reconstructed probabilities of traits at the root of the tree. For each number or categorical feature in the column "CFID" (Categorical Feature ID), the top-most level of the hierarchically organized features in the original dataset (Grids in the dataset) are given by their name (e.g., Alignment), followed by a vertical line and the name of a Feature (e.g.. Noun: Present Progressive). This Feature may or may not correspond to a categorical feature; for this matter the combination of Grid|Feature may recur in several multistate characters. Each categorical feature has a unique ID number (1-64) and each trait has a unique ID number (A1-WO50). Each categorical feature is described as Grid|Feature|Variant (dataset terms), which are given next to the categorical feature ID. The combination of values (1/0), which represent the unique trait, are given in the row below each cell with a value of a categorical feature. The ID in the row gives the unique trait ID and the unique label describes the trait.

CFID = ID of multistate character block (for reference in text). Label = descriptive property label; ID = unique trait ID (A = alignment, NM = nominal morphology, T = tense, VM = verbal morphology, WO = word order); Variant (1-4) = Variant of multistate character from DiACL, given as Grid|Feature|Variant (see S2a); Result = reconstructed probability of presence for each trait at the protolanguage state.

Label	ID	(Feature 1)	(Feature 2)	(Feature 3)	(Feature 4)
Present-Past		ALIGNMENT; Compare PROG-PAST; PAST_A=PROG_O	ALIGNMENT; Compare PROG-PAST; PROG_A=PST_O	ALIGNMENT; Compare PROG-PAST; PROG_So=PST_So	
Present-Past: No marking difference	A1	0	0	0	
Present-Past: A marking in Present Progressive and Past	A2	0	0	1	
Present-Past: Active-ergative in Present Prog and Past	A3	0	1	1	
Present-Past: All systems	A4	1	1	1	
Alignment; Noun: Present Progressive		ALIGNMENT; Noun: Present Progressive; N:PROG: O=So?	ALIGNMENT; Noun: Present Progressive; N:PROG:A=O?	ALIGNMENT; Noun: Present Progressive; N:PROG:A=Sa?	ALIGNMENT; Noun: Present Progressive; N:PROG:Sa=So?
Noun, Present progressive: Tripartite	A5	0	0	0	1

Noun, Present progressive: Nominative-accusative	A6	0	0	1	1
Noun, Present progressive: No marking	A7	1	1	1	1
Alignment; Noun: Simple Past		ALIGNMENT; Noun: Simple Past; N:PST: O=So?	ALIGNMENT; Noun: Simple Past; N:PST:A=O?	ALIGNMENT; Noun: Simple Past; N:PST:A=Sa?	ALIGNMENT; Noun: Simple Past; N:PST:Sa=So?
Noun, Simple past: Tripartite	A8	0	0	0	1
Noun, Simple past: Nominative-accusative	A9	0	0	1	1
Noun, Simple past: Ergative	A10	1	0	0	1
Noun, Simple past: No marking	A11	1	1	1	1
Alignment; Pronoun: Present Progressive		ALIGNMENT; Pronoun: Present Progressive; P:PROG: O=So?	ALIGNMENT; Pronoun: Present Progressive; P:PROG:A=O?	ALIGNMENT; Pronoun: Present Progressive; P:PROG:A=Sa?	ALIGNMENT; Pronoun: Present Progressive; P:PROG:Sa=So?
Pronoun, Present progressive: Nominative-accusative	A13	0	0	1	1
Pronoun, Present progressive: No marking	A14	1	1	1	1
Alignment; Pronoun: Simple Past		ALIGNMENT; Pronoun: Simple Past; P:PST: O=So?	ALIGNMENT; Pronoun: Simple Past; P:PST:A=O?	ALIGNMENT; Pronoun: Simple Past; P:PST:A=Sa?	ALIGNMENT; Pronoun: Simple Past; P:PST:Sa=So?
Pronoun, Simple past: Tripartite	A15	0	0	0	1
Pronoun, Simple past: Nominative-accusative	A16	0	0	1	1
Pronoun, Simple past: Ergative	A18	1	0	0	1
Pronoun, Simple past: No marking	A19	1	1	1	1
Alignment; Reflexive pronoun in transitive clause, A		ALIGNMENT; Reflexive Pronoun in trans. Clause; REFL-ref-A			
Reflexive not with Agent	A20	0			
Reflexive with Agent	A21	1			

Alignment; Reflexive pronoun in transitive clause, O		ALIGNMENT; Reflexive Pronoun in trans. Clause; REFL-ref-O			
Reflexive not with Object	A22	0			
Reflexive with Object	A23	1			
Alignment; Verb: Present Progressive		ALIGNMENT; Verb: Present Progressive; V:PROG: O=So?	ALIGNMENT; Verb: Present Progressive; V:PROG:A=O?	ALIGNMENT; Verb: Present Progressive; V:PROG:A=Sa?	ALIGNMENT; Verb: Present Progressive; V:PROG:Sa=So?
Verb, Present progressive: Tripartite	A24	0	0	0	1
Verb, Present progressive: Nominative-Accusative	A25	0	0	1	1
Verb, Present progressive: No marking	A26	1	1	1	1
Alignment; Verb: Simple Past		ALIGNMENT; Verb: Simple Past; V:PST: O=So?	ALIGNMENT; Verb: Simple Past; V:PST:A=O?	ALIGNMENT; Verb: Simple Past; V:PST:A=Sa?	ALIGNMENT; Verb: Simple Past; V:PST:Sa=So?
Verb, Simple past: Tripartite	A27	0	0	0	1
Verb, Simple past: Nominative-accusative	A28	0	0	1	1
Verb, Simple past: Ergative	A29	1	0	0	1
Verb, Simple past: No marking	A30	1	1	1	1
Nominal morphology; Case on adjective		Nominal morphology; Case marking; CASE-ADJ			
No case on adjective	NM1	0			
Case on adjective	NM2	1			
Nominal morphology; Case on article		Nominal morphology; Case marking; CASE-ART			
No case on article	NM3	0			
Case on article	NM4	1			
Nominal morphology; Case on NP		Nominal morphology; Case marking; CASE-FIRST	Nominal morphology; Case marking; CASE-LAST		
No rule of case on last/first member of NP	NM5	0	0		
Case on last member of NP	NM6	0	1		
Nominal morphology; Case on noun		Nominal morphology; Case marking; CASE-N			

No case on noun	NM7	0	
Case on noun	NM8	1	
Nominal morphology; Definite suffix on adjective		Nominal morphology; Definiteness marking; ADJ-DEF	
No definite suffix on adjective	NM9	0	
Definite suffix on adjective	NM10	1	
Nominal morphology; Definiteness obligatory		Nominal morphology; Definiteness marking; DEF-FIRST	Nominal morphology; Definiteness marking; DEF-LAST
No obligatory definiteness	NM11	0	0
Obligatory definiteness on last member of NP	NM12	0	1
Obligatory definiteness on first member of NP	NM13	1	0
Nominal morphology; Definite article		Nominal morphology; Definiteness marking; DEF.ART	
No definite article	NM14	0	
Definite article	NM15	1	
Nominal morphology; Definite suffix noun		Nominal morphology; Definiteness marking; N-DEF	
No definite suffix on noun	NM16	0	
Definite suffix on noun	NM17	1	
Nominal morphology; 5 genders		Nominal morphology; Gender / Noun class; <5 GENDER	
Fewer than five genders	NM18	0	
More than five genders	NM19	1	
Nominal morphology; Noun class for animates		Nominal morphology; Gender / Noun class; ANIM	
No noun class for animates	NM20	0	
Noun class for animates	NM21	1	
Nominal morphology; Masculine/feminine		Nominal morphology; Gender / Noun class; M/F	

No masculine/feminine distinction	NM22	0		
Masculine/feminine distinction	NM23	1		
Nominal morphology; Neuter		Nominal morphology; Gender / Noun class; NEUTR		
No neuter gender	NM24	0		
Neuter gender	NM25	1		
Nominal morphology; Gender on predicative adjective		Nominal morphology; Gender agreement; PRED-ADJ		
No gender on predicative adjective	NM26	0		
Gender on predicative adjective	NM27	1		
Nominal morphology; More than 7 cases		Nominal morphology; Nominal cases; <7 Cases		
Not more than 7 cases	NM28	0		
More than 7 cases	NM29	1		
Nominal morphology; Agglutination for number		Nominal morphology; Nominal cases; AGG.CASE.NR		
No agglutination for number	NM30	0		
Agglutination for number	NM31	1		
Nominal morphology; Agglutination for case		Nominal morphology; Nominal cases; AGGL.CASE		
No agglutination for case	NM32	0		
Agglutination for case	NM33	1		
Nominal morphology; Genitive/dative		Nominal morphology; Nominal cases; DAT	Nominal morphology; Nominal cases; GEN	Nominal morphology; Nominal cases; GEN/DAT
No genitive or dative	NM34	0	0	0
Genitive but no dative	NM35	0	1	0
Genitive but no dative	NM36	0	1	1
Dative but no genitive	NM37	1	0	0

Genitive and dative	NM38	1	1	1
Nominal morphology; Case difference A/O		Nominal morphology; Nominal cases; O-case		
No case difference A and O	NM39	0		
Case difference A and O	NM40	1		
Nominal morphology; Peripheral cases		Nominal morphology; Nominal cases; OBL-Cases		
No peripheral cases	NM41	0		
Peripheral cases	NM42	1		
Nominal morphology; Vocative		Nominal morphology; Nominal cases; VOC		
No Vocative	NM43	0		
Vocative	NM44	1		
Nominal morphology; Agreement on prepositions		Nominal morphology; Preposition agreement; PRON-AGR		
No agreement on prepositions	NM45	0		
Agreement on prepositions	NM46	1		
Nominal morphology; More tha 7 pronominal cases		Nominal morphology; Pronominal Cases; <7 Cases		
Less than 7 pronominal cases (pronouns)	NM47	0		
More than 7 pronominal cases (pronouns)	NM48	1		
Nominal morphology; Agglutination for cases (pronouns)		Nominal morphology; Pronominal Cases; AGGL.CASE		
No agglutination for number (pronouns)	NM49	0		
Agglutination for number (pronouns)	NM50	1		
Nominal morphology; Agglutination for case (pronouns)		Nominal morphology; Pronominal Cases; AGGL.CASE.NR		

Agglutination for case (pronouns)	NM51	0
No agglutination for case (pronouns)	NM52	1
Nominal morphology; Difference A/O (pronouns)		Nominal morphology; Pronominal Cases; A $\neq$ O
No difference A and O (pronouns)	NM53	0
Difference A and O (pronouns)	NM54	1
Nominal morphology; Difference O/DAT (pronouns)		Nominal morphology; Pronominal Cases; DAT $\neq$ O
No difference O and Dative (pronouns)	NM55	0
Difference, O and Dative (pronouns)	NM56	1
Nominal morphology; Peripheral cases (pronouns)		Nominal morphology; Pronominal Cases; OBL-Cases
No peripheral cases (pronouns)	NM57	0
Peripheral cases (pronouns)	NM58	1
Nominal morphology; Vocative (pronouns)		Nominal morphology; Pronominal Cases; VOC
No Vocative (pronouns)	NM59	0
Vocative (pronouns)	NM60	1
Tense; Synthetic Present progressive		TENSE; Continous present; Present ?
No synthetic Present progressive	T1	0
Synthetic Present progressive	T2	1
Tense; Present progressive by auxiliary		TENSE; Continous present; Progressive present
No Present progressive by auxiliary	T3	0

Present progressive by auxiliary	T4	1		
Tense; Future by auxiliary		TENSE; Future; FUT.AUX		
No Future by auxiliary	T5	0		
Future by auxiliary	T6	1		
Tense; Future by participle		TENSE; Future; FUT.Participle		
No Future by participle	T7	0		
Future by participle	T8	1		
Tense; Future by particle		TENSE; Future; FUT.Particle		
No Future by particle	T9	0		
Future by particle	T10	1		
Tense; Synthetic future		TENSE; Future; FUT.Synth		
No synthetic Future	T11	0		
Synthetic Future	T12	1		
Tense; Future by perfect		TENSE; Future; PERF.FUT		
No Future by perfect	T13	0		
Future by perfect	T14	1		
Verbal morphology; Present progressive A Agreement		Verbal morphology; present progressive, A; PROG:A-AGR-FULL	Verbal morphology; present progressive, A; PROG:A-Gender-AGR	Verbal morphology; present progressive, A; PROG:NO-A-AGR
Present progressive: Syncretic A Agrement	VM1	0	0	0
Present progressive: No A Agreement	VM2	0	0	1
Present progressive: Gender A Agreement	VM3	0	1	0
Present progressive: Full A Agreement	VM4	1	0	0
Present progressive: Full and Gender A Agreement	VM5	1	1	0

Verbal morphology; Present progressive Dative Agreement		Verbal morphology; present progressive, DAT; PROG:DAT-AGR-FULL	Verbal morphology; present progressive, DAT; PROG:DAT-Gender-AGR	Verbal morphology; present progressive, DAT; PROG:NO-DAT-AGR
Present progressive: Syncretic Dative Agreement	VM6	0	0	0
Present progressive: No Dative Agreement	VM7	0	0	1
Present progressive: Full Dative Agreement	VM8	1	0	0
Verbal morphology; Present progressive O Agreement		Verbal morphology; present progressive, O; PROG:NO-O-AGR	Verbal morphology; present progressive, O; PROG:O-AGR-FULL	Verbal morphology; present progressive, O; PROG:O-Gender-AGR
Present progressive: Syncretic O Agreement	VM9	0	0	0
Present progressive: Full O Agreement	VM10	0	1	0
Present progressive: No O Agreement	VM11	1	0	0
Verbal morphology; Simple past A Agreement		Verbal morphology; simple PAST, A; PST:A-AGR-FULL	Verbal morphology; simple PAST, A; PST:A-Gender-AGR	Verbal morphology; simple PAST, A; PST:NO-A-AGR
Simple past: Syncretic A Agreement	VM12	0	0	0
Simple past: No A Agreement	VM13	0	0	1
Simple past: Gender A Agreement	VM14	0	1	0
Simple past: Full A Agreement	VM16	1	0	0
Simple past: Full and Gender A Agreement	VM17	1	1	0
Verbal morphology; Simple past Dative Agreement		Verbal morphology; simple PAST, DAT; PST:DAT-AGR-FULL	Verbal morphology; simple PAST, DAT; PST:DAT-Gender-AGR	Verbal morphology; simple PAST, DAT; PST:NO-DAT-AGR
Simple past: Syncretic Dative Agreement	VM18	0	0	0

Simple past: No Dative Agreement	VM19	0	0	1
Simple past: Gender Dative Agreement	VM20	0	1	0
Simple past: Full Dative Agreement	VM21	1	0	0
Verbal morphology; Simple past O Agreement		Verbal morphology; simple PAST, O; PST:NO-O-AGR	Verbal morphology; simple PAST, O; PST:O-AGR-FULL	Verbal morphology; simple PAST, O; PST:O-Gender-AGR
Simple past: Syncretic O Agreement	VM22	0	0	0
Simple past: Gender O Agreement	VM23	0	1	0
Simple past: Gender and full O Agreement	VM24	0	1	1
Simple past: No O Agreement	VM25	1	0	0
Word order; Postpostions		Word order; Adpositions; Post		
No Postpositions	WO1	0		
Postpositions	WO2	1		
Word order; Prepositions		Word order; Adpositions; Prep		
No Prepositions	WO3	0		
Prepositions	WO4	1		
Word order; Clitic finite V		Word order; Clitic pronouns finite verb; 2nd position	Word order; Clitic pronouns finite verb; OV	Word order; Clitic pronouns finite verb; VO
Clitic finite V: Category irrelevant	WO5	0	0	0
Clitic finite V: VO	WO6	0	0	1
Clitic finite V: OV	WO7	0	1	0
Clitic finite V: VO and OV	WO8	0	1	1
Clitic finite V: V2	WO9	1	0	0
Word order; Infinitive V		Word order; Clitic pronouns infinitive; 2nd position	Word order; Clitic pronouns infinitive; OV	Word order; Clitic pronouns infinitive; VO
Clitic Infinitive V: Category irrelevant	WO10	0	0	0
Clitic Infinitive V: VO	WO11	0	0	1
Clitic Infinitive V: OV	WO12	0	1	0
Clitic Infinitive V: VO and OV	WO13	0	1	1
Clitic Infinitive V: V2	WO14	1	0	0

Word order; Clitic particle V		Word order; Clitic pronouns participle; 2nd position	Word order; Clitic pronouns participle; OV	Word order; Clitic pronouns participle; VO	
Clitic Participle V: Category irrelevant	WO15	0	0	0	
Clitic Participle V: VO	WO16	0	0	1	
Clitic Participle V: OV	WO17	0	1	0	
Clitic Participle V: VO and OV	WO18	0	1	1	
Clitic Participle V: V2	WO19	1	0	0	

Word order; Infinitive		Word order; Infinitive; OV	Word order; Infinitive; VO		
Infinitive: Category irrelevant	WO20	0	0		
Infinitive: VO	WO21	0	1		
Infinitive: OV	WO22	1	0		
Infinitive: OV and VO	WO23	1	1		

Word order; Main clause		Word order; Main clauses; SOV	Word order; Main clauses; SVO	Word order; Main clauses; V2	Word order; Main clauses; VSO
Main clause: VSO	WO24	0	0	0	1
Main clause: V2	WO25	0	0	1	0
Main clause: SVO	WO26	0	1	0	0
Main clause: SOV	WO27	1	0	0	0
Main clause: SOV/SVO	WO28	1	1	0	0

Word order; Posessor-Noun		Word order; Noun-Possessor; N-Poss	Word order; Noun-Possessor; Poss-N
Possessor - Noun	WO29	0	1
Noun - Possessor	WO30	1	0
Possessor – Noun and Noun – Possessor	WO31	1	1

Word order; Noun-Adjective		Word order; Noun-adjective; AN	Word order; Noun-adjective; NA
Noun – Adjective	WO32	0	1
Adjective - Noun	WO33	1	0
Adjective – Noun and Noun – Adjective	WO34	1	1

Word order; Nound - Relative		Word order; Noun-relative clause; NRel	Word order; Noun-relative clause; RelN
NRel: Category irrelevant	WO35	0	0
NRel: Relative - Noun	WO36	0	1
NRel: Noun - Relative	WO37	1	0

NRel: Relative - Noun and Noun - Relative	WO38	1	1

Word order; Participle		Word order; Participle; OV	Word order; Participle; VO
Participle - O	WO39	0	1
O - Participle	WO40	1	0
Participle - O and O - Participle	WO41	1	1

Word order; Subordinate clause		Word order; Subordinate clause; SOV	Word order; Subordinate clause; SVO	Word order; Subordinate clause; V2	Word order; Subordinate clause; VSO
Subordinate clause: VSO	WO42	0	0	0	1
Subordinate clause: V2	WO43	0	0	1	0
Subordinate clause: SVO	WO44	0	1	0	0
Subordinate clause: SOV	WO45	1	0	0	0
Subordinate clause: SOV and SVO	WO46	1	1	0	0

Word order; WH - Verb		Word order; WH-element; WH-V
No WH - Verb (initial or not)	WO47	0
WH - Verb (initial or not)	WO48	1

Word order; WH-initial		Word order; WH-element; WH-initial
No WH-initial (no inversion)	WO49	0
WH-initial (no inversion)	WO50	1

S3b

Coded variants of S2b (different schools) transformed into the traits of S3a, organized by their categorical features. ID = trait ID of S3a.

ID	Label	Canonical	Isolating	Active-stative
A1	Present-Past: No marking difference	0	0	0
A2	Present-Past: A marking in Present Progressive and Past	1	1	1
A3	Present-Past: Active-ergative in Present Prog and Past	0	0	0
A4	Present-Past: All systems	0	0	0
A5	Noun, Present progressive: Tripartite	0	0	0
A6	Noun, Present progressive: Nominative-accusative	1	0	0
A7	Noun, Present progressive: No marking	0	1	0
A8	Noun, Simple past: Tripartite	0	0	0
A9	Noun, Simple past: Nominative-accusative	1	0	0
A10	Noun, Simple past: Ergative	0	0	0
A11	Noun, Simple past: No marking	0	1	0
A13	Pronoun, Present progressive: Nominative-accusative	1	NA	0
A14	Pronoun, Present progressive: No marking	0	NA	0
A15	Pronoun, Simple past: Tripartite	0	NA	0
A16	Pronoun, Simple past: Nominative-accusative	1	NA	0

A18	Pronoun, Simple past: Ergative	0	NA	0
A19	Pronoun, Simple past: No marking	0	NA	0
A20	Reflexive not with Agent	0	NA	NA
A21	Reflexive with Agent	1	NA	NA
A22	Reflexive not with Object	1	NA	NA
A23	Reflexive with Object	0	NA	NA
A24	Verb, Present progressive: Tripartite	0	0	0
A25	Verb, Present progressive: Nominative-Accusative	1	0	0
A26	Verb, Present progressive: No marking	0	1	0
A27	Verb, Simple past: Tripartite	0	0	0
A28	Verb, Simple past: Nominative-accusative	1	0	0
A29	Verb, Simple past: Ergative	0	0	0
A30	Verb, Simple past: No marking	0	1	0
NM1	No case on adjective	0	1	0
NM2	Case on adjective	1	0	1
NM3	No case on article	1	1	1
NM4	Case on article	0	0	0
NM5	No rule of case on last/first member of NP	1	1	0
NM6	Case on last member of NP	0	0	1
NM7	No case on noun	0	1	0
NM8	Case on noun	1	0	1
NM9	No definite suffix on adjective	1	1	1
NM10	Definite suffix on adjective	0	0	0
NM11	No obligatory definiteness	1		1
NM12	Obligatory definiteness on last member of NP	0	1	0
NM13	Obligatory definiteness on first member of NP	0	0	0
NM14	No definite article	1	0	1
NM15	Definite article	0	1	0
NM16	No definite suffix on noun	1	1	1
NM17	Definite suffix on noun	0	0	0
NM18	Less than five genders	1	1	1
NM19	More than five genders	0	0	0
NM20	No noun class for animates	0	0	0
NM21	Noun class for animates	1	1	1
NM22	No masculine/feminine distinction	0	1	1
NM23	Masculine/feminine distinction	1	0	0
NM24	No neuter gender	0	0	0
NM25	Neuter gender	1	1	1
NM26	No gender on predicative adjective	0	NA	0
NM27	Gender on predicative adjective	1	NA	1
NM28	Not more than 7 cases	0	0	0
NM29	More than 7 cases	1	1	1
NM30	No agglutination for number	1	1	0
NM31	Agglutination for number	0	0	1
NM32	No agglutination for case	1	1	0
NM33	Agglutination for case	0	0	1
NM34	No genitive or dative	0	1	0
NM35	Genitive but no dative	0	0	0
NM36	Genitive but no dative	0	0	0
NM37	Dative but no genitive	0	0	0
NM38	Genitive and dative	1	0	0
NM39	No case difference A and O	0	1	0
NM40	Case difference A and O	1	0	1
NM41	No peripheral cases	0	1	1
NM42	Peripheral cases	1	0	0
NM43	No Vocative	0	1	NA
NM44	Vocative	1	0	NA
NM45	No agreement on prepositions	1	NA	1
NM46	Agreement on prepositions	0	NA	0
NM47	Less than 7 pronominal cases (pronouns)	1	1	1
NM48	More than 7 pronominal cases (pronouns)	0	0	0
NM49	No agglutination for number (pronouns)	1	1	NA
NM50	Agglutination for number (pronouns)	0	0	NA

NM51	Agglutination for case (pronouns)	0	0	NA
NM52	No agglutination for case (pronouns)	1	1	NA
NM53	No difference A and O (pronouns)	0	1	0
NM54	Difference A and O (pronouns)	1	0	1
NM55	No difference O and Dative (pronouns)	0	1	1
NM56	Difference, O and Dative (pronouns)	1	0	0
NM57	No peripheral cases (pronouns)	0	1	1
NM58	Peripheral cases (pronouns)	1	0	0
NM59	No Vocative (pronouns)	1	1	NA
NM60	Vocative (pronouns)	0	0	NA
T1	No synthetic Present progressive	1	1	1
T2	Synthetic Present progressive	0	0	0
T3	No Present progressive by auxiliary	1	1	1
T4	Present progressive by auxiliary	0	0	0
T5	No Future by auxiliary	1	1	1
T6	Future by auxiliary	0	0	0
T7	No Future by participle	1	1	1
T8	Future by participle	0	0	0
T9	No Future by particle	1	1	1
T10	Future by particle	0	0	0
T11	No synthetic Future	1	1	1
T12	Synthetic Future	0	0	0
T13	No Future by perfect	1	1	1
T14	Future by perfect	0	0	0
VM1	Present progressive: Syncretic A Agrement	0	1	0
VM2	Present progressive: No A Agreement	0	0	0
VM3	Present progressive: Gender A Agreement	0	0	0
VM4	Present progressive: Full A Agreement	1	0	1
VM5	Present progressive: Full and Gender A Agreement	0	0	0
VM6	Present progressive: Syncretic Dative Agreement	0	NA	NA
VM7	Present progressive: No Dative Agreement	1	NA	NA
VM8	Present progressive: Full Dative Agreement	0	NA	NA
VM9	Present progressive: Syncretic O Agreement	0	NA	1
VM10	Present progressive: Full O Agreement	0	NA	0
VM11	Present progressive: No O Agreement	1	NA	0
VM12	Simple past: Syncretic A Agreement	0	1	1
VM13	Simple past: No A Agreement	0	0	0
VM14	Simple past: Gender A Agreement	0	0	0
VM16	Simple past: Full A Agreement	1	0	0
VM17	Simple past: Full and Gender A Agreement	0	0	0
VM18	Simple past: Syncretic Dative Agreement	0	NA	0
VM19	Simple past: No Dative Agreement	1	NA	1
VM20	Simple past: Gender Dative Agreement	0	NA	0
VM21	Simple past: Full Dative Agreement	0	NA	0
VM22	Simple past: Syncretic O Agreement	0	NA	1
VM23	Simple past: Gender O Agreement	0	NA	0
VM24	Simple past: Gender and full O Agreement	0	NA	0
VM25	Simple past: No O Agreement	1	NA	0
WO1	No Postpositions	0	0	0
WO2	Postpositions	1	1	1
WO3	No Prepositions	0	0	0
WO4	Prepositions	1	0	0
WO5	Clitic finite V: Category irrelevant	0	0	0
WO6	Clitic finite V: VO	0	0	0
WO7	Clitic finite V: OV	0	0	1
WO8	Clitic finite V: VO and OV	0	0	0
WO9	Clitic finite V: V2	0	0	0
WO10	Clitic Infinitive V: Category irrelevant	NA	NA	NA
WO11	Clitic Infinitive V: VO	NA	NA	NA
WO12	Clitic Infinitive V: OV	NA	NA	NA
WO13	Clitic Infinitive V: VO and OV	NA	NA	NA
WO14	Clitic Infinitive V: V2	NA	NA	NA
WO15	Clitic Participle V: Category irrelevant	NA	NA	NA

WO16	Clitic Participle V: VO	NA	NA	NA
WO17	Clitic Participle V: OV	NA	NA	NA
WO18	Clitic Participle V: VO and OV	NA	NA	NA
WO19	Clitic Participle V: V2	NA	NA	NA
WO20	Infinitive: Category irrelevant	0	0	0
WO21	Infinitive: VO	0	1	0
WO22	Infinitive: OV	1	1	1
WO23	Infinitive: OV and VO	0	1	0
WO24	Main clause: VSO	0	1	0
WO25	Main clause: V2	0	0	0
WO26	Main clause: SVO	0	1	0
WO27	Main clause: SOV	1	1	1
WO28	Main clause: SOV/SVO	0	1	0
WO29	Possessor - Noun	1	1	1
WO30	Noun - Possessor	0	1	0
WO31	Possessor ? Noun and Noun ? Possessor	0	1	0
WO32	Noun ? Adjective	0	1	0
WO33	Adjective - Noun	1	1	1
WO34	Adjective ? Noun and Noun ? Adjective	0	1	0
WO35	NRel: Category irrelevant	0	0	0
WO36	NRel: Relative - Noun	1	1	1
WO37	NRel: Noun - Relative	0	1	0
WO38	NRel: Relative - Noun and Noun - Relative	0	1	0
WO39	Participle - O	0	1	0
WO40	O - Participle	1	1	1
WO41	Participle - O and O - Participle	0	1	0
WO42	Subordinate clause: VSO	0	1	0
WO43	Subordinate clause: V2	0	0	0
WO44	Subordinate clause: SVO	0	1	0
WO45	Subordinate clause: SOV	1	1	1
WO46	Subordinate clause: SOV and SVO	0	1	0
WO47	No WH - Verb (initial or not)	1	NA	NA
WO48	WH - Verb (initial or not)	0	NA	NA
WO49	No WH-initial (no inversion)	0	NA	NA
WO50	WH-initial (no inversion)	1	NA	NA

S4

Reconstruction probabilities for values of each variable.

ID	Value	SP	Dir	Unif
A1	Present-Past: A marking in Present Progressive and Past	0.781	0.312	0.311
A2	Present-Past: Active-ergative in Present Progressive and Past	0.100	0.200	0.199
A3	Present-Past: All systems	0.091	0.251	0.251
A4	Present-Past: No marking difference	0.026	0.236	0.237
A5	Noun, Present progressive: No marking	0.314	0.315	0.311
A6	Noun, Present progressive: Nominative-accusative	0.654	0.363	0.360
A7	Noun, Present progressive: Tripartite	0.030	0.321	0.327
A8	Noun, Simple past: Ergative	0.125	0.189	0.186
A9	Noun, Simple past: No marking	0.302	0.271	0.273
A10	Noun, Simple past: Nominative-accusative	0.537	0.297	0.301
A11	Noun, Simple past: Tripartite	0.034	0.241	0.238
A12	Pronoun, Present progressive: No marking	0.092	0.463	0.460
A13	Pronoun, Present progressive: Nominative-accusative	0.907	0.536	0.539

A14	Pronoun, Simple past: Ergative	0.082	0.142	0.145
A15	Pronoun, Simple past: No marking	0.060	0.197	0.196
A16	Pronoun, Simple past: Nominative-accusative	0.830	0.446	0.447
A17	Pronoun, Simple past: Tripartite	0.026	0.213	0.210
A18	Reflexive not with Agent	0.040	0.505	0.505
A19	Reflexive with Agent	0.959	0.494	0.494
A20	Reflexive not with Object	0.957	0.546	0.548
A21	Reflexive with Object	0.042	0.453	0.451
A22	Verb, Present progressive: A=O/So=O	0.024	0.240	0.242
A23	Verb, Present progressive: No marking	0.060	0.244	0.243
A24	Verb, Present progressive: Nominative-Accusative	0.873	0.276	0.276
A25	Verb, Present progressive: Tripartite	0.041	0.238	0.237
A26	Verb, Simple past: Double oblique	0.030	0.192	0.194
A27	Verb, Simple past: Ergative	0.105	0.168	0.174
A28	Verb, Simple past: No marking	0.105	0.202	0.197
A29	Verb, Simple past: Nominative-accusative	0.674	0.251	0.252
A30	Verb, Simple past: Tripartite	0.083	0.185	0.181
NM1	Case on adjective	0.559	0.617	0.608
NM2	No case on adjective	0.440	0.382	0.391
NM3	Case on article	0.267	0.618	0.621
NM4	No case on article	0.732	0.381	0.378
NM5	Case on last member of NP	0.075	0.120	0.124
NM6	No rule of case on last/first member of NP	0.924	0.879	0.875
NM7	Case on noun	0.744	0.517	0.521
NM8	No case on noun	0.255	0.482	0.478
NM9	Definite suffix on adjective	0.076	0.301	0.302
NM10	No definite suffix on adjective	0.923	0.698	0.697
NM11	No obligatory definiteness	0.857	0.556	0.563
NM12	Obligatory definiteness on first member of NP	0.075	0.096	0.097
NM13	Obligatory definiteness on last member of NP	0.067	0.347	0.339
NM14	Definite article	0.059	0.086	0.084
NM15	No definite article	0.940	0.913	0.915
NM16	Definite suffix on noun	0.087	0.196	0.194
NM17	No definite suffix on noun	0.912	0.803	0.805
NM18	Fewer than five genders	0.987	0.582	0.588
NM19	More than five genders	0.012	0.417	0.411
NM20	No noun class for animates	0.913	0.591	0.588
NM21	Noun class for animates	0.087	0.408	0.411
NM22	Masculine/feminine distinction	0.683	0.596	0.597
NM23	No masculine/feminine distinction	0.316	0.403	0.402
NM24	Neuter gender	0.854	0.965	0.968
NM25	No neuter gender	0.145	0.034	0.031

NM26	Gender on predicative adjective	0.672	0.624	0.620
NM27	No gender on predicative adjective	0.327	0.375	0.379
NM28	More than 7 cases	0.416	0.950	0.948
NM29	Not more than 7 cases	0.583	0.049	0.051
NM30	Agglutination for number	0.229	0.360	0.356
NM31	No agglutination for number	0.770	0.639	0.643
NM32	Agglutination for case	0.218	0.429	0.425
NM33	No agglutination for case	0.781	0.570	0.574
NM34	Dative but no genitive	0.069	0.250	0.251
NM35	Genitive and dative	0.607	0.398	0.404
NM36	Genitive but no dative	0.138	0.234	0.233
NM37	No genitive or dative	0.183	0.115	0.109
NM38	Case difference A and O	0.715	0.561	0.561
NM39	No case difference A and O	0.284	0.438	0.438
NM40	No peripheral cases	0.017	0.003	0.003
NM41	Peripheral cases	0.982	0.996	0.996
NM42	No Vocative	0.114	0.054	0.053
NM43	Vocative	0.885	0.945	0.946
NM44	Agreement on prepositions	0.032	0.111	0.109
NM45	No agreement on prepositions	0.967	0.888	0.890
NM46	Fewer than 7 pronominal cases (pronouns)	0.969	0.543	0.546
NM47	More than 7 pronominal cases (pronouns)	0.031	0.456	0.453
NM48	Agglutination for number (pronouns)	0.081	0.189	0.184
NM49	No agglutination for number (pronouns)	0.918	0.810	0.815
NM50	Agglutination for case (pronouns)	0.962	0.759	0.758
NM51	No agglutination for case (pronouns)	0.037	0.240	0.241
NM52	Difference A and O (pronouns)	0.934	0.742	0.741
NM53	No difference A and O (pronouns)	0.065	0.257	0.258
NM54	Difference, O and Dative (pronouns)	0.429	0.529	0.519
NM55	No difference O and Dative (pronouns)	0.570	0.470	0.480
NM56	No peripheral cases (pronouns)	0.077	0.012	0.011
NM57	Peripheral cases (pronouns)	0.922	0.987	0.988
NM58	No Vocative (pronouns)	0.985	0.549	0.551
NM59	Vocative (pronouns)	0.014	0.450	0.449
T1	No synthetic Present progressive	0.078	0.382	0.378
T2	Synthetic Present progressive	0.922	0.617	0.621
T3	No Present progressive by auxiliary	0.676	0.561	0.558
T4	Present progressive by auxiliary	0.324	0.438	0.441
T5	Future by auxiliary	0.371	0.425	0.427
T6	No Future by auxiliary	0.628	0.574	0.573
T7	Future by participle	0.276	0.762	0.767
T8	No Future by participle	0.724	0.237	0.232

T9	Future by particle	0.106	0.367	0.368
T10	No Future by particle	0.893	0.632	0.631
T11	No synthetic Future	0.730	0.661	0.661
T12	Synthetic Future	0.269	0.339	0.338
T13	Future by perfect	0.019	0.271	0.278
T14	No Future by perfect	0.980	0.728	0.721
VM1	Present progressive: Full A Agreement	0.657	0.328	0.326
VM2	Present progressive: Full and Gender A Agreement	0.024	0.190	0.194
VM3	Present progressive: Gender A Agreement	0.040	0.189	0.189
VM4	Present progressive: No A Agreement	0.064	0.183	0.178
VM5	Present progressive: Syncretic A Agreement	0.212	0.107	0.110
VM6	Present progressive: Full Dative Agreement	0.020	0.303	0.308
VM7	Present progressive: No Dative Agreement	0.922	0.419	0.419
VM8	Present progressive: Syncretic Dative Agreement	0.057	0.276	0.271
VM9	Present progressive: Full O Agreement	0.021	0.302	0.308
VM10	Present progressive: No O Agreement	0.920	0.417	0.415
VM11	Present progressive: Syncretic O Agreement	0.057	0.280	0.276
VM12	Simple past: Full A Agreement	0.031	0.201	0.203
VM13	Simple past: Full and Gender A Agreement	0.270	0.197	0.199
VM14	Simple past: Gender A Agreement	0.048	0.199	0.200
VM15	Simple past: No A Agreement	0.055	0.194	0.198
VM16	Simple past: Syncretic A Agreement	0.207	0.206	0.198
VM17	Simple past: Full A Agreement	0.386	0.201	0.203
VM18	Simple past: Full Dative Agreement	0.022	0.239	0.237
VM19	Simple past: Gender Dative Agreement	0.020	0.236	0.238
VM20	Simple past: No Dative Agreement	0.893	0.301	0.302
VM21	Simple past: Syncretic Dative Agreement	0.064	0.221	0.221
VM22	Simple past: Gender and full O Agreement	0.027	0.189	0.182
VM23	Simple past: Gender O Agreement	0.036	0.184	0.182
VM24	Simple past: No O Agreement	0.785	0.288	0.296
VM25	Simple Past: Syncretic and gender A Agreement	0.067	0.172	0.170
VM26	Simple past: Syncretic O Agreement	0.083	0.165	0.167
WO1	No Postpositions	0.150	0.051	0.050
WO2	Postpositions	0.849	0.948	0.949
WO3	No Prepositions	0.738	0.819	0.816
WO4	Prepositions	0.261	0.180	0.183
WO5	Clitic finite V: Category irrelevant	0.442	0.170	0.167
WO6	Clitic finite V: OV	0.141	0.164	0.164
WO7	Clitic finite V: V2	0.161	0.269	0.270
WO8	Clitic finite V: VO	0.181	0.182	0.185
WO9	Clitic finite V: VO and OV	0.073	0.212	0.211
WO10	Clitic Infinitive V: Category irrelevant	0.433	0.137	0.136

WO11	Clitic Infinitive V: OV	0.189	0.208	0.205
WO12	Clitic Infinitive V: V2	0.161	0.249	0.254
WO13	Clitic Infinitive V: VO	0.169	0.204	0.209
WO14	Clitic Infinitive V: VO and OV	0.045	0.2	0.194
WO15	Clitic Participle V: Category irrelevant	0.418	0.140	0.149
WO16	Clitic Participle V: OV	0.197	0.201	0.196
WO17	Clitic Participle V: V2	0.148	0.240	0.236
WO18	Clitic Participle V: VO	0.181	0.207	0.214
WO19	Clitic Participle V: VO and OV	0.054	0.21	0.203
WO20	Infinitive: Category irrelevant	0.063	0.261	0.261
WO21	Infinitive: OV	0.805	0.404	0.404
WO22	Infinitive: OV and VO	0.038	0.245	0.248
WO23	Infinitive: VO	0.092	0.087	0.086
WO24	Main clause: SOV	0.904	0.640	0.638
WO25	Main clause: SOV/SVO	0.009	0.134	0.137
WO26	Main clause: SVO	0.034	0.027	0.029
WO27	Main clause: V2	0.030	0.058	0.056
WO28	Main clause: VSO	0.021	0.139	0.138
WO29	Noun - Possessor	0.346	0.337	0.336
WO30	Possessor - Noun	0.585	0.285	0.288
WO31	Possessor - Noun and Noun - Possessor	0.068	0.377	0.374
WO32	Adjective - Noun	0.869	0.535	0.538
WO33	Adjective - Noun and Noun - Adjective	0.053	0.306	0.305
WO34	Noun - Adjective	0.076	0.157	0.155
WO35	NRel: Category irrelevant	0.040	0.28	0.282
WO36	NRel: Noun - Relative	0.627	0.118	0.113
WO37	NRel: Relative - Noun	0.291	0.316	0.315
WO38	NRel: Relative - Noun and Noun - Relative	0.040	0.285	0.289
WO39	O - Participle	0.893	0.619	0.618
WO40	Participle - O	0.049	0.053	0.054
WO41	Participle - O and O - Participle	0.056	0.326	0.327
WO42	Subordinate clause: SOV	0.899	0.558	0.560
WO43	Subordinate clause: SOV and SVO	0.012	0.162	0.160
WO44	Subordinate clause: SVO	0.037	0.024	0.026
WO45	Subordinate clause: V2	0.013	0.091	0.089
WO46	Subordinate clause: VSO	0.037	0.163	0.163
WO47	No WH - Verb (initial or not)	0.778	0.801	0.799
WO48	WH - Verb (initial or not)	0.221	0.198	0.200
WO49	No WH-initial (no inversion)	0.212	0.177	0.176
WO50	WH-initial (no inversion)	0.787	0.822	0.823

S5

Transition rates for features, organized by category (S3a).

Feature 1	Feature 1	Rate(F1→F2)
Present-Past: A marking in Present Progressive and Past	Present-Past: Active-ergative in Present Progressive and Past	0.090
Present-Past: A marking in Present Progressive and Past	Present-Past: All systems	0.237
Present-Past: A marking in Present Progressive and Past	Present-Past: No marking difference	0.057
Present-Past: Active-ergative in Present Progressive and Past	Present-Past: A marking in Present Progressive and Past	0.563
Present-Past: Active-ergative in Present Progressive and Past	Present-Past: All systems	0.609
Present-Past: Active-ergative in Present Progressive and Past	Present-Past: No marking difference	0.330
Present-Past: All systems	Present-Past: A marking in Present Progressive and Past	1.663
Present-Past: All systems	Present-Past: Active-ergative in Present Progressive and Past	0.906
Present-Past: All systems	Present-Past: No marking difference	0.462
Present-Past: No marking difference	Present-Past: A marking in Present Progressive and Past	1.552
Present-Past: No marking difference	Present-Past: Active-ergative in Present Progressive and Past	1.569
Present-Past: No marking difference	Present-Past: All systems	1.526
Noun, Present progressive: No marking	Noun, Present progressive: Nominative-accusative	1.044
Noun, Present progressive: No marking	Noun, Present progressive: Tripartite	0.175
Noun, Present progressive: Nominative-accusative	Noun, Present progressive: No marking	0.593
Noun, Present progressive: Nominative-accusative	Noun, Present progressive: Tripartite	0.082
Noun, Present progressive: Tripartite	Noun, Present progressive: No marking	1.563
Noun, Present progressive: Tripartite	Noun, Present progressive: Nominative-accusative	1.819
Noun, Simple past: Ergative	Noun, Simple past: No marking	0.365
Noun, Simple past: Ergative	Noun, Simple past: Nominative-accusative	0.533
Noun, Simple past: Ergative	Noun, Simple past: Tripartite	0.310

Noun, Simple past: No marking	Noun, Simple past: Ergative	0.282
Noun, Simple past: No marking	Noun, Simple past: Nominative-accusative	1.105
Noun, Simple past: No marking	Noun, Simple past: Tripartite	0.194
Noun, Simple past: Nominative-accusative	Noun, Simple past: Ergative	0.113
Noun, Simple past: Nominative-accusative	Noun, Simple past: No marking	0.751
Noun, Simple past: Nominative-accusative	Noun, Simple past: Tripartite	0.095
Noun, Simple past: Tripartite	Noun, Simple past: Ergative	1.455
Noun, Simple past: Tripartite	Noun, Simple past: No marking	1.336
Noun, Simple past: Tripartite	Noun, Simple past: Nominative-accusative	1.629
Pronoun, Present progressive: No marking	Pronoun, Present progressive: Nominative-accusative	0.807
Pronoun, Present progressive: Nominative-accusative	Pronoun, Present progressive: No marking	0.118
Pronoun, Simple past: Ergative	Pronoun, Simple past: No marking	0.301
Pronoun, Simple past: Ergative	Pronoun, Simple past: Nominative-accusative	0.326
Pronoun, Simple past: Ergative	Pronoun, Simple past: Tripartite	0.311
Pronoun, Simple past: No marking	Pronoun, Simple past: Ergative	0.828
Pronoun, Simple past: No marking	Pronoun, Simple past: Nominative-accusative	0.641
Pronoun, Simple past: No marking	Pronoun, Simple past: Tripartite	0.502
Pronoun, Simple past: Nominative-accusative	Pronoun, Simple past: Ergative	0.061
Pronoun, Simple past: Nominative-accusative	Pronoun, Simple past: No marking	0.140
Pronoun, Simple past: Nominative-accusative	Pronoun, Simple past: Tripartite	0.057
Pronoun, Simple past: Tripartite	Pronoun, Simple past: Ergative	1.547
Pronoun, Simple past: Tripartite	Pronoun, Simple past: No marking	1.156
Pronoun, Simple past: Tripartite	Pronoun, Simple past: Nominative-accusative	1.261
Reflexive not with Agent	Reflexive with Agent	0.888
Reflexive with Agent	Reflexive not with Agent	0.042
Reflexive not with Object	Reflexive with Object	0.035
Reflexive with Object	Reflexive not with Object	0.548
Verb, Present progressive: A=O/So=O	Verb, Present progressive: No marking	1.299

Verb, Present progressive: A=O/So=O	Verb, Present progressive: Nominative-Accusative	1.546
Verb, Present progressive: A=O/So=O	Verb, Present progressive: Tripartite	1.192
Verb, Present progressive: No marking	Verb, Present progressive: A=O/So=O	0.587
Verb, Present progressive: No marking	Verb, Present progressive: Nominative-Accusative	1.552
Verb, Present progressive: No marking	Verb, Present progressive: Tripartite	0.734
Verb, Present progressive: Nominative-Accusative	Verb, Present progressive: A=O/So=O	0.037
Verb, Present progressive: Nominative-Accusative	Verb, Present progressive: No marking	0.129
Verb, Present progressive: Nominative-Accusative	Verb, Present progressive: Tripartite	0.044
Verb, Present progressive: Tripartite	Verb, Present progressive: A=O/So=O	0.734
Verb, Present progressive: Tripartite	Verb, Present progressive: No marking	0.963
Verb, Present progressive: Tripartite	Verb, Present progressive: Nominative-Accusative	1.214
Verb, Simple past: Double oblique	Verb, Simple past: Ergative	1.480
Verb, Simple past: Double oblique	Verb, Simple past: No marking	1.291
Verb, Simple past: Double oblique	Verb, Simple past: Nominative-accusative	1.431
Verb, Simple past: Double oblique	Verb, Simple past: Tripartite	1.443
Verb, Simple past: Ergative	Verb, Simple past: Double oblique	0.374
Verb, Simple past: Ergative	Verb, Simple past: No marking	0.463
Verb, Simple past: Ergative	Verb, Simple past: Nominative-accusative	0.536
Verb, Simple past: Ergative	Verb, Simple past: Tripartite	0.705
Verb, Simple past: No marking	Verb, Simple past: Double oblique	0.437
Verb, Simple past: No marking	Verb, Simple past: Ergative	0.753
Verb, Simple past: No marking	Verb, Simple past: Nominative-accusative	1.485
Verb, Simple past: No marking	Verb, Simple past: Tripartite	0.775
Verb, Simple past: Nominative-accusative	Verb, Simple past: Double oblique	0.053
Verb, Simple past: Nominative-accusative	Verb, Simple past: Ergative	0.064
Verb, Simple past: Nominative-accusative	Verb, Simple past: No marking	0.371

Verb, Simple past: Nominative-accusative	Verb, Simple past: Tripartite	0.073
Verb, Simple past: Tripartite	Verb, Simple past: Double oblique	0.508
Verb, Simple past: Tripartite	Verb, Simple past: Ergative	1.002
Verb, Simple past: Tripartite	Verb, Simple past: No marking	0.676
Verb, Simple past: Tripartite	Verb, Simple past: Nominative-accusative	0.811
Case on adjective	No case on adjective	0.562
No case on adjective	Case on adjective	0.408
Case on article	No case on article	0.722
No case on article	Case on article	0.182
Case on last member of NP	No rule of case on last/first member of NP	0.258
No rule of case on last/first member of NP	Case on last member of NP	0.138
Case on noun	No case on noun	0.316
No case on noun	Case on noun	0.718
Definite suffix on adjective	No definite suffix on adjective	0.274
No definite suffix on adjective	Definite suffix on adjective	0.061
No obligatory definiteness	Obligatory definiteness on first member of NP	0.093
No obligatory definiteness	Obligatory definiteness on last member of NP	0.182
Obligatory definiteness on first member of NP	No obligatory definiteness	0.185
Obligatory definiteness on first member of NP	Obligatory definiteness on last member of NP	0.132
Obligatory definiteness on last member of NP	No obligatory definiteness	1.551
Obligatory definiteness on last member of NP	Obligatory definiteness on first member of NP	0.639
Definite article	No definite article	0.135
No definite article	Definite article	0.102
Definite suffix on noun	No definite suffix on noun	0.308
No definite suffix on noun	Definite suffix on noun	0.124
Fewer than five genders	More than five genders	0.017
More than five genders	Fewer than five genders	0.650
No noun class for animates	Noun class for animates	0.070
Noun class for animates	No noun class for animates	0.508
Masculine/feminine distinction	No masculine/feminine distinction	0.214
No masculine/feminine distinction	Masculine/feminine distinction	0.373
Neuter gender	No neuter gender	0.361

No neuter gender	Neuter gender	0.094
Gender on predicative adjective	No gender on predicative adjective	0.311
No gender on predicative adjective	Gender on predicative adjective	0.395
More than 7 cases	Not more than 7 cases	0.728
Not more than 7 cases	More than 7 cases	0.057
Agglutination for number	No agglutination for number	0.353
No agglutination for number	Agglutination for number	0.187
Agglutination for case	No agglutination for case	0.520
No agglutination for case	Agglutination for case	0.218
Dative but no genitive	Genitive and dative	1.332
Dative but no genitive	Genitive but no dative	1.218
Dative but no genitive	No genitive or dative	0.859
Genitive and dative	Dative but no genitive	0.165
Genitive and dative	Genitive but no dative	0.346
Genitive and dative	No genitive or dative	0.161
Genitive but no dative	Dative but no genitive	0.579
Genitive but no dative	Genitive and dative	0.930
Genitive but no dative	No genitive or dative	0.737
No genitive or dative	Dative but no genitive	0.164
No genitive or dative	Genitive and dative	0.137
No genitive or dative	Genitive but no dative	0.238
Case difference A and O	No case difference A and O	0.355
No case difference A and O	Case difference A and O	0.583
No peripheral cases	Peripheral cases	0.046
Peripheral cases	No peripheral cases	0.150
No Vocative	Vocative	0.128
Vocative	No Vocative	0.253
Agreement on prepositions	No agreement on prepositions	0.110
No agreement on prepositions	Agreement on prepositions	0.039
Fewer than 7 pronominal cases (pronouns)	More than 7 pronominal cases (pronouns)	0.038
More than 7 pronominal cases (pronouns)	Fewer than 7 pronominal cases (pronouns)	0.740
Agglutination for number (pronouns)	No agglutination for number (pronouns)	0.324
No agglutination for number (pronouns)	Agglutination for number (pronouns)	0.132
Agglutination for case (pronouns)	No agglutination for case (pronouns)	0.058
No agglutination for case (pronouns)	Agglutination for case (pronouns)	0.434
Difference A and O (pronouns)	No difference A and O (pronouns)	0.085
No difference A and O (pronouns)	Difference A and O (pronouns)	0.396
Difference, O and Dative (pronouns)	No difference O and Dative (pronouns)	0.535

No difference O and Dative (pronouns)	Difference, O and Dative (pronouns)	0.381
No peripheral cases (pronouns)	Peripheral cases (pronouns)	0.056
Peripheral cases (pronouns)	No peripheral cases (pronouns)	0.236
No Vocative (pronouns)	Vocative (pronouns)	0.019
Vocative (pronouns)	No Vocative (pronouns)	0.721
No synthetic Present progressive	Synthetic Present progressive	0.522
Synthetic Present progressive	No synthetic Present progressive	0.071
No Present progressive by auxiliary	Present progressive by auxiliary	0.270
Present progressive by auxiliary	No Present progressive by auxiliary	0.440
Future by auxiliary	No Future by auxiliary	0.463
No Future by auxiliary	Future by auxiliary	0.383
Future by participle	No Future by participle	0.653
No Future by participle	Future by participle	0.093
Future by particle	No Future by particle	0.581
No Future by particle	Future by particle	0.132
No synthetic Future	Synthetic Future	0.165
Synthetic Future	No synthetic Future	0.206
Future by perfect	No Future by perfect	0.360
No Future by perfect	Future by perfect	0.023
Present progressive: Full A Agreement	Present progressive: Full and Gender A Agreement	0.077
Present progressive: Full A Agreement	Present progressive: Gender A Agreement	0.112
Present progressive: Full A Agreement	Present progressive: No A Agreement	0.097
Present progressive: Full A Agreement	Present progressive: Syncretic A Agreement	0.207
Present progressive: Full and Gender A Agreement	Present progressive: Full A Agreement	1.548
Present progressive: Full and Gender A Agreement	Present progressive: Gender A Agreement	1.269
Present progressive: Full and Gender A Agreement	Present progressive: No A Agreement	1.367
Present progressive: Full and Gender A Agreement	Present progressive: Syncretic A Agreement	1.598
Present progressive: Gender A Agreement	Present progressive: Full A Agreement	1.350
Present progressive: Gender A Agreement	Present progressive: Full and Gender A Agreement	0.763
Present progressive: Gender A Agreement	Present progressive: No A Agreement	1.095

Present progressive: Gender A Agreement	Present progressive: Syncretic A Agreement	1.322
Present progressive: No A Agreement	Present progressive: Full A Agreement	1.047
Present progressive: No A Agreement	Present progressive: Full and Gender A Agreement	0.586
Present progressive: No A Agreement	Present progressive: Gender A Agreement	0.807
Present progressive: No A Agreement	Present progressive: Syncretic A Agreement	1.355
Present progressive: Syncretic A Agreement	Present progressive: Full A Agreement	0.115
Present progressive: Syncretic A Agreement	Present progressive: Full and Gender A Agreement	0.109
Present progressive: Syncretic A Agreement	Present progressive: Gender A Agreement	0.147
Present progressive: Syncretic A Agreement	Present progressive: No A Agreement	0.352
Present progressive: Full Dative Agreement	Present progressive: No Dative Agreement	1.441
Present progressive: Full Dative Agreement	Present progressive: Syncretic Dative Agreement	1.328
Present progressive: No Dative Agreement	Present progressive: Full Dative Agreement	0.032
Present progressive: No Dative Agreement	Present progressive: Syncretic Dative Agreement	0.069
Present progressive: Syncretic Dative Agreement	Present progressive: Full Dative Agreement	0.457
Present progressive: Syncretic Dative Agreement	Present progressive: No Dative Agreement	0.669
Present progressive: Full O Agreement	Present progressive: No O Agreement	1.435
Present progressive: Full O Agreement	Present progressive: Syncretic O Agreement	1.329
Present progressive: No O Agreement	Present progressive: Full O Agreement	0.032
Present progressive: No O Agreement	Present progressive: Syncretic O Agreement	0.069
Present progressive: Syncretic O Agreement	Present progressive: Full O Agreement	0.454
Present progressive: Syncretic O Agreement	Present progressive: No O Agreement	0.661

Simple past: Full A Agreement	Simple past: Full and Gender A Agreement	1.733
Simple past: Full A Agreement	Simple past: Gender A Agreement	1.217
Simple past: Full A Agreement	Simple past: No A Agreement	1.299
Simple past: Full A Agreement	Simple past: Syncretic A Agreement	1.729
Simple past: Full and Gender A Agreement	Simple past: Full A Agreement	0.179
Simple past: Full and Gender A Agreement	Simple past: Gender A Agreement	0.236
Simple past: Full and Gender A Agreement	Simple past: No A Agreement	0.284
Simple past: Full and Gender A Agreement	Simple past: Syncretic A Agreement	0.822
Simple past: Gender A Agreement	Simple past: Full A Agreement	0.748
Simple past: Gender A Agreement	Simple past: Full and Gender A Agreement	1.432
Simple past: Gender A Agreement	Simple past: No A Agreement	1.011
Simple past: Gender A Agreement	Simple past: Syncretic A Agreement	1.422
Simple past: No A Agreement	Simple past: Full A Agreement	0.725
Simple past: No A Agreement	Simple past: Full and Gender A Agreement	1.524
Simple past: No A Agreement	Simple past: Gender A Agreement	0.923
Simple past: No A Agreement	Simple past: Syncretic A Agreement	1.527
Simple past: Syncretic A Agreement	Simple past: Full A Agreement	0.265
Simple past: Syncretic A Agreement	Simple past: Full and Gender A Agreement	1.143
Simple past: Syncretic A Agreement	Simple past: Gender A Agreement	0.364
Simple past: Syncretic A Agreement	Simple past: No A Agreement	0.443
Simple past: Full Dative Agreement	Simple past: Gender Dative Agreement	1.016
Simple past: Full Dative Agreement	Simple past: No Dative Agreement	1.462
Simple past: Full Dative Agreement	Simple past: Syncretic Dative Agreement	1.347
Simple past: Gender Dative Agreement	Simple past: Full Dative Agreement	1.005
Simple past: Gender Dative Agreement	Simple past: No Dative Agreement	1.490
Simple past: Gender Dative Agreement	Simple past: Syncretic Dative Agreement	1.378
Simple past: No Dative Agreement	Simple past: Full Dative Agreement	0.033
Simple past: No Dative Agreement	Simple past: Gender Dative Agreement	0.035

Simple past: No Dative Agreement	Simple past: Syncretic Dative Agreement	0.069
Simple past: Syncretic Dative Agreement	Simple past: Full Dative Agreement	0.468
Simple past: Syncretic Dative Agreement	Simple past: Gender Dative Agreement	0.461
Simple past: Syncretic Dative Agreement	Simple past: No Dative Agreement	0.678
Simple past: Gender and full O Agreement	Simple past: Gender O Agreement	1.097
Simple past: Gender and full O Agreement	Simple past: No O Agreement	1.109
Simple past: Gender and full O Agreement	Simple Past: Syncretic and gender A Agreement	1.350
Simple past: Gender and full O Agreement	Simple past: Syncretic O Agreement	1.355
Simple past: Gender O Agreement	Simple past: Gender and full O Agreement	0.924
Simple past: Gender O Agreement	Simple past: No O Agreement	1.008
Simple past: Gender O Agreement	Simple Past: Syncretic and gender A Agreement	1.268
Simple past: Gender O Agreement	Simple past: Syncretic O Agreement	1.285
Simple past: No O Agreement	Simple past: Gender and full O Agreement	0.052
Simple past: No O Agreement	Simple past: Gender O Agreement	0.056
Simple past: No O Agreement	Simple Past: Syncretic and gender A Agreement	0.073
Simple past: No O Agreement	Simple past: Syncretic O Agreement	0.092
Simple Past: Syncretic and gender A Agreement	Simple past: Gender and full O Agreement	0.584
Simple Past: Syncretic and gender A Agreement	Simple past: Gender O Agreement	0.653
Simple Past: Syncretic and gender A Agreement	Simple past: No O Agreement	0.623
Simple Past: Syncretic and gender A Agreement	Simple past: Syncretic O Agreement	0.886
Simple past: Syncretic O Agreement	Simple past: Gender and full O Agreement	0.474
Simple past: Syncretic O Agreement	Simple past: Gender O Agreement	0.530
Simple past: Syncretic O Agreement	Simple past: No O Agreement	0.439
Simple past: Syncretic O Agreement	Simple Past: Syncretic and gender A Agreement	0.733

No Postpositions	Postpositions	0.034
Postpositions	No Postpositions	0.085
No Prepositions	Prepositions	0.060
Prepositions	No Prepositions	0.037
Clitic finite V: Category irrelevant	Clitic finite V: OV	0.076
Clitic finite V: Category irrelevant	Clitic finite V: V2	0.117
Clitic finite V: Category irrelevant	Clitic finite V: VO	0.543
Clitic finite V: Category irrelevant	Clitic finite V: VO and OV	0.097
Clitic finite V: OV	Clitic finite V: Category irrelevant	0.376
Clitic finite V: OV	Clitic finite V: V2	0.123
Clitic finite V: OV	Clitic finite V: VO	0.174
Clitic finite V: OV	Clitic finite V: VO and OV	0.130
Clitic finite V: V2	Clitic finite V: Category irrelevant	0.762
Clitic finite V: V2	Clitic finite V: OV	0.242
Clitic finite V: V2	Clitic finite V: VO	0.612
Clitic finite V: V2	Clitic finite V: VO and OV	0.431
Clitic finite V: VO	Clitic finite V: Category irrelevant	1.187
Clitic finite V: VO	Clitic finite V: OV	0.240
Clitic finite V: VO	Clitic finite V: V2	0.563
Clitic finite V: VO	Clitic finite V: VO and OV	0.364
Clitic finite V: VO and OV	Clitic finite V: Category irrelevant	0.833
Clitic finite V: VO and OV	Clitic finite V: OV	0.517
Clitic finite V: VO and OV	Clitic finite V: V2	0.836
Clitic finite V: VO and OV	Clitic finite V: VO	0.778
Clitic Infinitive V: Category irrelevant	Clitic Infinitive V: OV	0.096
Clitic Infinitive V: Category irrelevant	Clitic Infinitive V: V2	0.062
Clitic Infinitive V: Category irrelevant	Clitic Infinitive V: VO	0.097
Clitic Infinitive V: Category irrelevant	Clitic Infinitive V: VO and OV	0.081
Clitic Infinitive V: OV	Clitic Infinitive V: Category irrelevant	0.316
Clitic Infinitive V: OV	Clitic Infinitive V: V2	0.625
Clitic Infinitive V: OV	Clitic Infinitive V: VO	0.668
Clitic Infinitive V: OV	Clitic Infinitive V: VO and OV	0.388
Clitic Infinitive V: V2	Clitic Infinitive V: Category irrelevant	0.563
Clitic Infinitive V: V2	Clitic Infinitive V: OV	0.626
Clitic Infinitive V: V2	Clitic Infinitive V: VO	0.412
Clitic Infinitive V: V2	Clitic Infinitive V: VO and OV	0.426

Clitic Infinitive V: VO	Clitic Infinitive V: Category irrelevant	0.170
Clitic Infinitive V: VO	Clitic Infinitive V: OV	0.940
Clitic Infinitive V: VO	Clitic Infinitive V: V2	0.345
Clitic Infinitive V: VO	Clitic Infinitive V: VO and OV	0.272
Clitic Infinitive V: VO and OV	Clitic Infinitive V: Category irrelevant	1.434
Clitic Infinitive V: VO and OV	Clitic Infinitive V: OV	1.263
Clitic Infinitive V: VO and OV	Clitic Infinitive V: V2	1.191
Clitic Infinitive V: VO and OV	Clitic Infinitive V: VO	1.054
Clitic Participle V: Category irrelevant	Clitic Participle V: OV	0.139
Clitic Participle V: Category irrelevant	Clitic Participle V: V2	0.090
Clitic Participle V: Category irrelevant	Clitic Participle V: VO	0.116
Clitic Participle V: Category irrelevant	Clitic Participle V: VO and OV	0.101
Clitic Participle V: OV	Clitic Participle V: Category irrelevant	0.494
Clitic Participle V: OV	Clitic Participle V: V2	0.433
Clitic Participle V: OV	Clitic Participle V: VO	0.595
Clitic Participle V: OV	Clitic Participle V: VO and OV	0.361
Clitic Participle V: V2	Clitic Participle V: Category irrelevant	0.604
Clitic Participle V: V2	Clitic Participle V: OV	0.759
Clitic Participle V: V2	Clitic Participle V: VO	0.561
Clitic Participle V: V2	Clitic Participle V: VO and OV	0.475
Clitic Participle V: VO	Clitic Participle V: Category irrelevant	0.233
Clitic Participle V: VO	Clitic Participle V: OV	0.821
Clitic Participle V: VO	Clitic Participle V: V2	0.512
Clitic Participle V: VO	Clitic Participle V: VO and OV	0.408
Clitic Participle V: VO and OV	Clitic Participle V: Category irrelevant	0.864
Clitic Participle V: VO and OV	Clitic Participle V: OV	1.317
Clitic Participle V: VO and OV	Clitic Participle V: V2	1.111
Clitic Participle V: VO and OV	Clitic Participle V: VO	1.687
Infinitive: Category irrelevant	Infinitive: OV	0.545
Infinitive: Category irrelevant	Infinitive: OV and VO	0.532
Infinitive: Category irrelevant	Infinitive: VO	0.577
Infinitive: OV	Infinitive: Category irrelevant	0.038

Infinitive: OV	Infinitive: OV and VO	0.038
Infinitive: OV	Infinitive: VO	0.075
Infinitive: OV and VO	Infinitive: Category irrelevant	1.365
Infinitive: OV and VO	Infinitive: OV	0.567
Infinitive: OV and VO	Infinitive: VO	0.757
Infinitive: VO	Infinitive: Category irrelevant	0.106
Infinitive: VO	Infinitive: OV	0.080
Infinitive: VO	Infinitive: OV and VO	0.151
Main clause: SOV	Main clause: SOV/SVO	0.047
Main clause: SOV	Main clause: SVO	0.043
Main clause: SOV	Main clause: V2	0.035
Main clause: SOV	Main clause: VSO	0.031
Main clause: SOV/SVO	Main clause: SOV	1.082
Main clause: SOV/SVO	Main clause: SVO	1.547
Main clause: SOV/SVO	Main clause: V2	1.331
Main clause: SOV/SVO	Main clause: VSO	0.969
Main clause: SVO	Main clause: SOV	0.048
Main clause: SVO	Main clause: SOV/SVO	0.078
Main clause: SVO	Main clause: V2	0.197
Main clause: SVO	Main clause: VSO	0.057
Main clause: V2	Main clause: SOV	0.086
Main clause: V2	Main clause: SOV/SVO	0.133
Main clause: V2	Main clause: SVO	0.423
Main clause: V2	Main clause: VSO	0.145
Main clause: VSO	Main clause: SOV	0.419
Main clause: VSO	Main clause: SOV/SVO	0.451
Main clause: VSO	Main clause: SVO	0.540
Main clause: VSO	Main clause: V2	0.491
Noun - Possessor	Possessor - Noun	0.154
Noun - Possessor	Possessor - Noun and Noun - Possessor	0.159
Possessor - Noun	Noun - Possessor	0.088
Possessor - Noun	Possessor - Noun and Noun - Possessor	0.054
Possessor - Noun and Noun - Possessor	Noun - Possessor	0.929
Possessor - Noun and Noun - Possessor	Possessor - Noun	0.689
Adjective - Noun	Adjective - Noun and Noun - Adjective	0.086
Adjective - Noun	Noun - Adjective	0.024
Adjective - Noun and Noun - Adjective	Adjective - Noun	0.775

Adjective - Noun and Noun - Adjective	Noun - Adjective	1.035
Noun - Adjective	Adjective - Noun	0.079
Noun - Adjective	Adjective - Noun and Noun - Adjective	0.253
NRel: Category irrelevant	NRel: Noun - Relative	1.470
NRel: Category irrelevant	NRel: Relative - Noun	1.526
NRel: Category irrelevant	NRel: Relative - Noun and Noun - Relative	1.045
NRel: Noun - Relative	NRel: Category irrelevant	0.046
NRel: Noun - Relative	NRel: Relative - Noun	0.030
NRel: Noun - Relative	NRel: Relative - Noun and Noun - Relative	0.030
NRel: Relative - Noun	NRel: Category irrelevant	0.160
NRel: Relative - Noun	NRel: Noun - Relative	0.210
NRel: Relative - Noun	NRel: Relative - Noun and Noun - Relative	0.221
NRel: Relative - Noun and Noun - Relative	NRel: Category irrelevant	0.962
NRel: Relative - Noun and Noun - Relative	NRel: Noun - Relative	1.271
NRel: Relative - Noun and Noun - Relative	NRel: Relative - Noun	1.595
O - Participle	Participle - O	0.081
O - Participle	Participle - O and O - Participle	0.036
Participle - O	O - Participle	0.091
Participle - O	Participle - O and O - Participle	0.066
Participle - O and O - Participle	O - Participle	0.392
Participle - O and O - Participle	Participle - O	0.334
Subordinate clause: SOV	Subordinate clause: SOV and SVO	0.048
Subordinate clause: SOV	Subordinate clause: SVO	0.059
Subordinate clause: SOV	Subordinate clause: V2	0.054
Subordinate clause: SOV	Subordinate clause: VSO	0.028
Subordinate clause: SOV and SVO	Subordinate clause: SOV	1.322
Subordinate clause: SOV and SVO	Subordinate clause: SVO	1.398
Subordinate clause: SOV and SVO	Subordinate clause: V2	1.053
Subordinate clause: SOV and SVO	Subordinate clause: VSO	1.012
Subordinate clause: SVO	Subordinate clause: SOV	0.078
Subordinate clause: SVO	Subordinate clause: SOV and SVO	0.064
Subordinate clause: SVO	Subordinate clause: V2	0.053
Subordinate clause: SVO	Subordinate clause: VSO	0.051
Subordinate clause: V2	Subordinate clause: SOV	0.311

Subordinate clause: V2	Subordinate clause: SOV and SVO	0.407
Subordinate clause: V2	Subordinate clause: SVO	0.992
Subordinate clause: V2	Subordinate clause: VSO	0.309
Subordinate clause: VSO	Subordinate clause: SOV	0.303
Subordinate clause: VSO	Subordinate clause: SOV and SVO	0.247
Subordinate clause: VSO	Subordinate clause: SVO	0.199
Subordinate clause: VSO	Subordinate clause: V2	0.208
No WH - Verb (initial or not)	WH - Verb (initial or not)	0.267
WH - Verb (initial or not)	No WH - Verb (initial or not)	0.253
No WH-initial (no inversion)	WH-initial (no inversion)	0.213
WH-initial (no inversion)	No WH-initial (no inversion)	0.219

S6

Entry and exit (gain and loss) rates of features. Type = A (alignment), NM (nominal morphology), T (tense), VM (verbal morphology), WO (word order).

Type	Feature	Gain/entry rate	Loss/exit rat
A1	Present-Past: A marking in Present Progressive and Past	1.087	0.385
A2	Present-Past: Active-ergative in Present Progressive and Past	0.232	1.503
A3	Present-Past: All systems	0.333	3.032
A4	Present-Past: No marking difference	0.136	4.647
A5	Noun, Present progressive: No marking	0.642	1.219
A6	Noun, Present progressive: Nominative-accusative	1.112	0.675
A7	Noun, Present progressive: Tripartite	0.115	3.383
A8	Noun, Simple past: Ergative	0.231	1.209
A9	Noun, Simple past: No marking	0.692	1.582
A10	Noun, Simple past: Nominative-accusative	0.959	0.960
A11	Noun, Simple past: Tripartite	0.162	4.421
A12	Pronoun, Present progressive: No marking	0.118	0.807
A13	Pronoun, Present progressive: Nominative-accusative	0.807	0.118
A14	Pronoun, Simple past: Ergative	0.224	0.938
A15	Pronoun, Simple past: No marking	0.217	1.973
A16	Pronoun, Simple past: Nominative-accusative	0.527	0.259
A17	Pronoun, Simple past: Tripartite	0.154	3.964
A18	Reflexive not with Agent	0.042	0.888
A19	Reflexive with Agent	0.888	0.042
A20	Reflexive not with Object	0.548	0.035
A21	Reflexive with Object	0.035	0.548
A22	Verb, Present progressive: A=O/So=O	0.101	4.038
A23	Verb, Present progressive: No marking	0.195	2.874
A24	Verb, Present progressive: Nominative-Accusative	1.446	0.211

A25	Verb, Present progressive: Tripartite	0.120	2.912
A26	Verb, Simple past: Double oblique	0.183	5.646
A27	Verb, Simple past: Ergative	0.303	2.078
A28	Verb, Simple past: No marking	0.448	3.451
A29	Verb, Simple past: Nominative-accusative	0.981	0.562
A30	Verb, Simple past: Tripartite	0.298	2.998
NM1	Case on adjective	0.408	0.562
NM2	No case on adjective	0.562	0.408
NM3	Case on article	0.182	0.722
NM4	No case on article	0.722	0.182
NM5	Case on last member of NP	0.138	0.258
NM6	No rule of case on last/first member of NP	0.258	0.138
NM7	Case on noun	0.718	0.316
NM8	No case on noun	0.316	0.718
NM9	Definite suffix on adjective	0.061	0.274
NM10	No definite suffix on adjective	0.274	0.061
NM11	No obligatory definiteness	0.431	0.275
NM12	Obligatory definiteness on first member of NP	0.149	0.318
NM13	Obligatory definiteness on last member of NP	0.165	2.191
NM14	Definite article	0.102	0.135
NM15	No definite article	0.135	0.102
NM16	Definite suffix on noun	0.124	0.308
NM17	No definite suffix on noun	0.308	0.124
NM18	Fewer than five genders	0.650	0.017
NM19	More than five genders	0.017	0.650
NM20	No noun class for animates	0.508	0.070
NM21	Noun class for animates	0.070	0.508
NM22	Masculine/feminine distinction	0.373	0.214
NM23	No masculine/feminine distinction	0.214	0.373
NM24	Neuter gender	0.094	0.361
NM25	No neuter gender	0.361	0.094
NM26	Gender on predicative adjective	0.395	0.311
NM27	No gender on predicative adjective	0.311	0.395
NM28	More than 7 cases	0.057	0.728
NM29	Not more than 7 cases	0.728	0.057
NM30	Agglutination for number	0.187	0.353
NM31	No agglutination for number	0.353	0.187
NM32	Agglutination for case	0.218	0.520
NM33	No agglutination for case	0.520	0.218
NM34	Dative but no genitive	0.225	3.409
NM35	Genitive and dative	0.440	0.673
NM36	Genitive but no dative	0.358	2.247

NM37	No genitive or dative	0.367	0.540
NM38	Case difference A and O	0.583	0.355
NM39	No case difference A and O	0.355	0.583
NM40	No peripheral cases	0.150	0.046
NM41	Peripheral cases	0.046	0.150
NM42	No Vocative	0.253	0.128
NM43	Vocative	0.128	0.253
NM44	Agreement on prepositions	0.039	0.110
NM45	No agreement on prepositions	0.110	0.039
NM46	Fewer than 7 pronominal cases (pronouns)	0.740	0.038
NM47	More than 7 pronominal cases (pronouns)	0.038	0.740
NM48	Agglutination for number (pronouns)	0.132	0.324
NM49	No agglutination for number (pronouns)	0.324	0.132
NM50	Agglutination for case (pronouns)	0.434	0.058
NM51	No agglutination for case (pronouns)	0.058	0.434
NM52	Difference A and O (pronouns)	0.396	0.085
NM53	No difference A and O (pronouns)	0.085	0.396
NM54	Difference, O and Dative (pronouns)	0.381	0.535
NM55	No difference O and Dative (pronouns)	0.535	0.381
NM56	No peripheral cases (pronouns)	0.236	0.056
NM57	Peripheral cases (pronouns)	0.056	0.236
NM58	No Vocative (pronouns)	0.721	0.019
NM59	Vocative (pronouns)	0.019	0.721
T1	No synthetic Present progressive	0.071	0.522
T2	Synthetic Present progressive	0.522	0.071
T3	No Present progressive by auxiliary	0.440	0.270
T4	Present progressive by auxiliary	0.270	0.440
T5	Future by auxiliary	0.383	0.463
T6	No Future by auxiliary	0.463	0.383
T7	Future by participle	0.093	0.653
T8	No Future by participle	0.653	0.093
T9	Future by particle	0.132	0.581
T10	No Future by particle	0.581	0.132
T11	No synthetic Future	0.206	0.165
T12	Synthetic Future	0.165	0.206
T13	Future by perfect	0.023	0.360
T14	No Future by perfect	0.360	0.023
VM1	Present progressive: Full A Agreement	0.408	0.494
VM2	Present progressive: Full and Gender A Agreement	0.159	5.784
VM3	Present progressive: Gender A Agreement	0.211	4.530
VM4	Present progressive: No A Agreement	0.293	3.796
VM5	Present progressive: Syncretic A Agreement	0.491	0.724

VM6	Present progressive: Full Dative Agreement	0.068	2.770
VM7	Present progressive: No Dative Agreement	0.842	0.102
VM8	Present progressive: Syncretic Dative Agreement	0.103	1.127
VM9	Present progressive: Full O Agreement	0.068	2.764
VM10	Present progressive: No O Agreement	0.833	0.101
VM11	Present progressive: Syncretic O Agreement	0.102	1.116
VM12	Simple past: Full A Agreement	0.247	7.639
VM13	Simple past: Full and Gender A Agreement	0.766	2.130
VM14	Simple past: Gender A Agreement	0.308	5.932
VM15	Simple past: No A Agreement	0.352	6.085
VM16	Simple past: Syncretic A Agreement	0.784	2.993
VM17	Simple past: Full Dative Agreement	0.091	3.825
VM18	Simple past: Gender Dative Agreement	0.092	3.874
VM19	Simple past: No Dative Agreement	0.978	0.138
VM20	Simple past: Syncretic Dative Agreement	0.134	1.607
VM21	Simple past: Gender and full O Agreement	0.194	4.912
VM22	Simple past: Gender O Agreement	0.211	4.485
VM23	Simple past: No O Agreement	0.673	0.273
VM24	Simple Past: Syncretic and gender A Agreement	0.270	2.748
VM25	Simple past: Syncretic O Agreement	0.287	2.177
VM26	No Postpositions	0.085	0.034
WO1	Postpositions	0.034	0.085
WO2	No Prepositions	0.037	0.060
WO3	Prepositions	0.060	0.037
WO4	Clitic finite V: Category irrelevant	0.786	0.835
WO5	Clitic finite V: OV	0.165	0.805
WO6	Clitic finite V: V2	0.254	2.048
WO7	Clitic finite V: VO	0.493	2.355
WO8	Clitic finite V: VO and OV	0.192	2.965
WO9	Clitic Infinitive V: Category irrelevant	0.429	0.338
WO10	Clitic Infinitive V: OV	0.358	1.999
WO11	Clitic Infinitive V: V2	0.252	2.029
WO12	Clitic Infinitive V: VO	0.281	1.729
WO13	Clitic Infinitive V: VO and OV	0.197	4.943
WO14	Clitic Participle V: Category irrelevant	0.471	0.447
WO15	Clitic Participle V: OV	0.414	1.884
WO16	Clitic Participle V: V2	0.285	2.400
WO17	Clitic Participle V: VO	0.360	1.975
WO18	Clitic Participle V: VO and OV	0.241	4.980
WO19	Infinitive: Category irrelevant	0.119	1.655
WO20	Infinitive: OV	0.196	0.152
WO21	Infinitive: OV and VO	0.112	2.691

WO22	Infinitive: VO	0.166	0.339
WO23	Main clause: SOV	0.115	0.158
WO24	Main clause: SOV/SVO	0.089	4.931
WO25	Main clause: SVO	0.211	0.381
WO26	Main clause: V2	0.153	0.788
WO27	Main clause: VSO	0.077	1.902
WO28	Noun - Possessor	0.154	0.313
WO29	Possessor - Noun	0.228	0.142
WO30	Possessor - Noun and Noun - Possessor	0.091	1.618
WO31	Adjective - Noun	0.223	0.110
WO32	Adjective - Noun and Noun - Adjective	0.133	1.811
WO33	Noun - Adjective	0.118	0.333
WO34	NRel: Category irrelevant	0.082	4.043
WO35	NRel: Noun - Relative	0.458	0.107
WO36	NRel: Relative - Noun	0.103	0.592
WO37	NRel: Relative - Noun and Noun - Relative	0.080	3.829
WO38	O - Participle	0.132	0.118
WO39	Participle - O	0.109	0.158
WO40	Participle - O and O - Participle	0.049	0.726
WO41	Subordinate clause: SOV	0.161	0.191
WO42	Subordinate clause: SOV and SVO	0.082	4.786
WO43	Subordinate clause: SVO	0.178	0.246
WO44	Subordinate clause: V2	0.082	2.020
WO45	Subordinate clause: VSO	0.068	0.958
WO46	No WH - Verb (initial or not)	0.253	0.267
WO47	WH - Verb (initial or not)	0.267	0.253
WO48	No WH-initial (no inversion)	0.219	0.213
WO49	WH-initial (no inversion)	0.213	0.219

S7

List of pairwise organized features, which are identified to be in a relation by a marking hierarchy. Feature 1 = Trait higher in hierarchy (more frequent, unmarked); Feature 2 = trait lower in hierarchy (less frequent, marked); Type = type of property distinguishing hierarchical relation between traits.

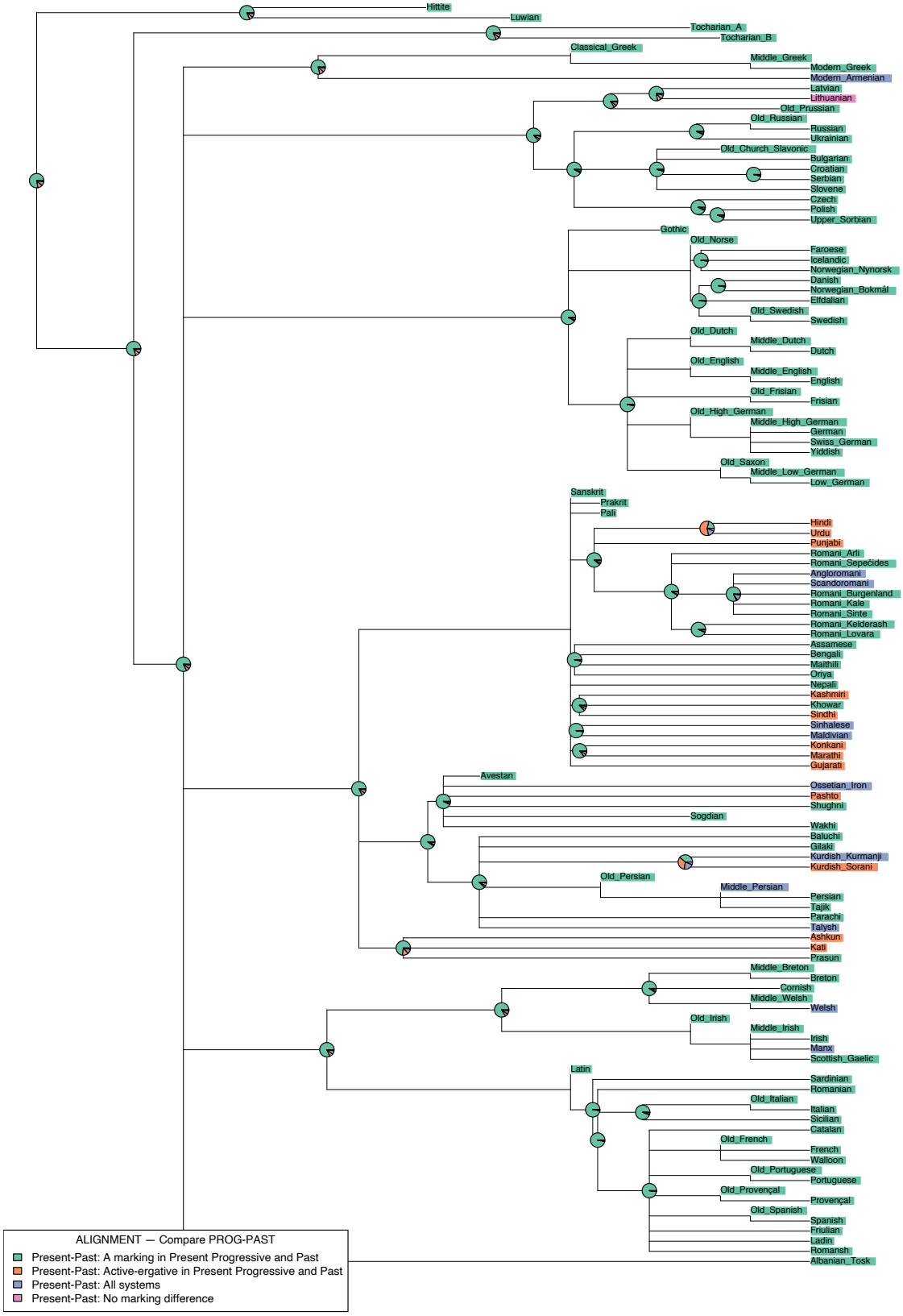
Feature 1 (Unmarked)	ID 1	Feature 2 (Marked)	ID 2	Type
Pronoun, Present progressive: Nominative-accusative	A13	Noun, Present progressive: Nominative-accusative	A6	pronoun < noun
Pronoun, Present progressive: No marking	A14	Noun, Present progressive: No marking	A7	pronoun < noun

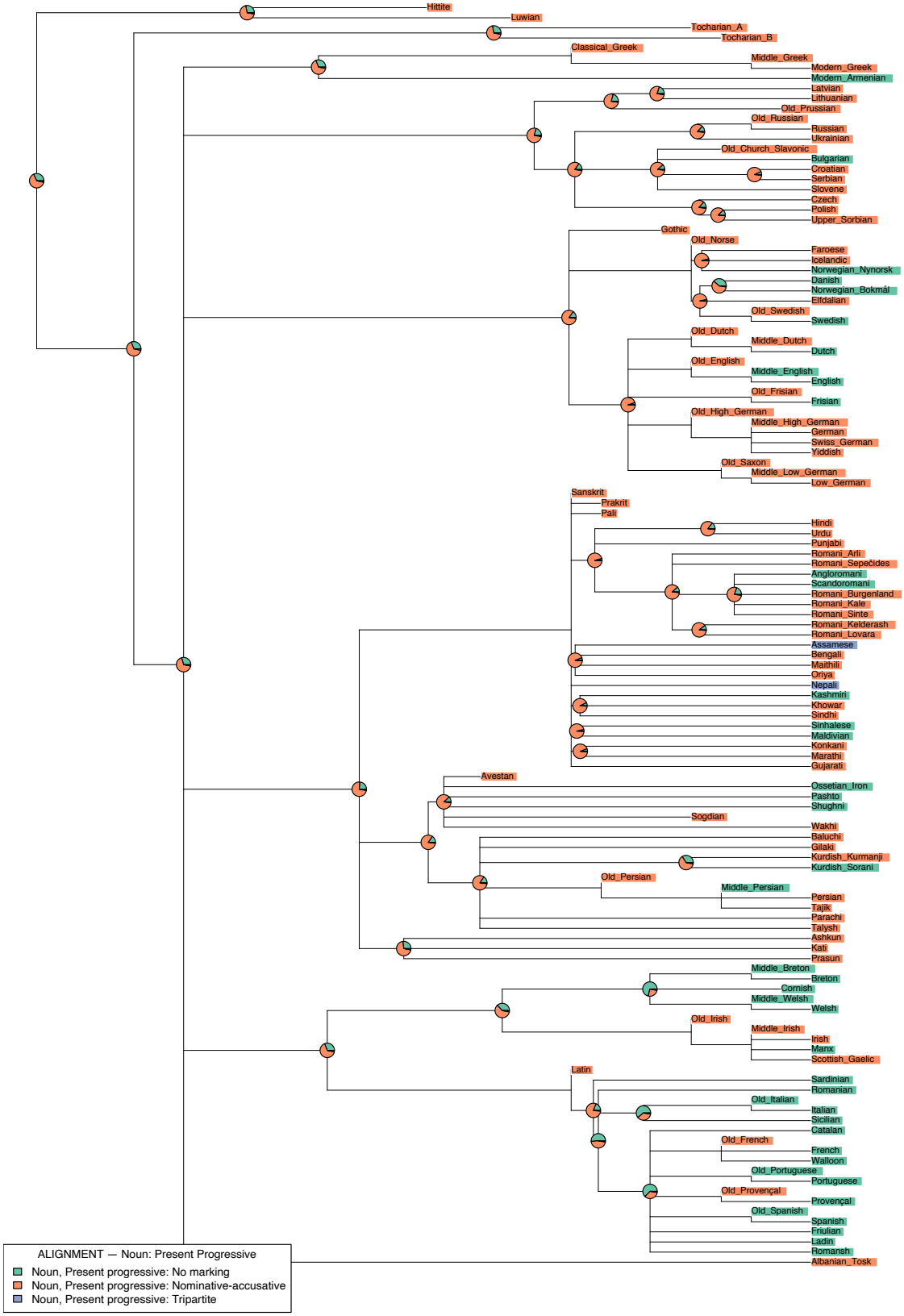
Pronoun, Simple past: Nominative-accusative	A16	Noun, Simple past: Nominative-accusative	A9	pronoun < noun
Pronoun, Simple past: No marking	A19	Noun, Simple past: No marking	A11	pronoun < noun
Pronoun, Simple past: Ergative	A18	Noun, Simple past: Ergative	A10	pronoun < noun
Pronoun, Simple past: Tripartite	A15	Noun, Simple past: Tripartite	A8	pronoun < noun
Agglutination for case (pronouns)	NM51	Agglutination for case	NM33	pronoun < noun
Agglutination for number (pronouns)	NM50	Agglutination for number	NM31	pronoun < noun
No agglutination for case (pronouns)	NM52	No agglutination for case	NM32	pronoun < noun
No agglutination for number (pronouns)	NM49	No agglutination for number	NM30	pronoun < noun
Difference A and O (pronouns)	NM54	Case difference A and O	NM40	pronoun < noun
No peripheral cases (pronouns)	NM57	Peripheral cases	NM42	pronoun < noun
Difference, O and Dative (pronouns)	NM56	No peripheral cases	NM41	pronoun < noun
Vocative (pronouns)	NM60	Vocative	NM44	pronoun < noun
No Vocative (pronouns)	NM59	No Vocative	NM43	pronoun < noun
More than 7 pronominal cases	NM48	More than 7 cases	NM29	pronoun < noun
Less than 7 pronominal cases	NM47	Not more than 7 cases	NM28	pronoun < noun
Synthetic Present progressive	T2	Synthetic Future	T12	present < future
No synthetic Present progressive	T1	No synthetic Future	T11	present < future
Present progressive by auxiliary	T4	Future by auxiliary	T6	present < future
No Present progressive by auxiliary	T3	No Future by auxiliary	T5	present < future
Pronoun, Present progressive: Nominative-accusative	A13	Pronoun, Simple past: Nominative-accusative	A16	present < past
Pronoun, Present progressive: No marking	A14	Pronoun, Simple past: No marking	A19	present < past
Noun, Present progressive: Nominative-accusative	A6	Noun, Simple past: Nominative-accusative	A9	present < past

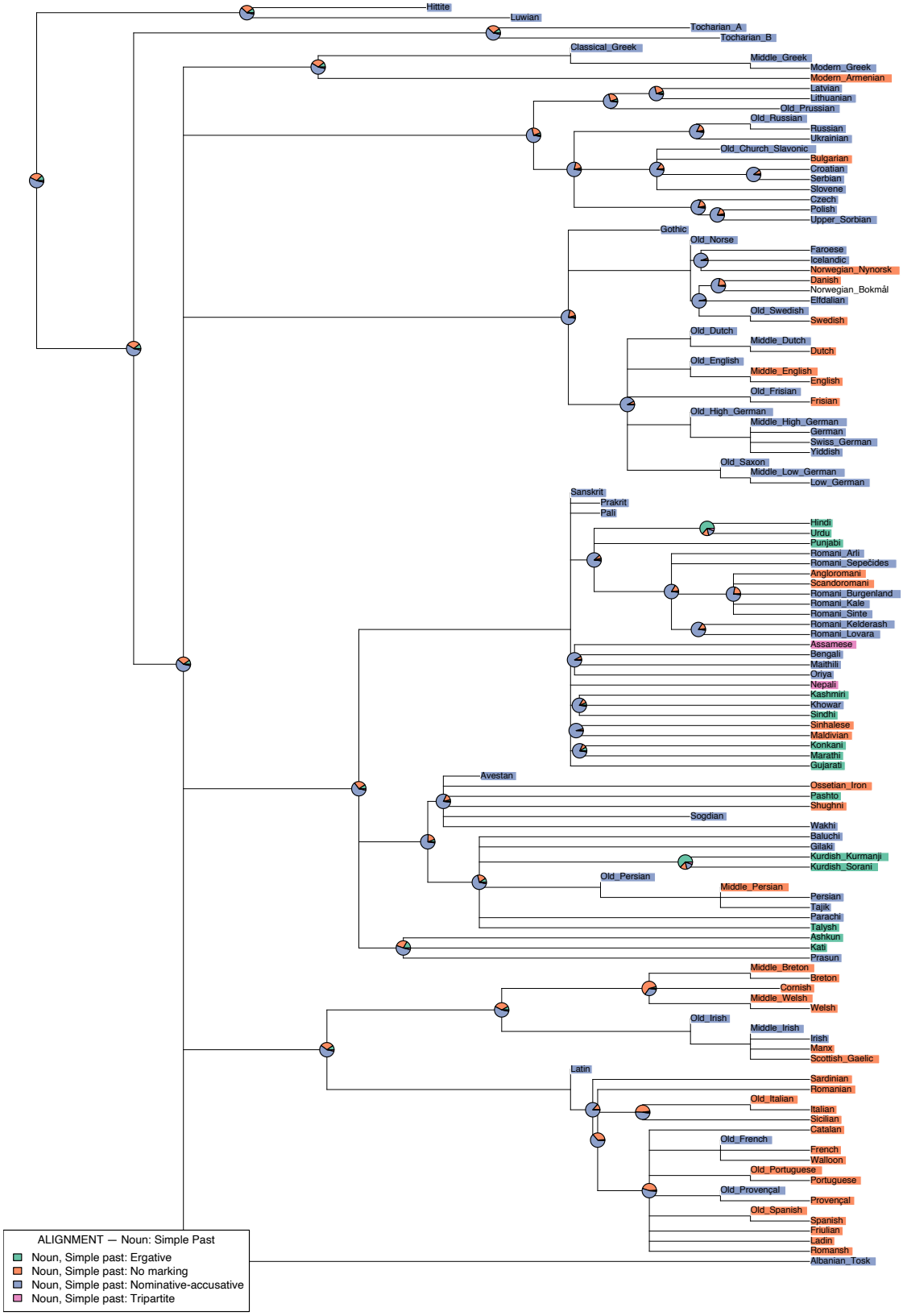
Noun, Present progressive: No marking	A7	Noun, Simple past: No marking	A11	present < past
Noun, Present progressive: Tripartite	A5	Noun, Simple past: Tripartite	A8	present < past
Verb, Present progressive: Nominative-Accusative	A25	Verb, Simple past: Nominative-accusative	A28	present < past
Verb, Present progressive: No marking	A26	Verb, Simple past: No marking	A30	present < past
Verb, Present progressive: Tripartite	A24	Verb, Simple past: Tripartite	A27	present < past
Present progressive: Syncretic A Agreement	VM1	Simple past: Syncretic A Agreement	VM12	present < past
Present progressive: No A Agreement	VM2	Simple past: No A Agreement	VM13	present < past
Present progressive: Gender A Agreement	VM3	Simple past: Gender A Agreement	VM14	present < past
Present progressive: Full A Agreement	VM4	Simple past: Full A Agreement	VM16	present < past
Present progressive: Full and Gender A Agreement	VM5	Simple past: Full and Gender A Agreement	VM17	present < past
Present progressive: Syncretic Dative Agreement	VM6	Simple past: Syncretic Dative Agreement	VM18	present < past
Present progressive: No Dative Agreement	VM7	Simple past: No Dative Agreement	VM19	present < past
Present progressive: Full Dative Agreement	VM8	Simple past: Full Dative Agreement	VM21	present < past
Present progressive: Syncretic O Agreement	VM9	Simple past: Syncretic O Agreement	VM22	present < past
Present progressive: No O Agreement	VM11	Simple past: No O Agreement	VM25	present < past
Reflexive with Agent	A21	Reflexive with Object	A23	agent < object
Reflexive not with Agent	A20	Reflexive not with Object	A22	agent < object
Case difference A and O	NM40	Genitive and dative	NM38	agent/object < oblique
Difference A and O (pronouns)	NM54	Difference, O and Dative (pronouns)	NM56	agent/object < oblique
No difference A and O (pronouns)	NM53	No difference O and Dative (pronouns)	NM55	agent/object < oblique
Masculine/feminine distinction	NM23	Neuter gender	NM25	masculine/feminine < neuter
No masculine/feminine distinction	NM22	No neuter gender	NM24	masculine/feminine < neuter

S8

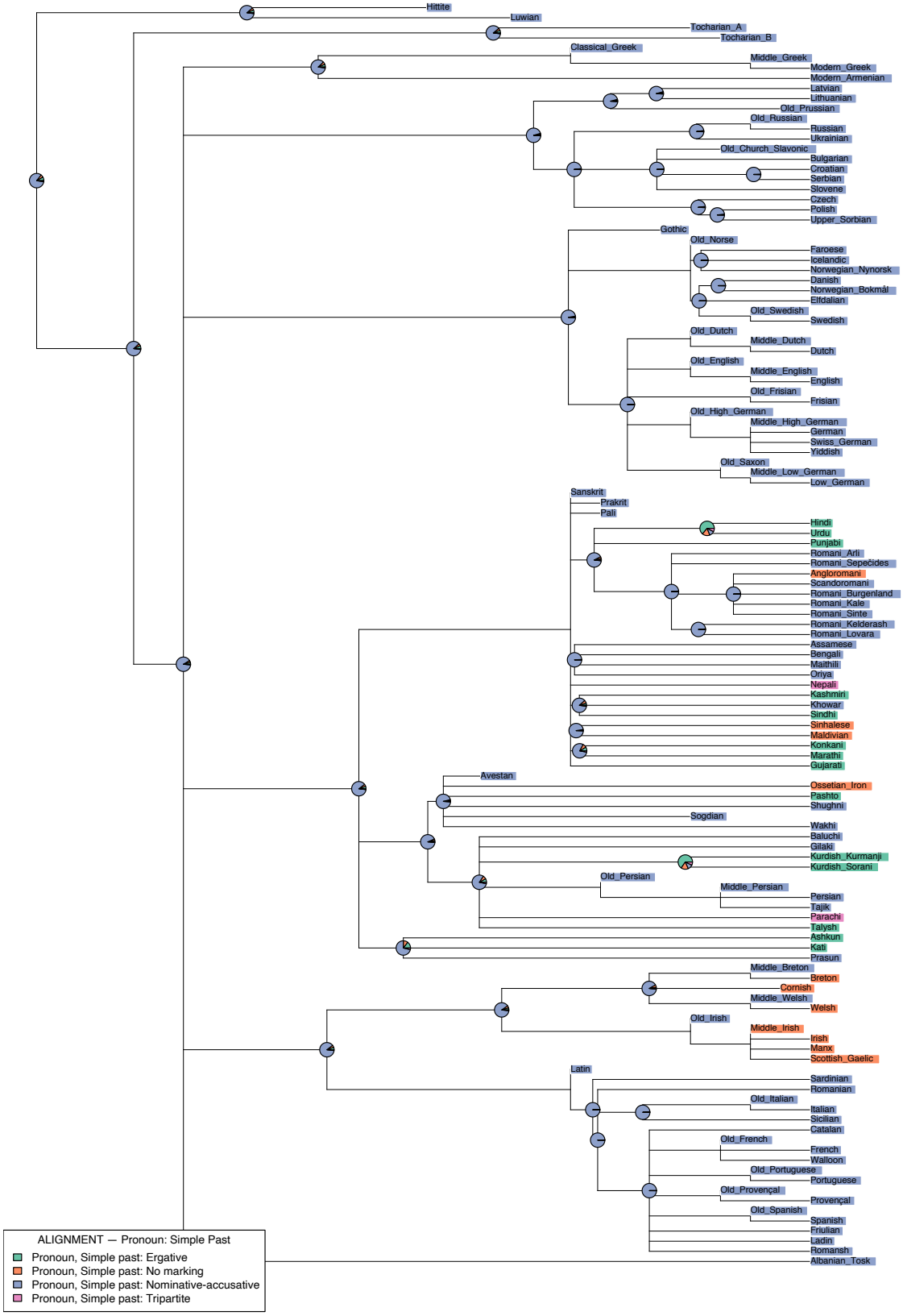
Reconstructed probability distributions of variables in our data set over all nodes of the phylogeny, visualized on a maximum clade credibility (MCC) tree of Indo-European.


Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
ALIGNMENT — Compare PROG-PAST
Present-Past: A marking in Present Progressive and Past
Present-Past: Active-ergative in Present Progressive and Past
Present-Past: All systems
Present-Past: No marking difference

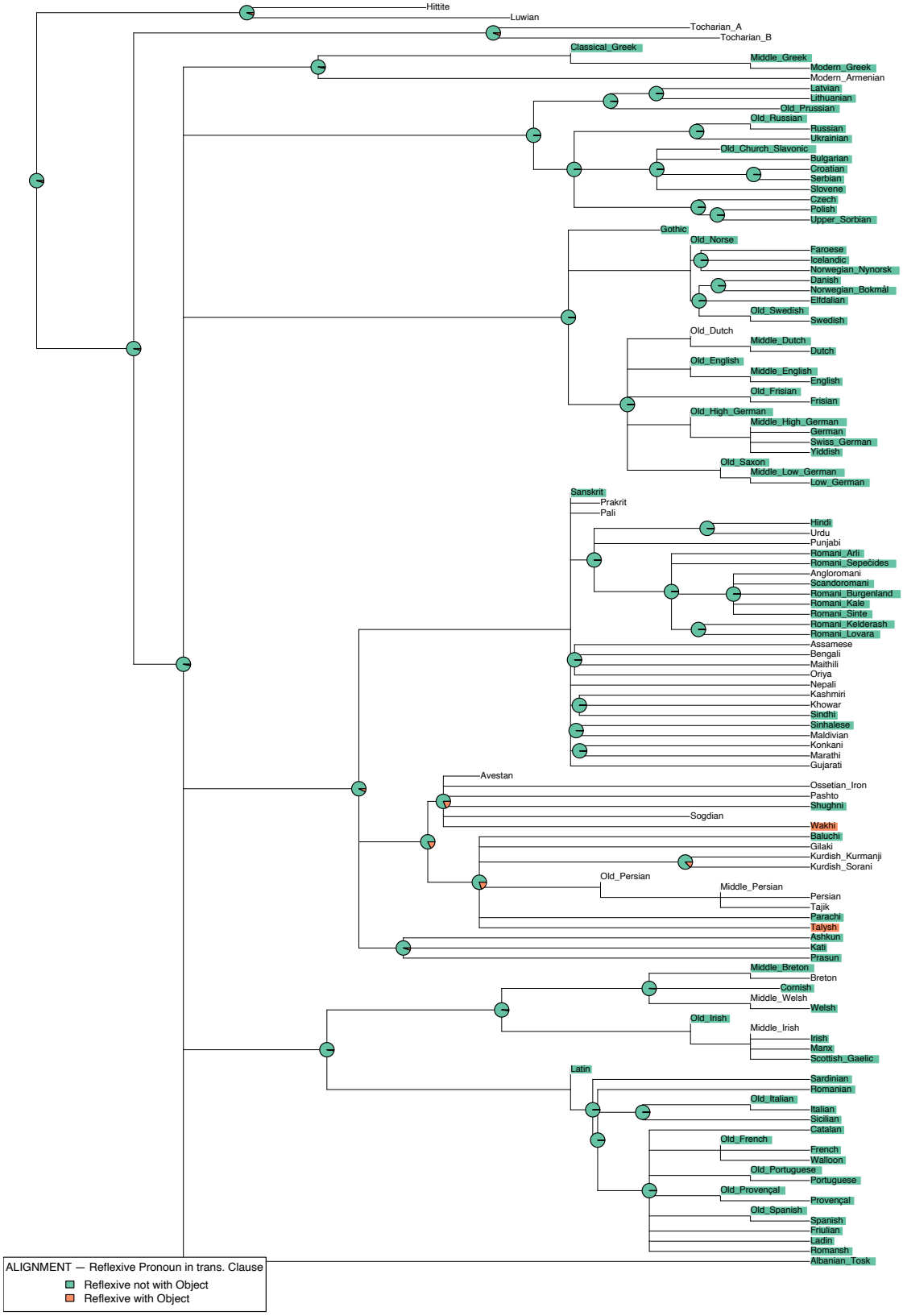

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
ALIGNMENT — Noun: Present Progressive
Noun, Present progressive: No marking
Noun, Present progressive: Nominative-accusative
Noun, Present progressive: Tripartite

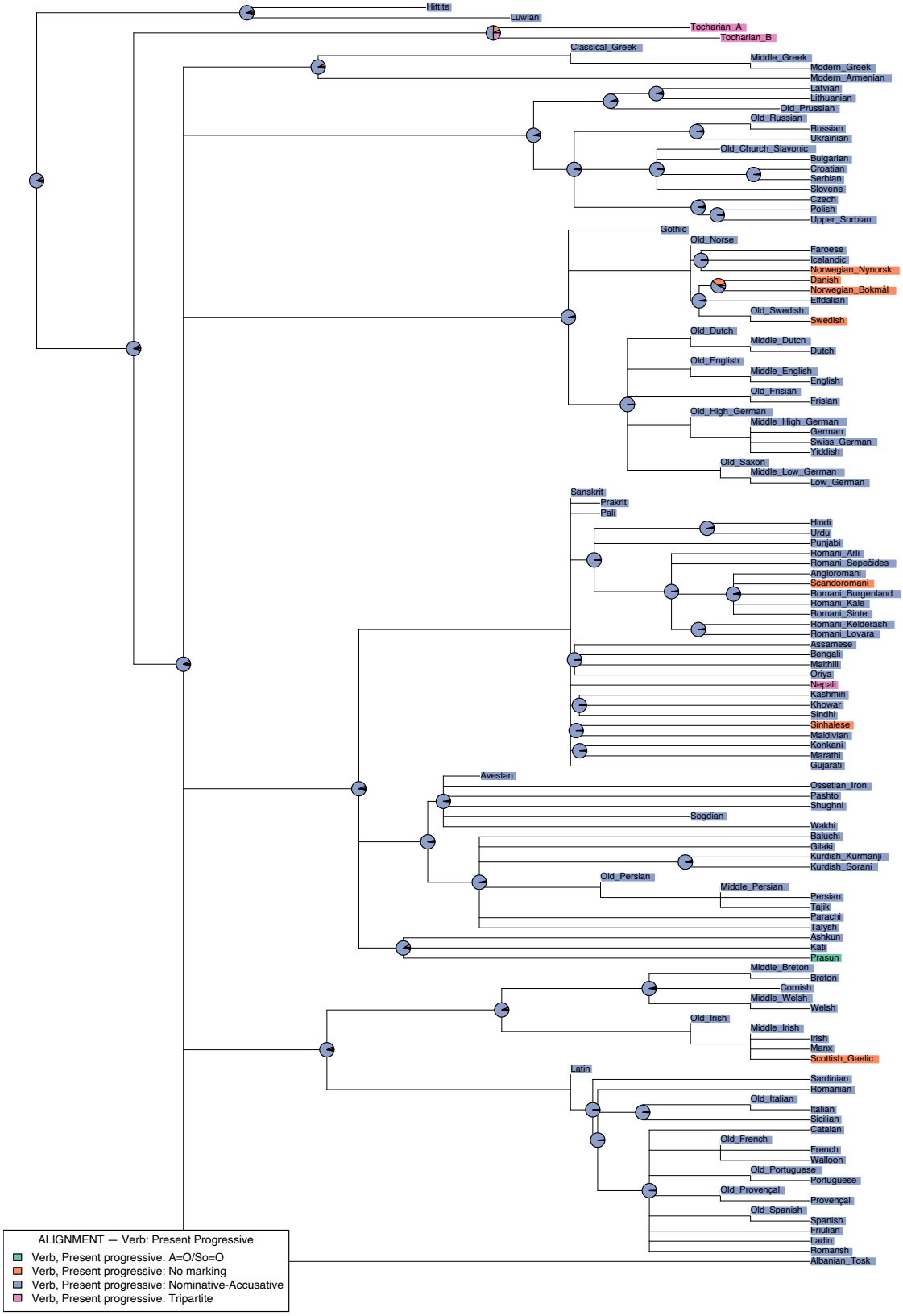

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
ALIGNMENT — Noun: Simple Past
Noun, Simple past: Ergative
Noun, Simple past: No marking
Noun, Simple past: Nominative-accusative
Noun, Simple past: Tripartite

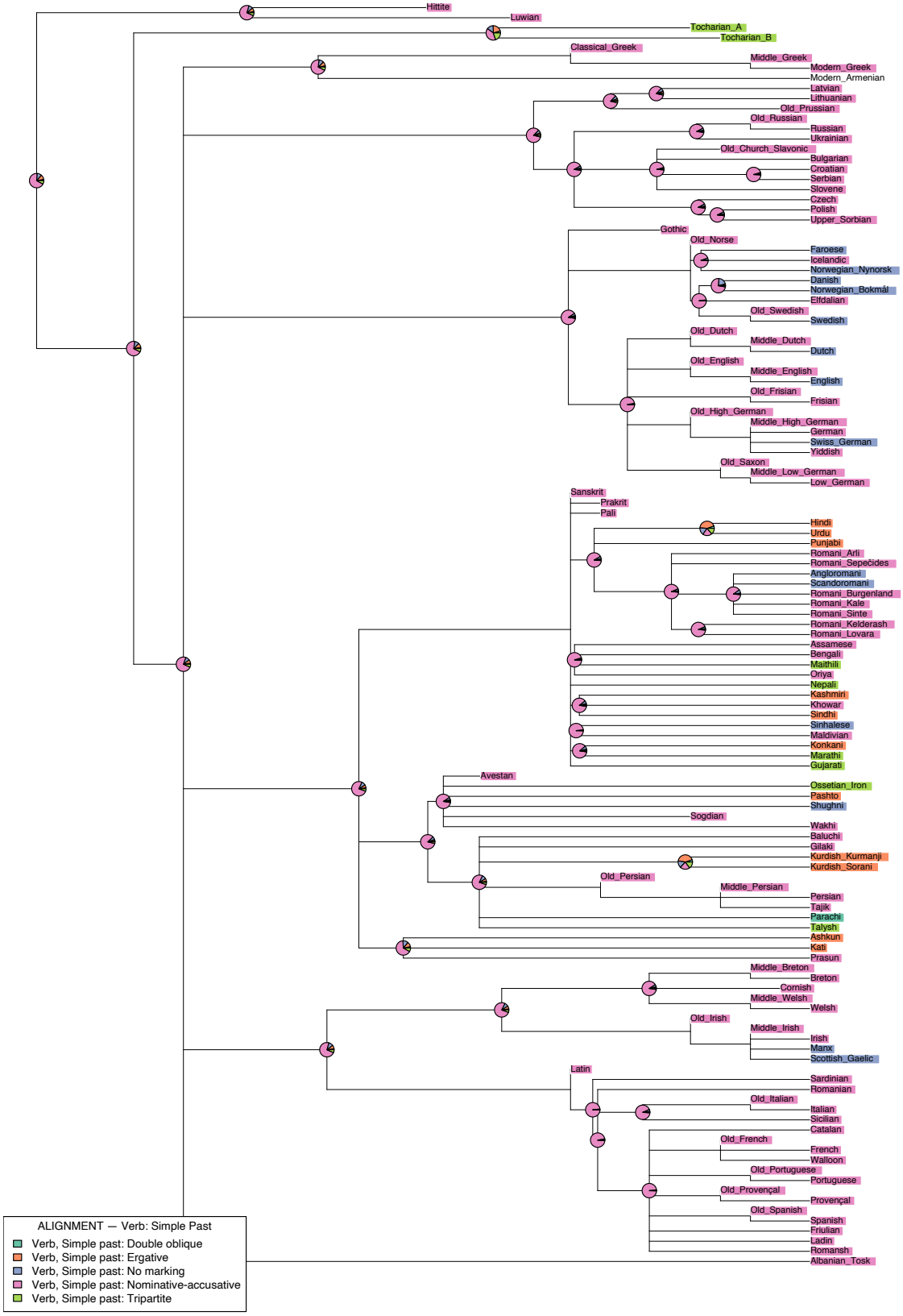

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
ALIGNMENT — Pronoun: Present Progressive
Pronoun, Present progressive: No marking
Pronoun, Present progressive: Nominative-accusative

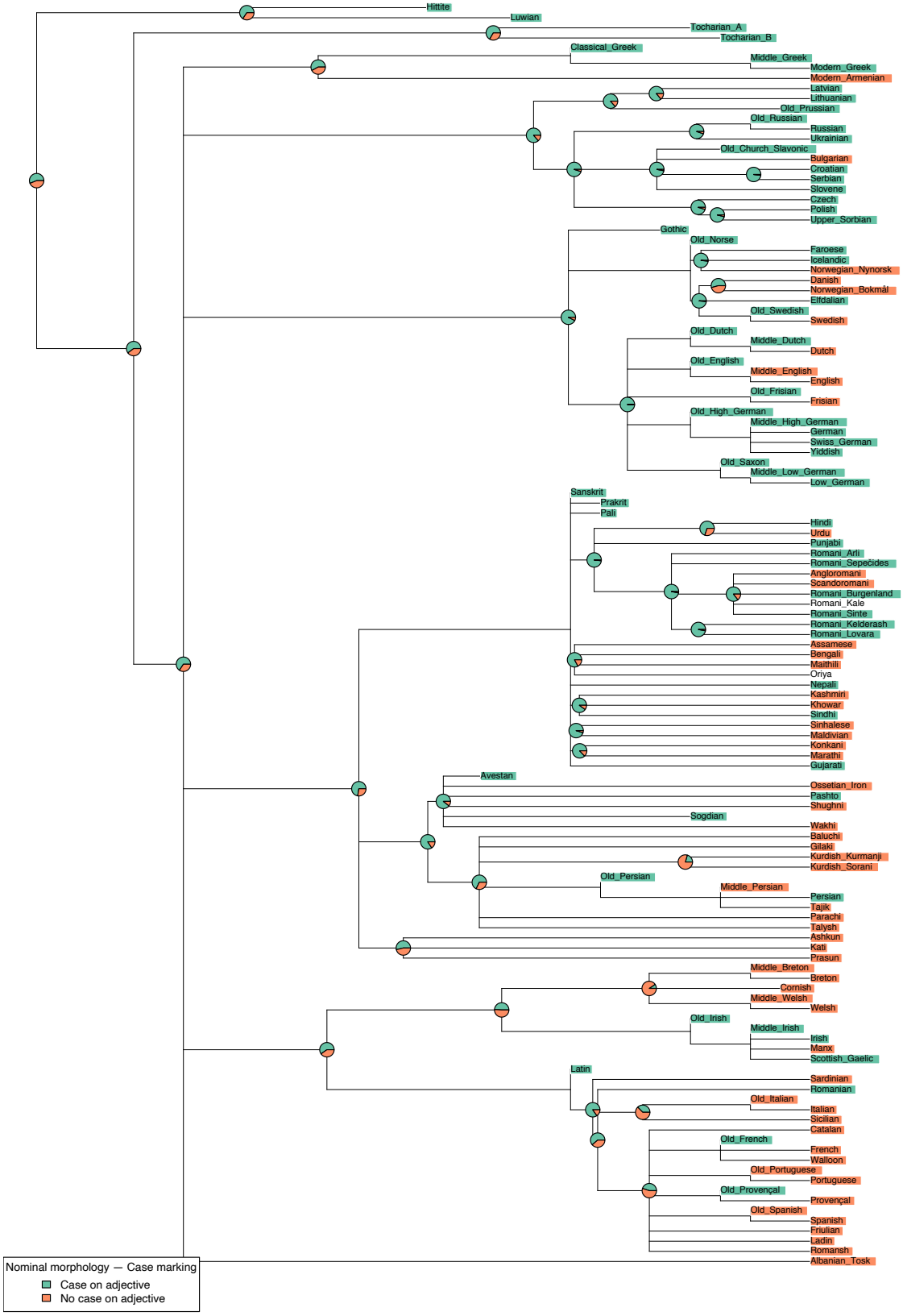

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
ALIGNMENT — Pronoun: Simple Past
Pronoun, Simple past: Ergative
Pronoun, Simple past: No marking
Pronoun, Simple past: Nominative-accusative
Pronoun, Simple past: Tripartite

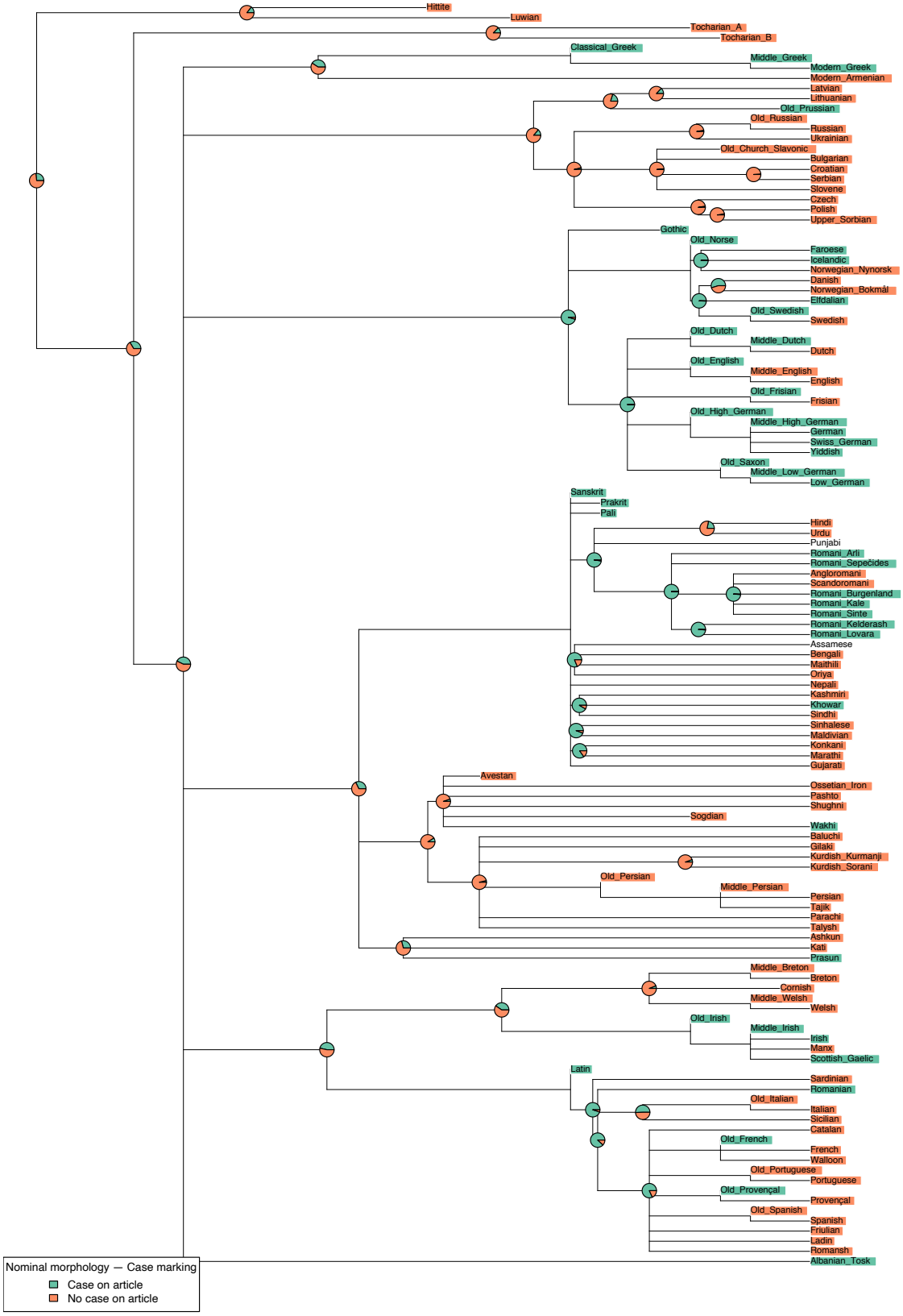

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
ALIGNMENT — Reflexive Pronoun in trans. Clause
Reflexive not with Agent
Reflexive with Agent

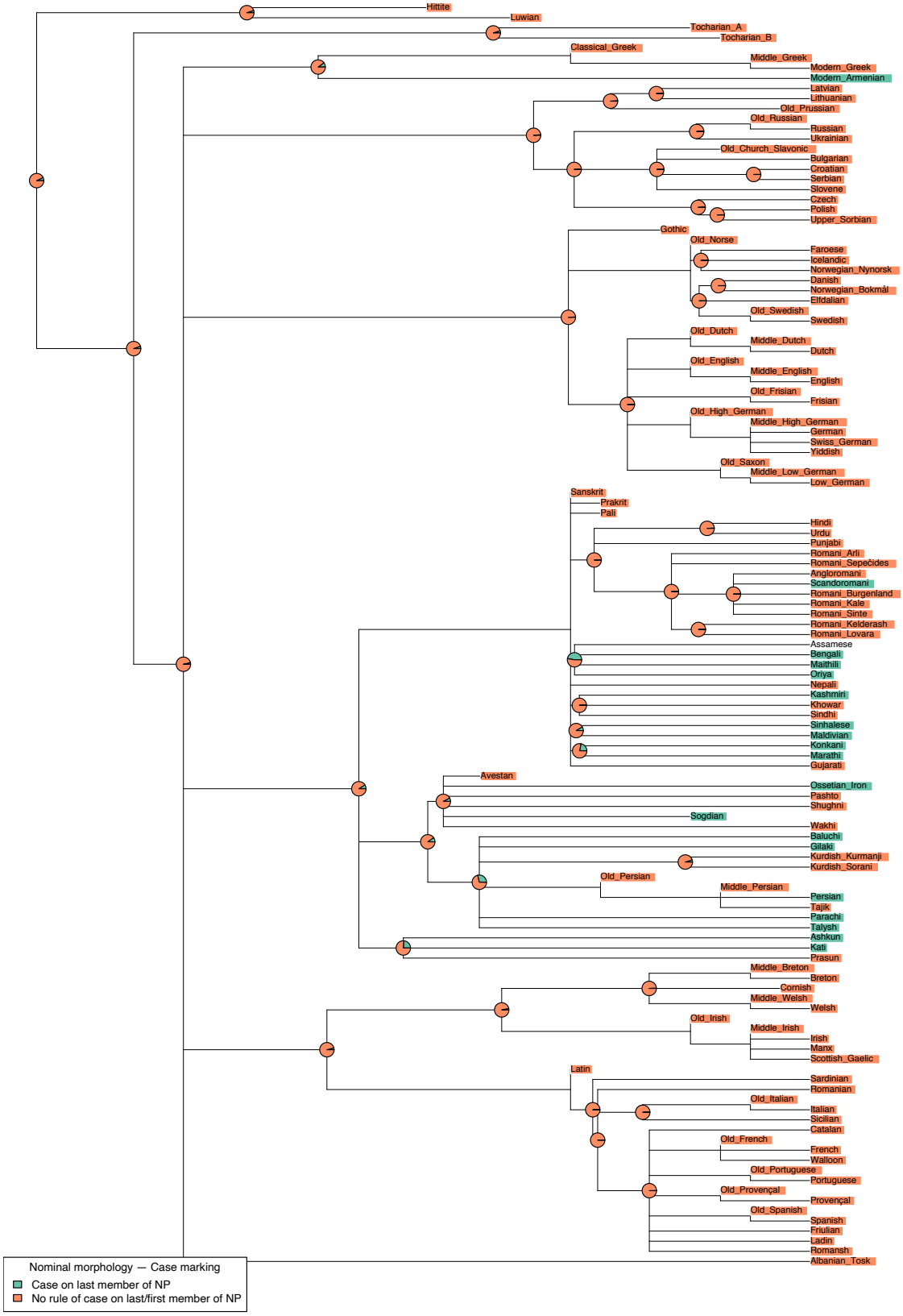

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
ALIGNMENT — Reflexive Pronoun in trans. Clause
Reflexive not with Object
Reflexive with Object

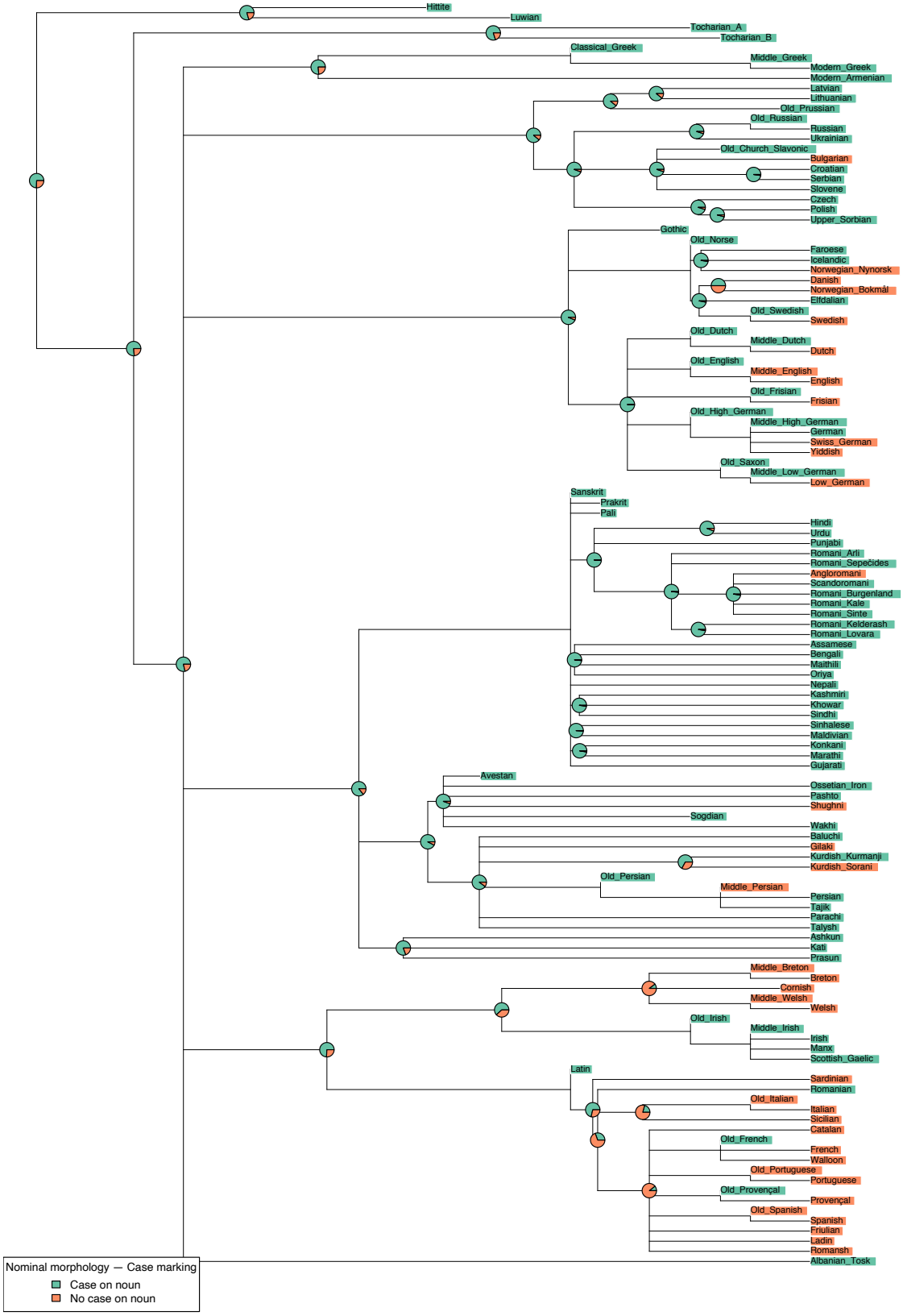

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
ALIGNMENT — Verb: Present Progressive
Verb, Present progressive: A=O/So=O
Verb, Present progressive: No marking
Verb, Present progressive: Nominative-Accusative
Verb, Present progressive: Tripartite

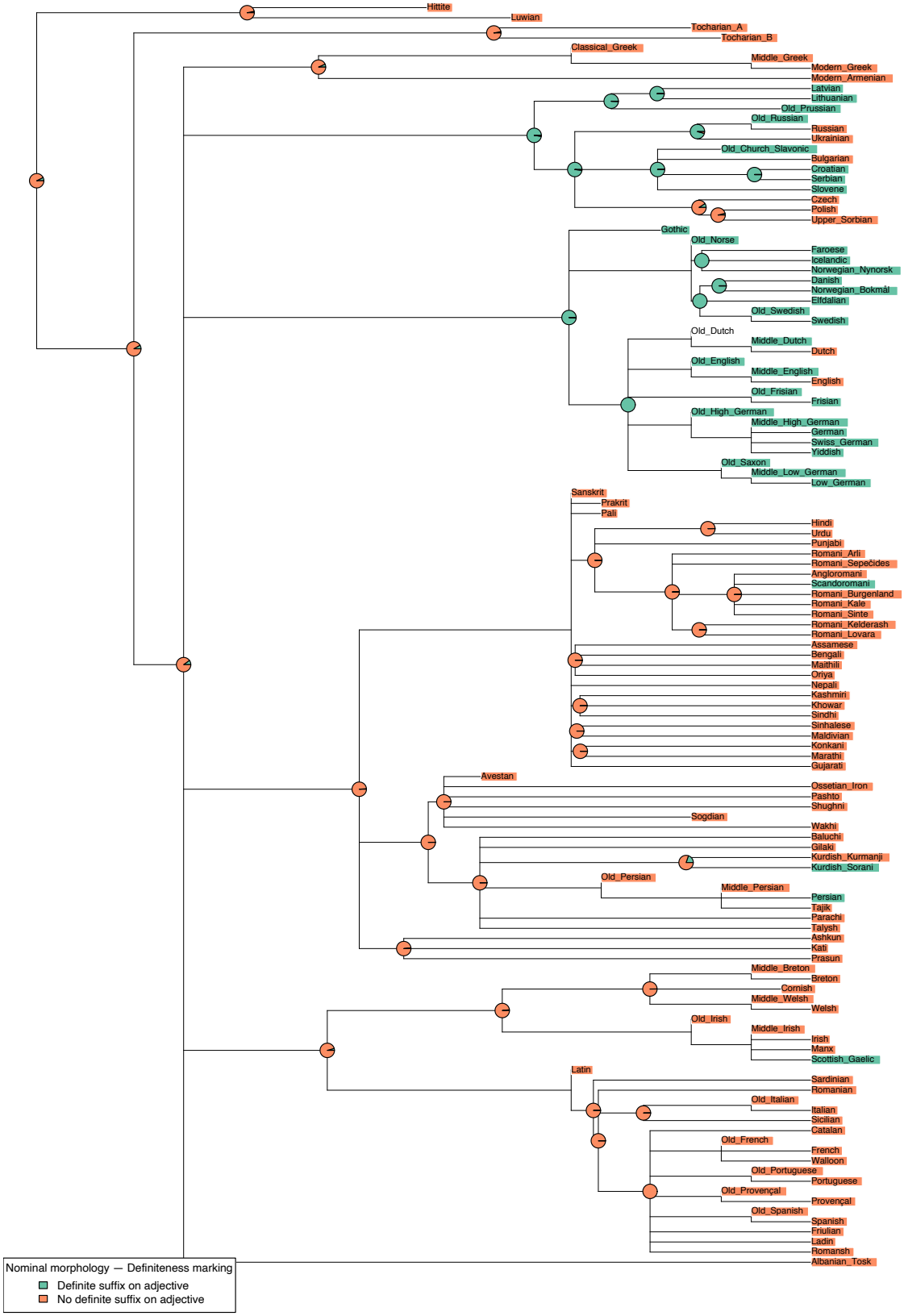

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
ALIGNMENT — Verb: Simple Past
Verb, Simple past: Double oblique
Verb, Simple past: Ergative
Verb, Simple past: No marking
Verb, Simple past: Nominative-accusative
Verb, Simple past: Tripartite

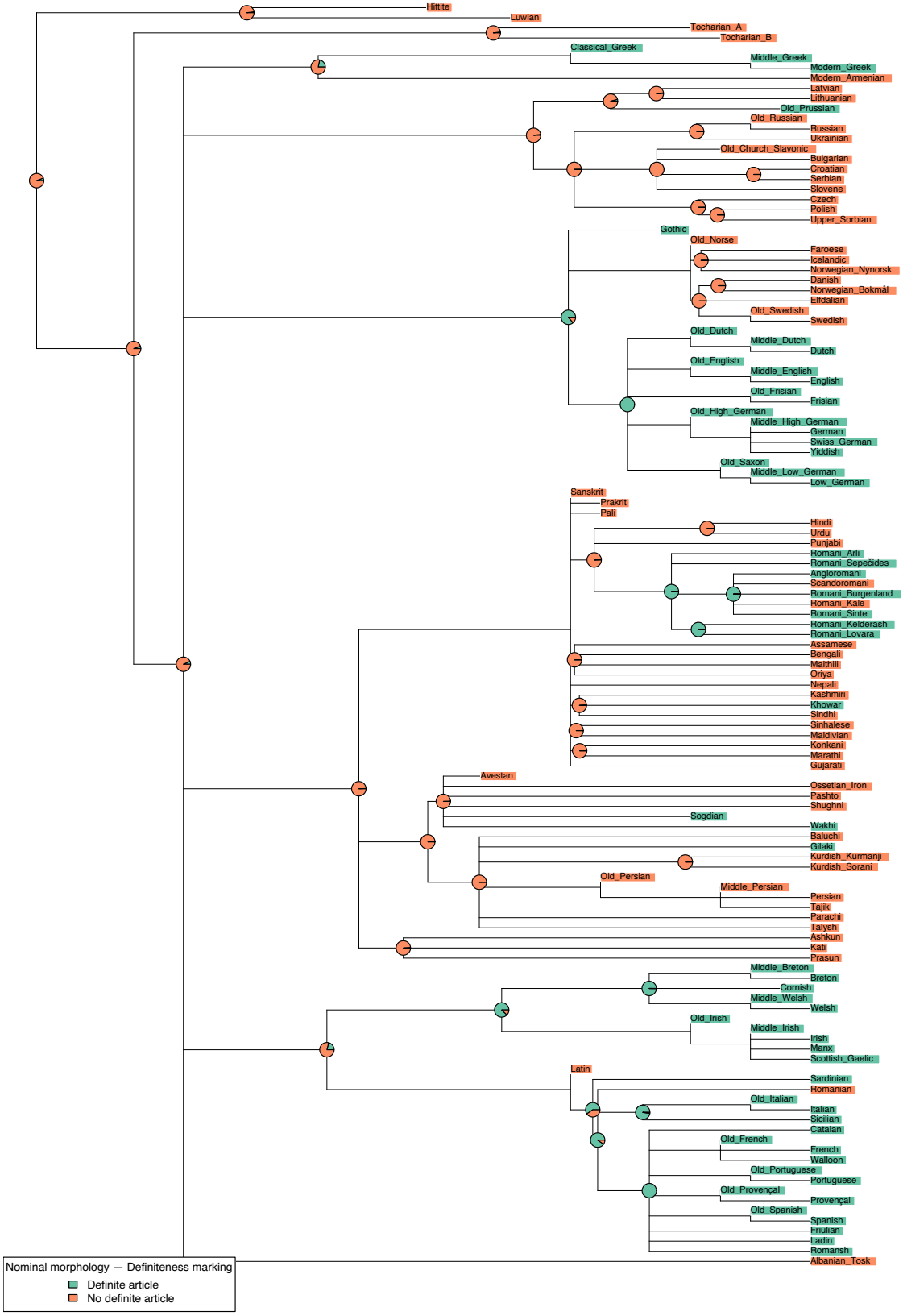

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Case marking
Case on adjective
No case on adjective

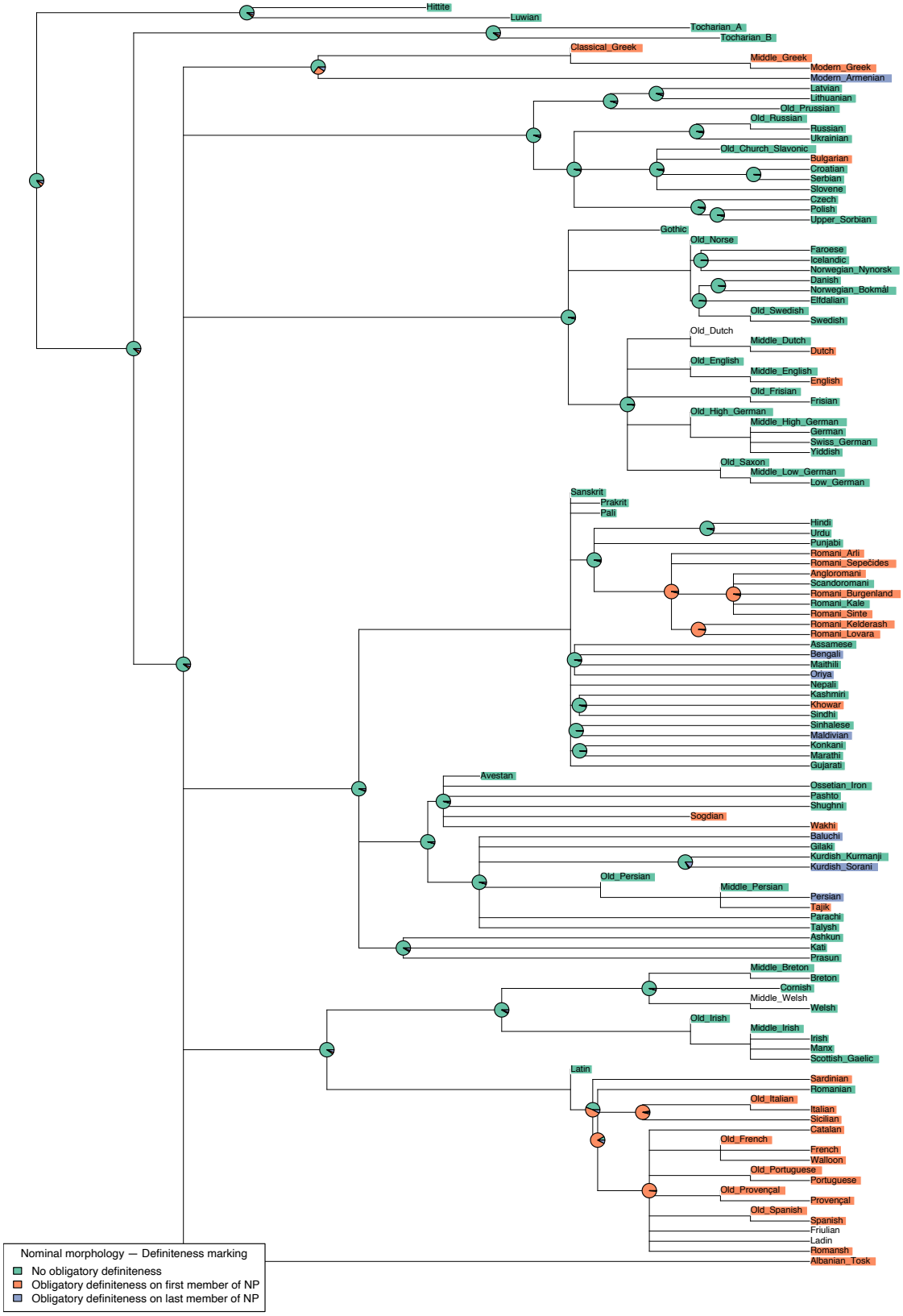

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Case marking
Case on article
No case on article

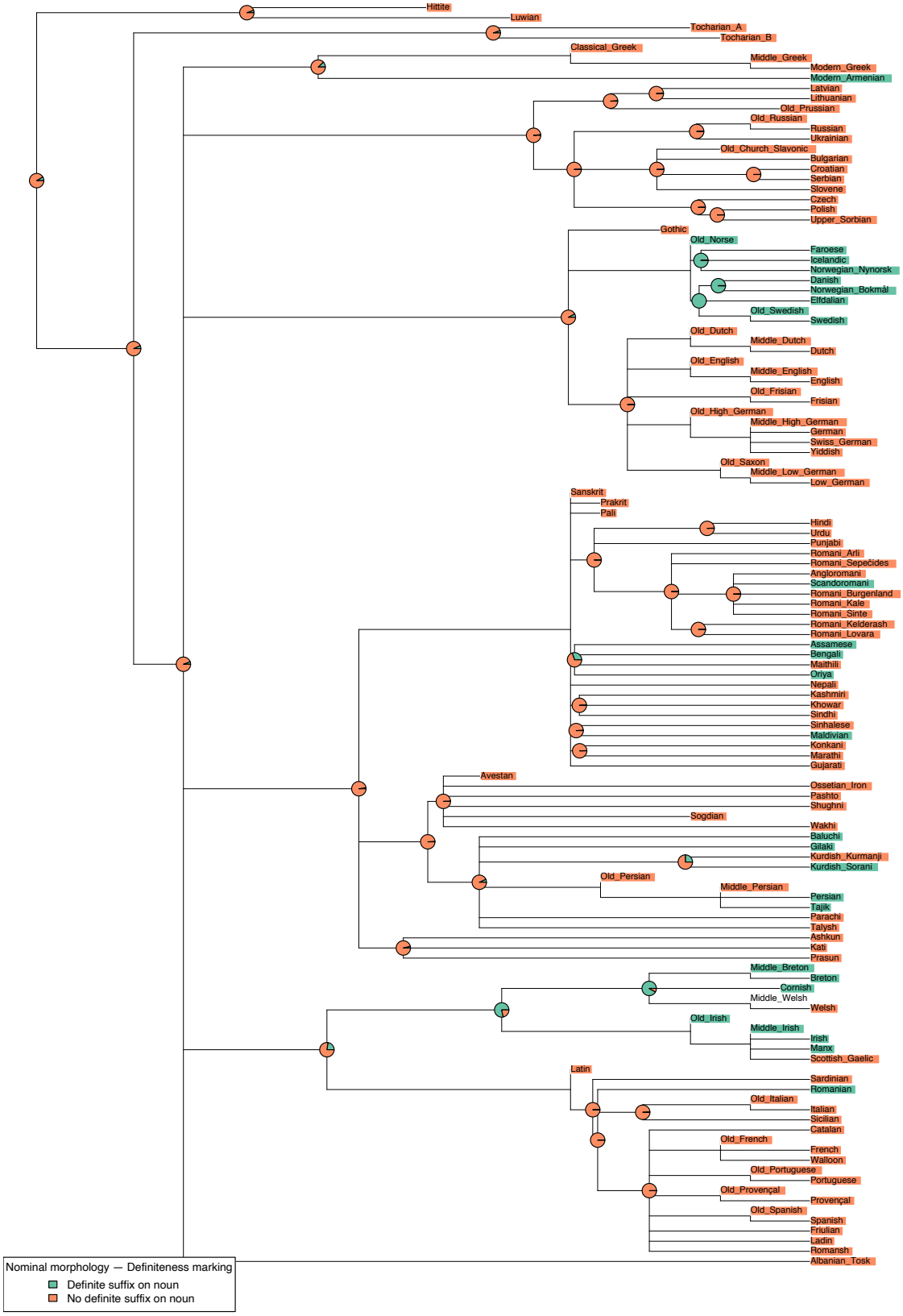

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Case marking
Case on last member of NP
No rule of case on last/first member of NP

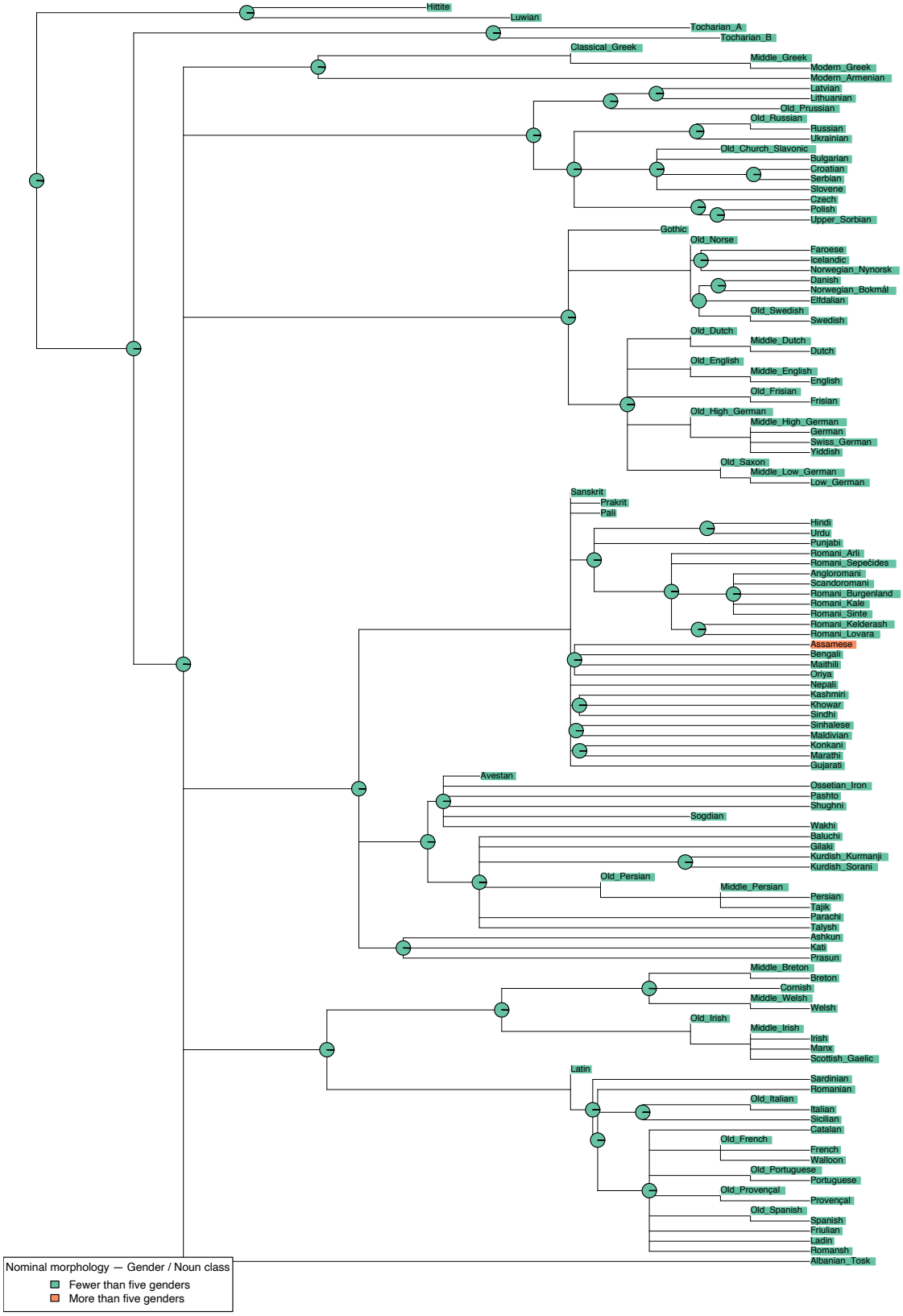

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Case marking
Case on noun
No case on noun

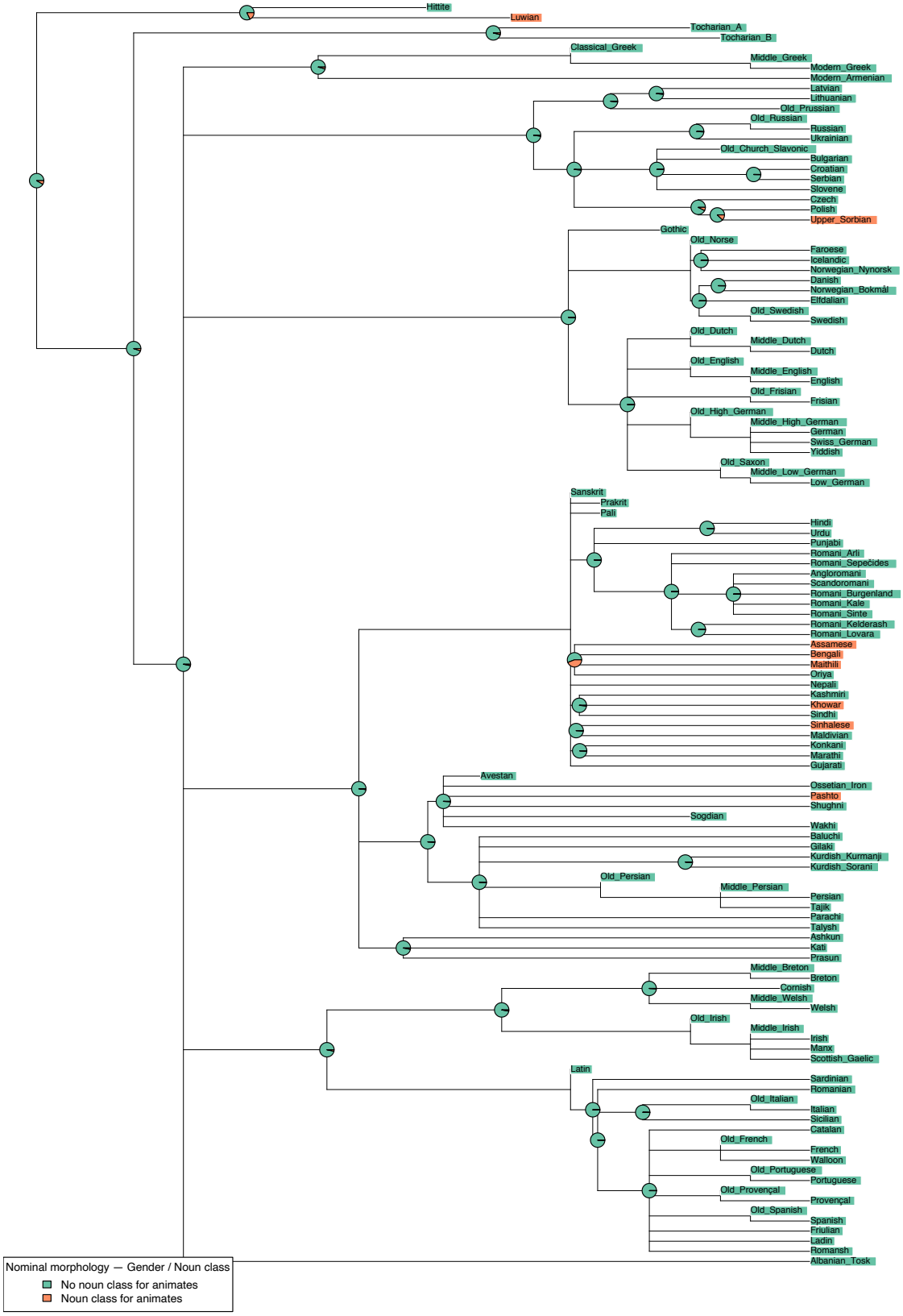

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Definiteness marking
Definite suffix on adjective
No definite suffix on adjective

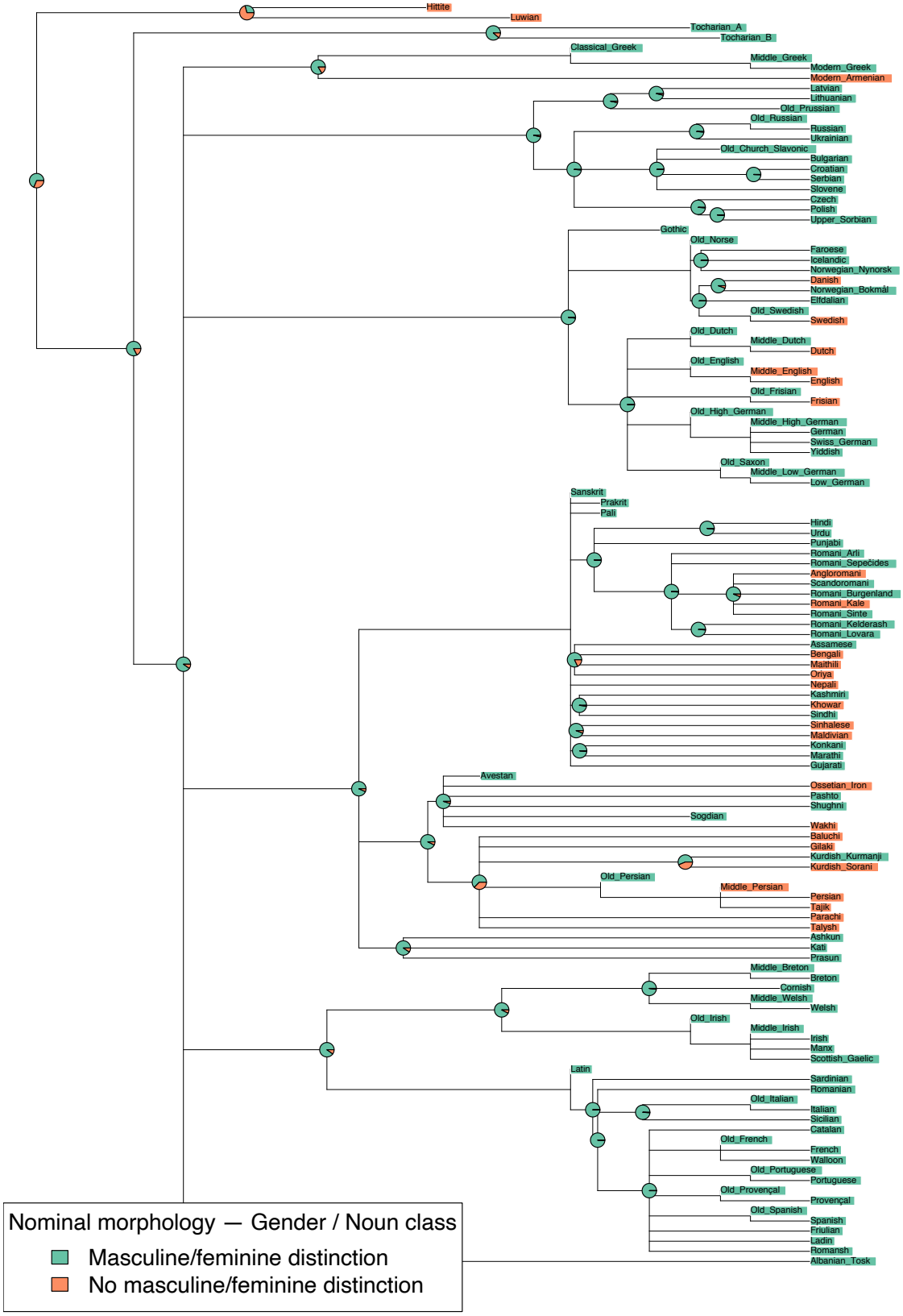

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talyshi
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Definiteness marking
Definite article
No definite article

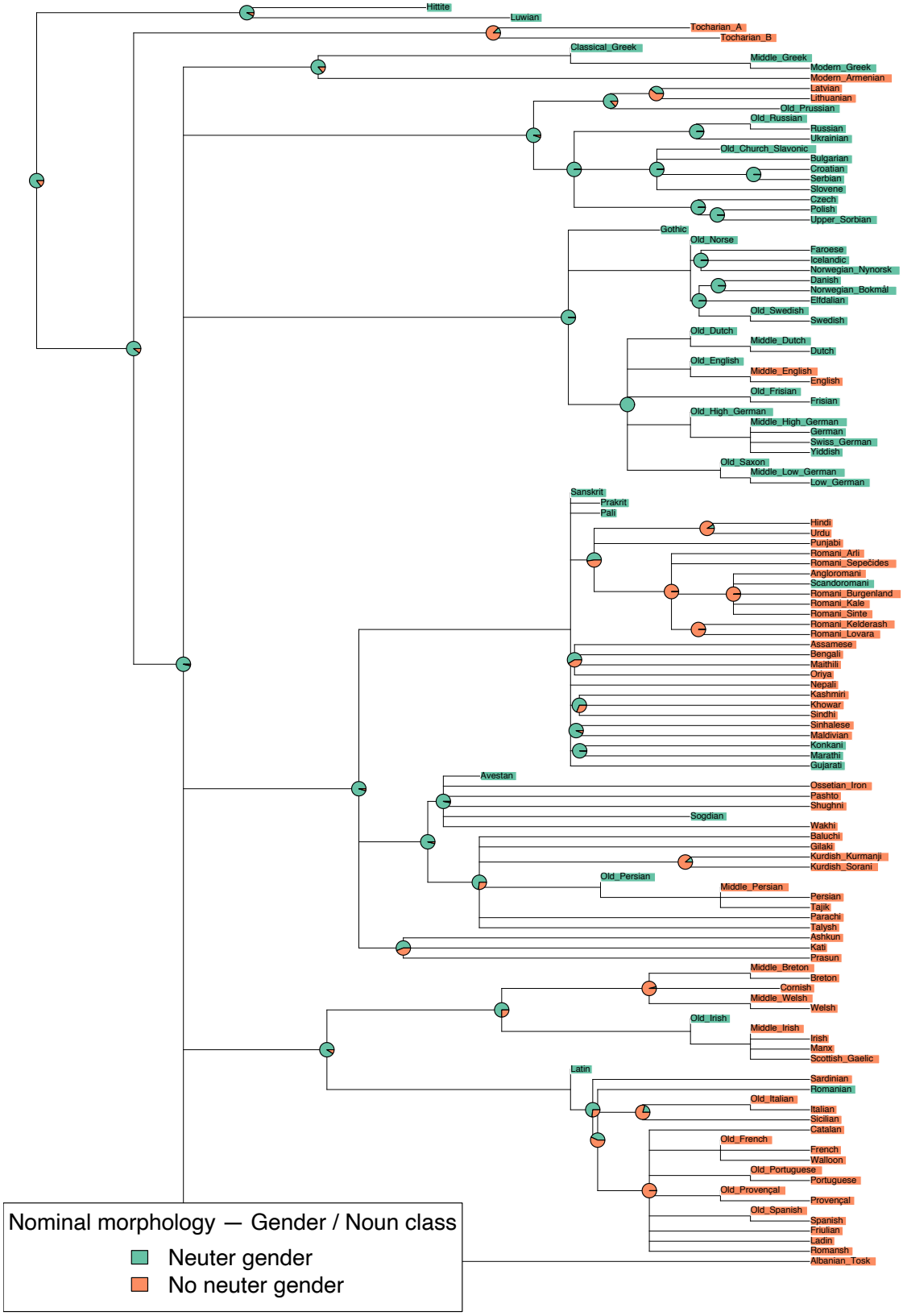

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Definiteness marking
No obligatory definiteness
Obligatory definiteness on first member of NP
Obligatory definiteness on last member of NP


Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Definiteness marking
Definite suffix on noun
No definite suffix on noun

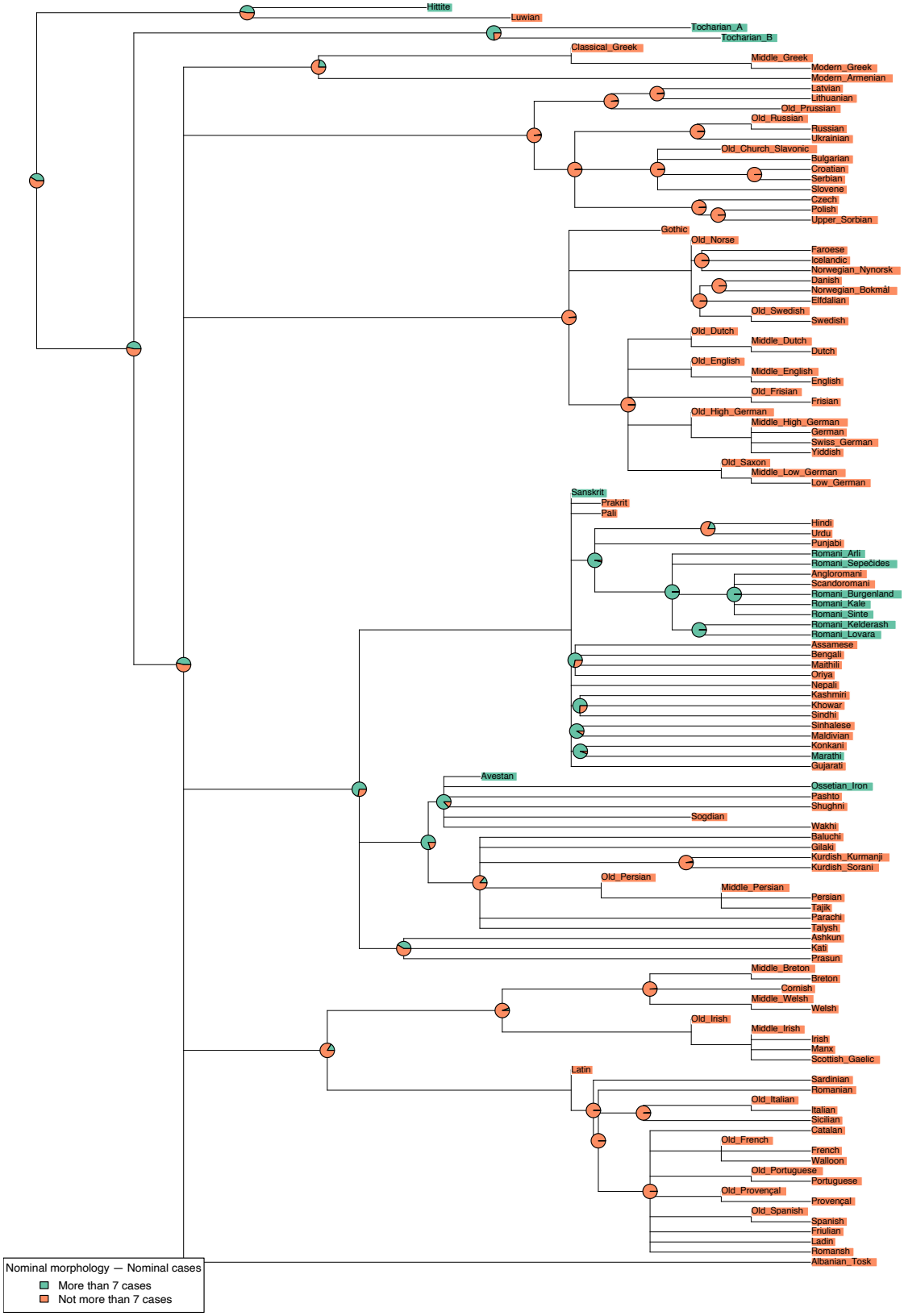

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Gender / Noun class
Fewer than five genders
More than five genders

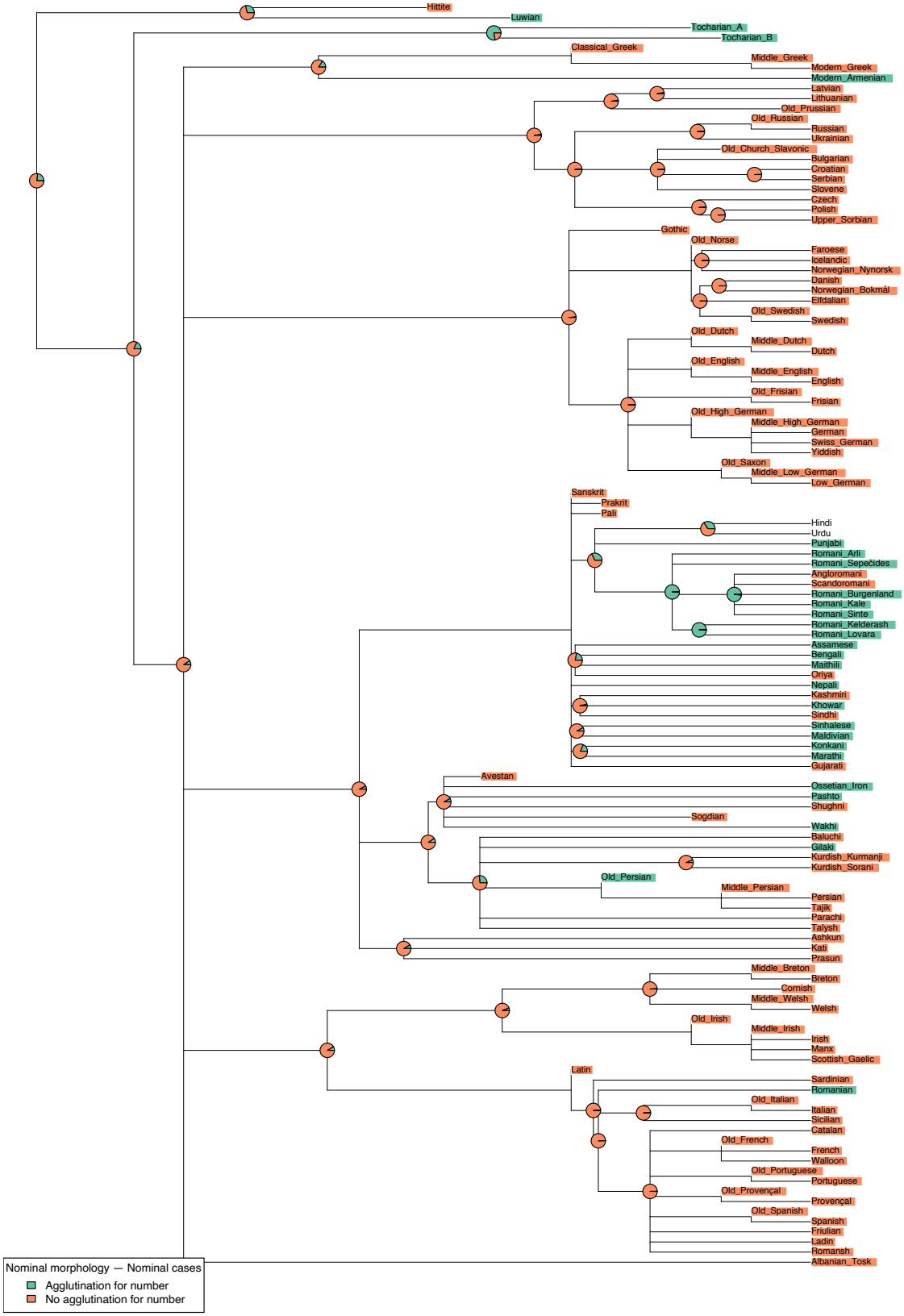

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Gender / Noun class
No noun class for animates
Noun class for animates

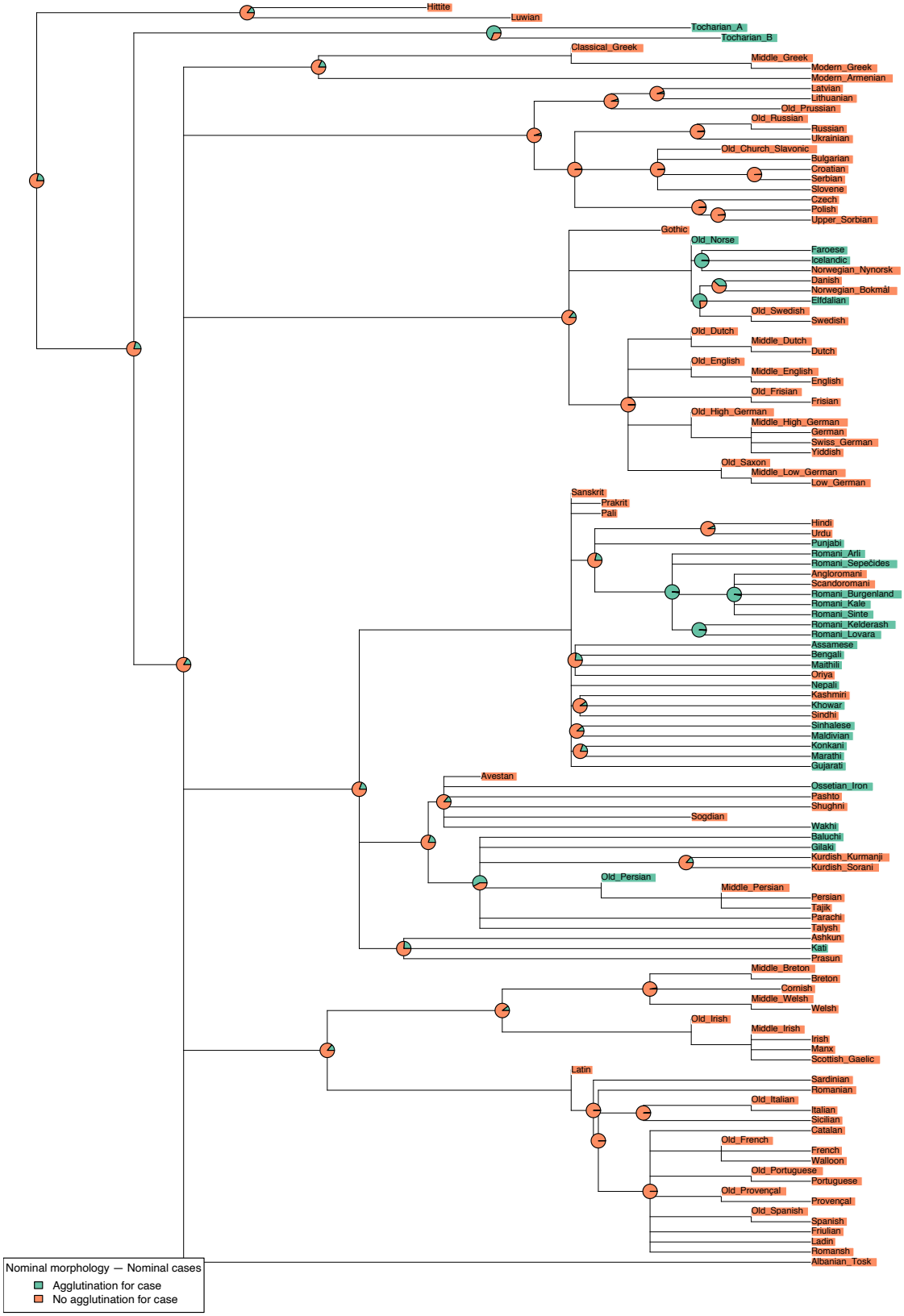

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Gender / Noun class
Masculine/feminine distinction
No masculine/feminine distinction

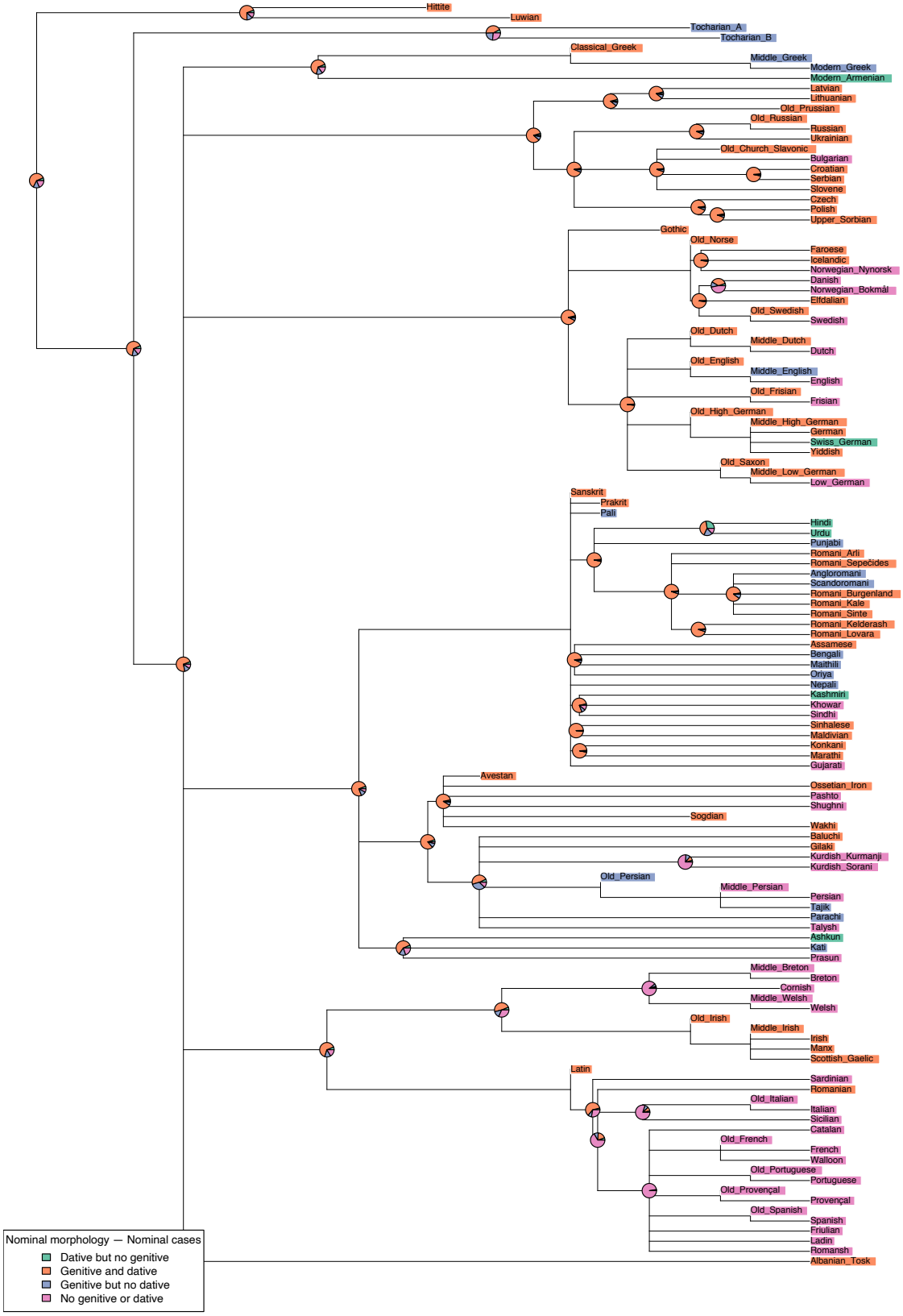

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Gender / Noun class
Neuter gender
No neuter gender

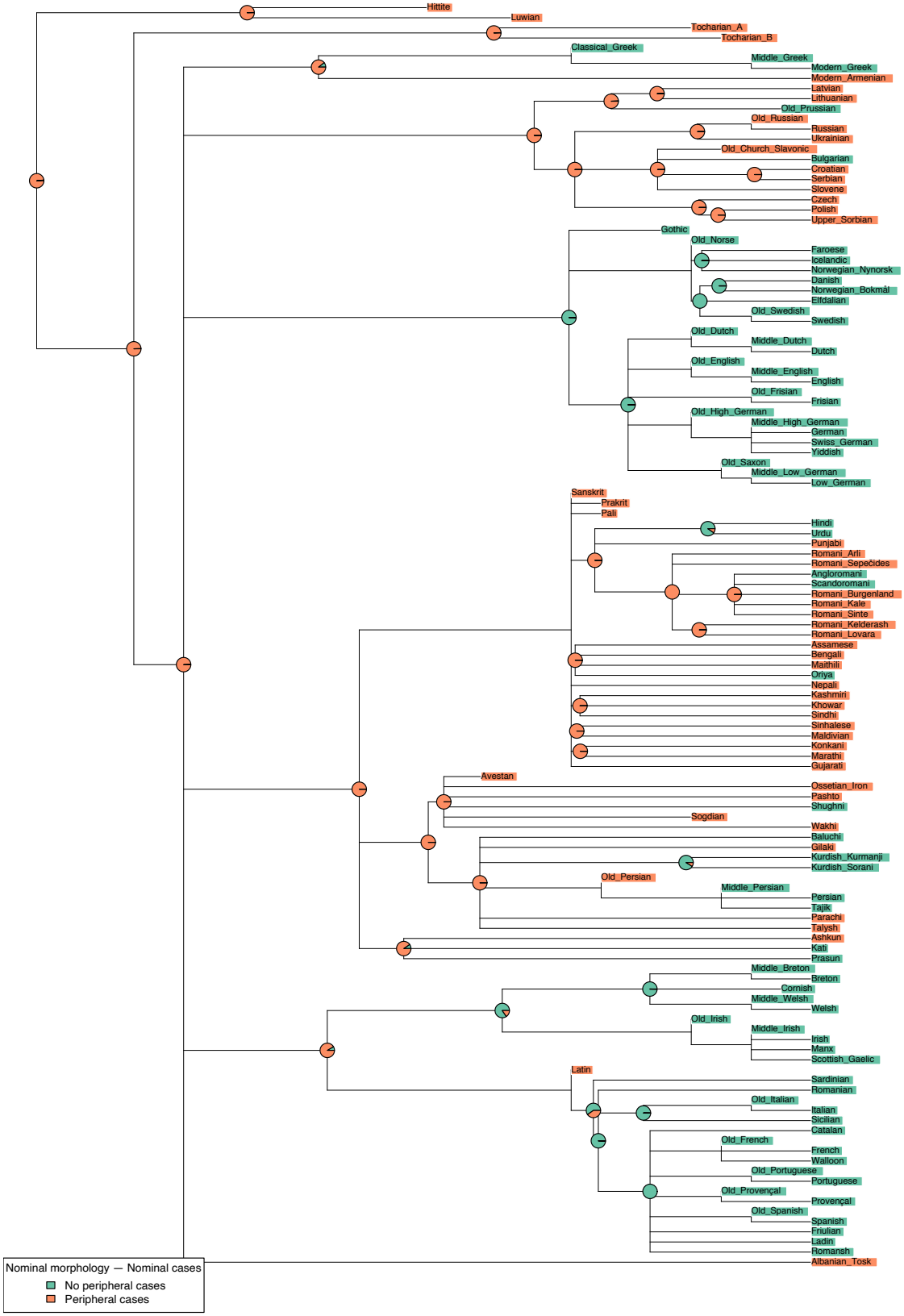

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Gender agreement
Gender on predicative adjective
No gender on predicative adjective

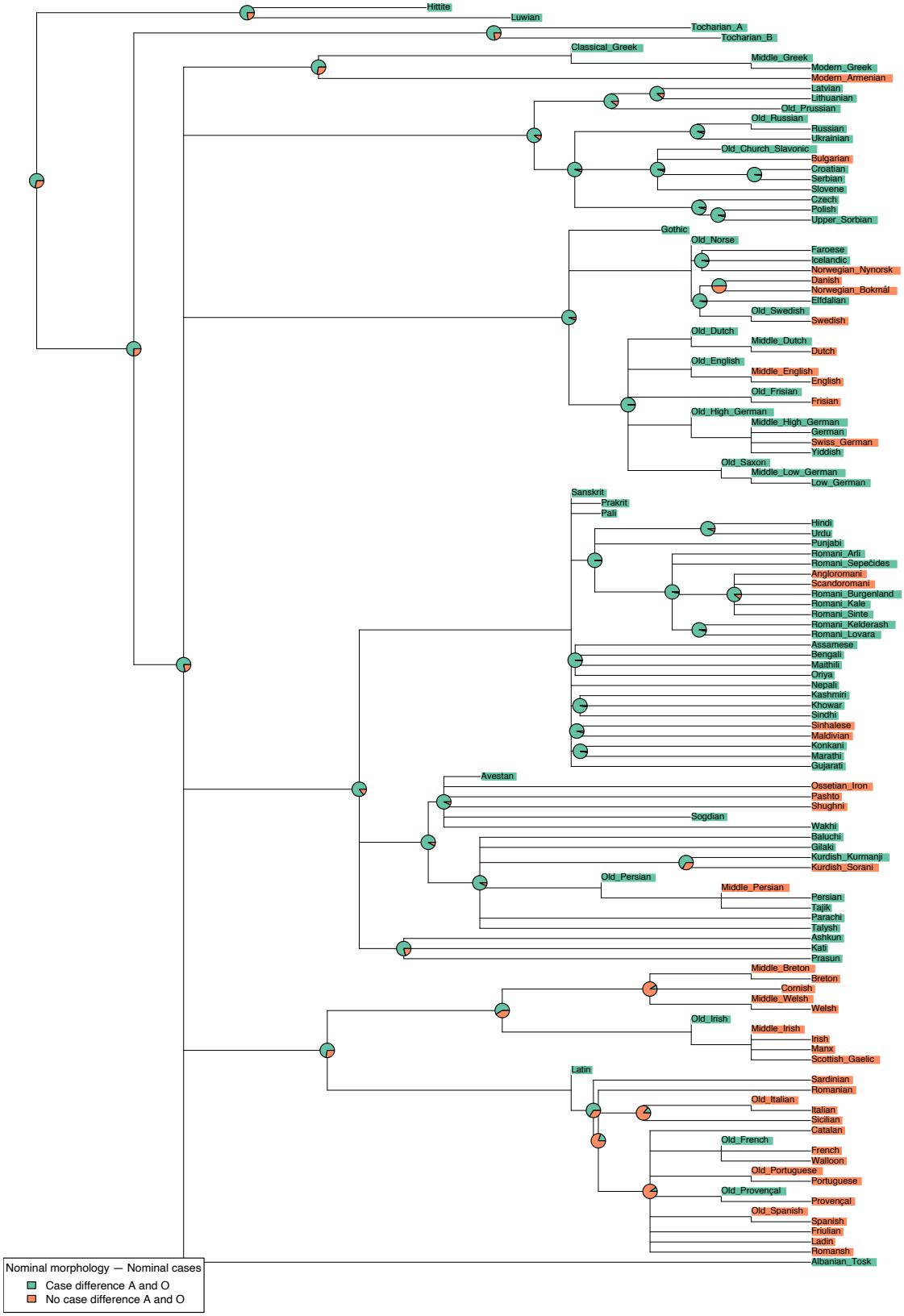

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Nominal cases
More than 7 cases
Not more than 7 cases

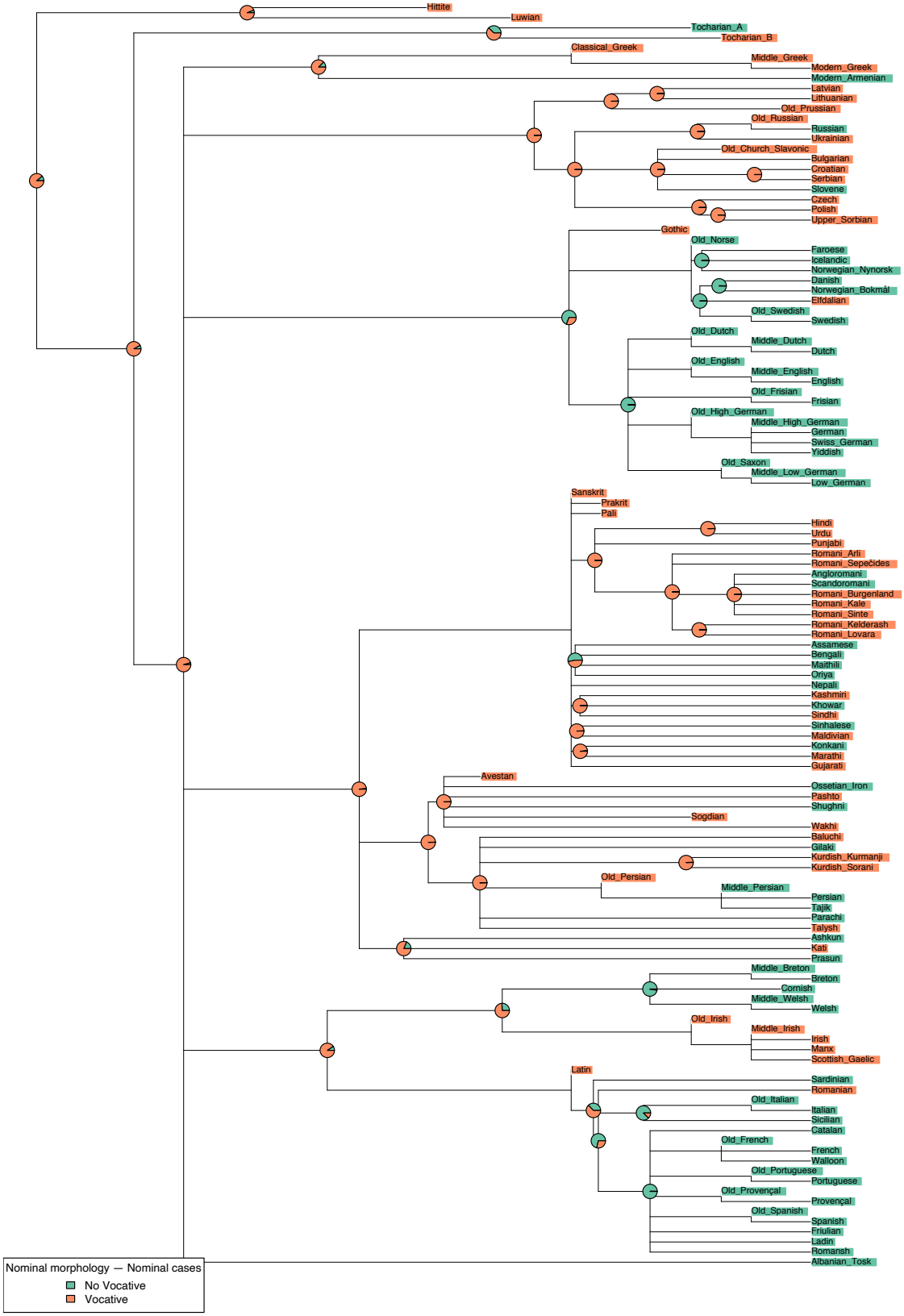

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Nominal cases
Agglutination for number
No agglutination for number

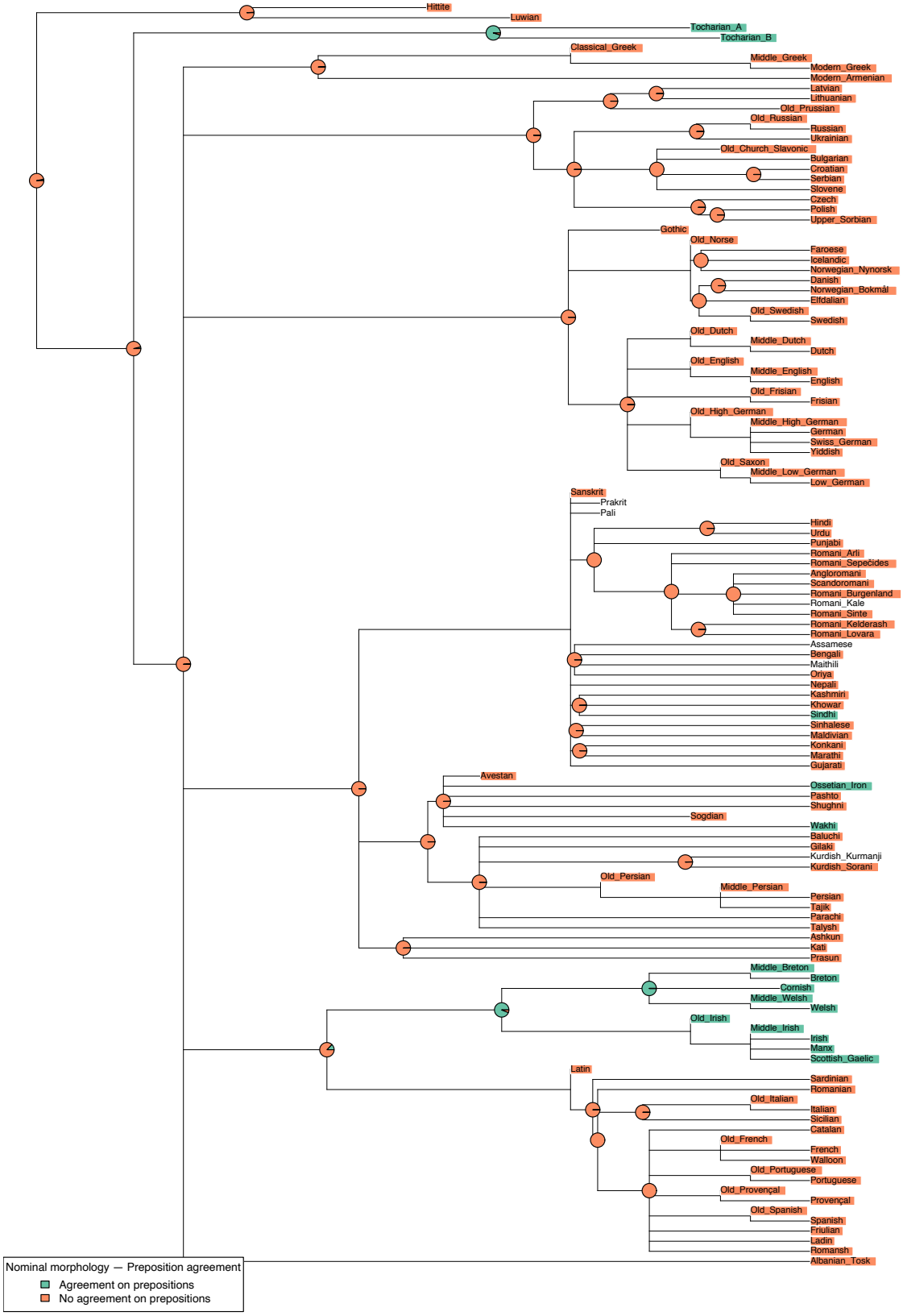

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Nominal cases
Agglutination for case
No agglutination for case

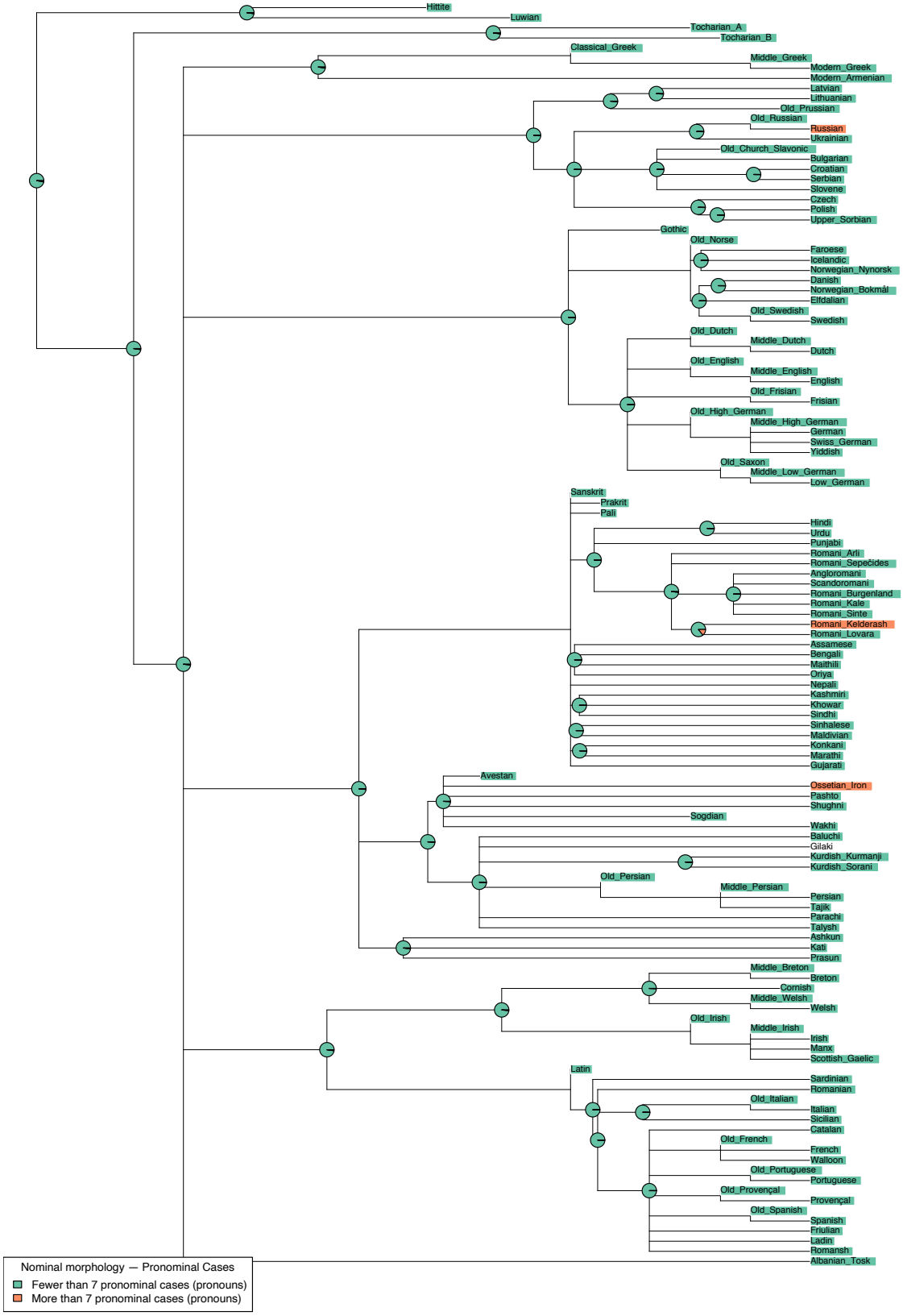

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Nominal cases
Dative but no genitive
Genitive and dative
Genitive but no dative
No genitive or dative

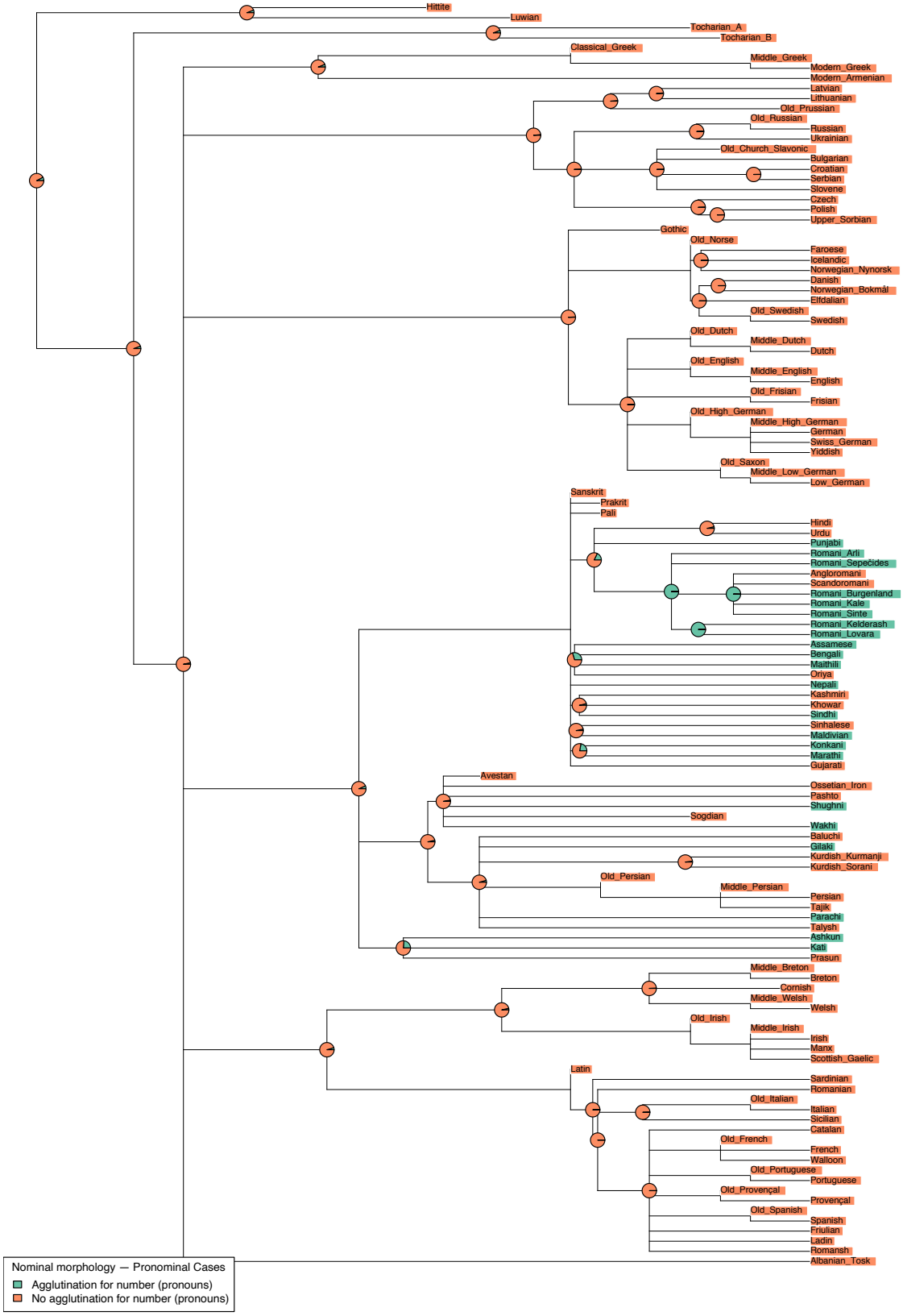

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Nominal cases
No peripheral cases
Peripheral cases

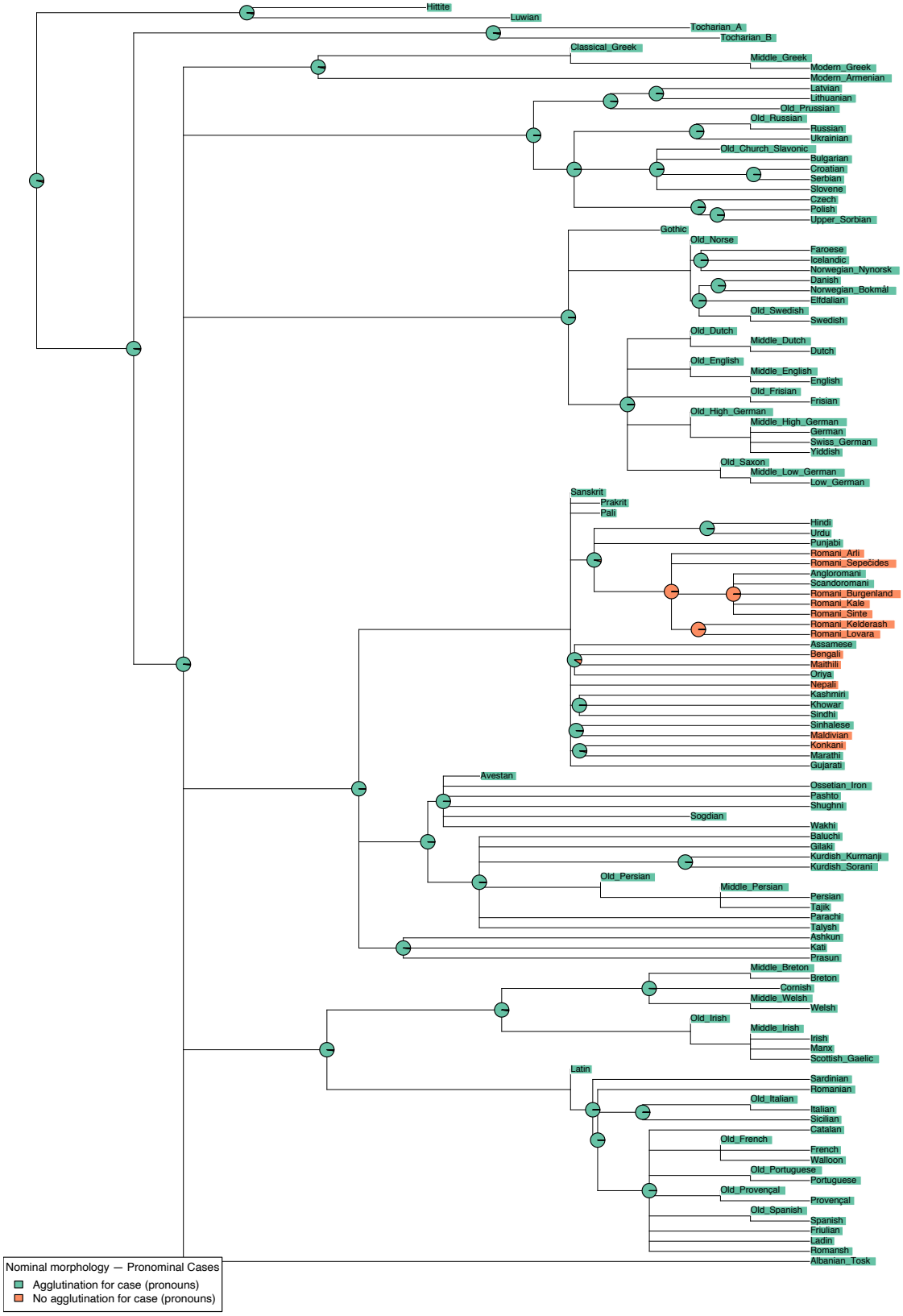

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Nominal cases
Case difference A and O
No case difference A and O

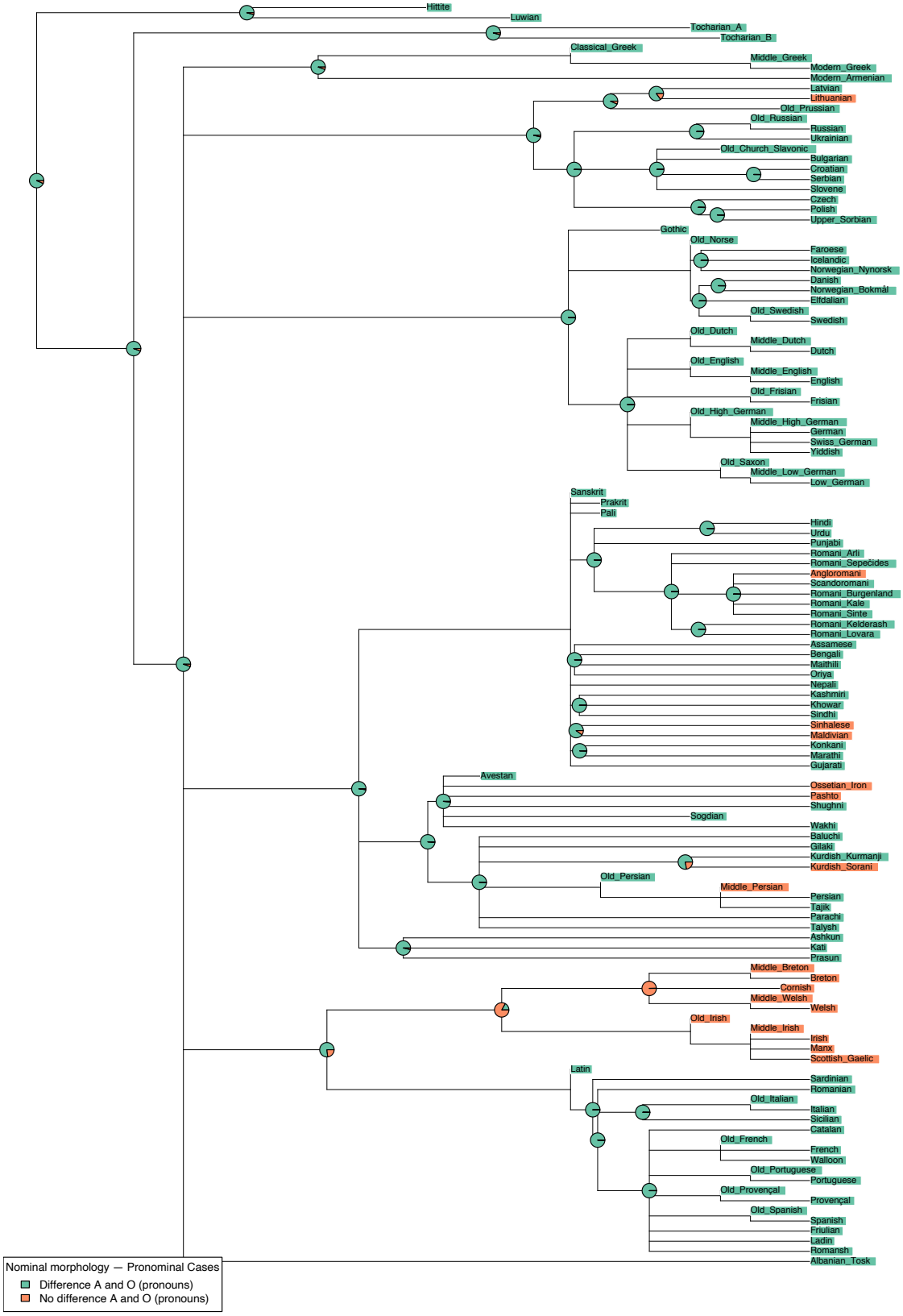

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Nominal cases
No Vocative
Vocative

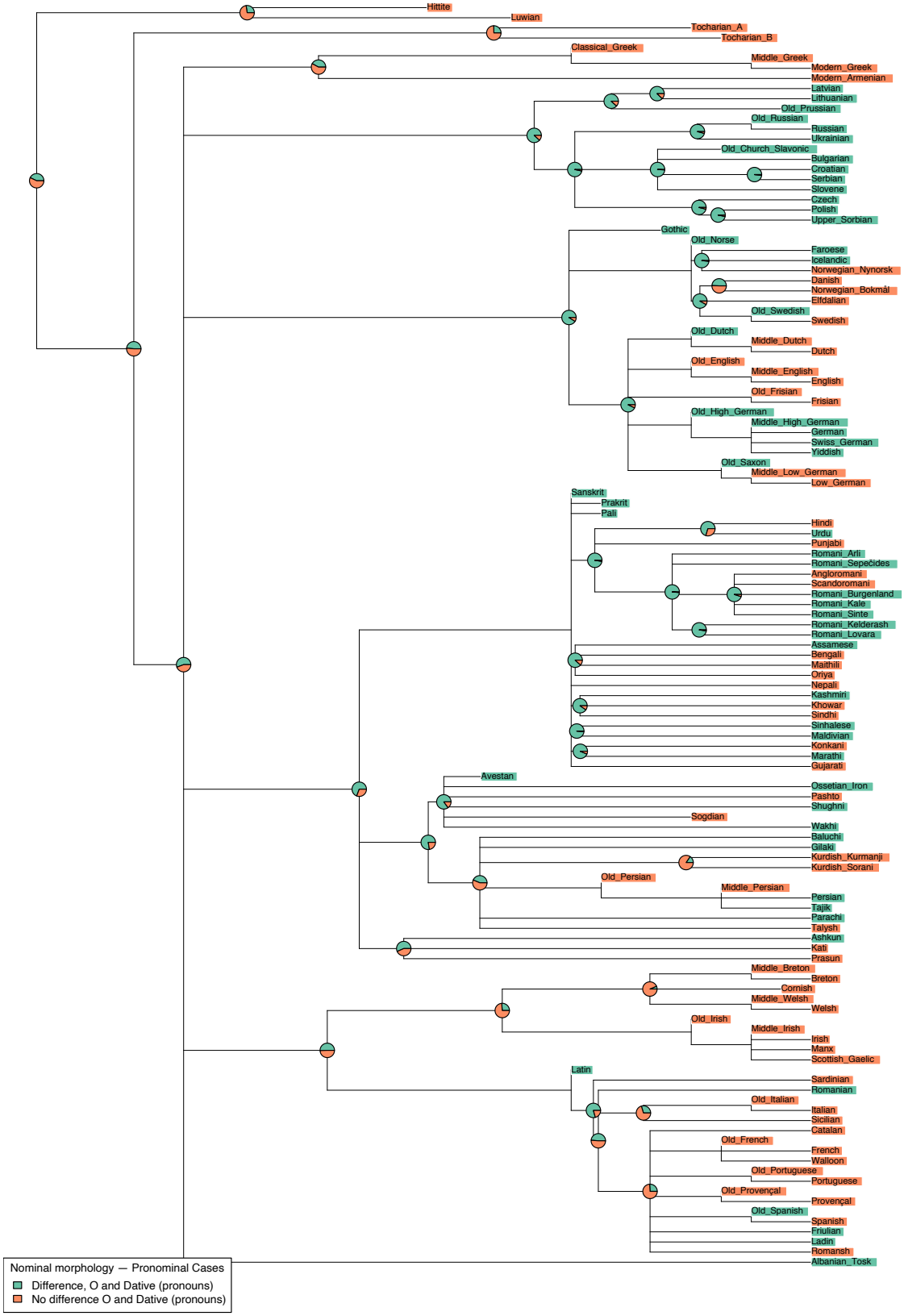

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Preposition agreement
Agreement on prepositions
No agreement on prepositions

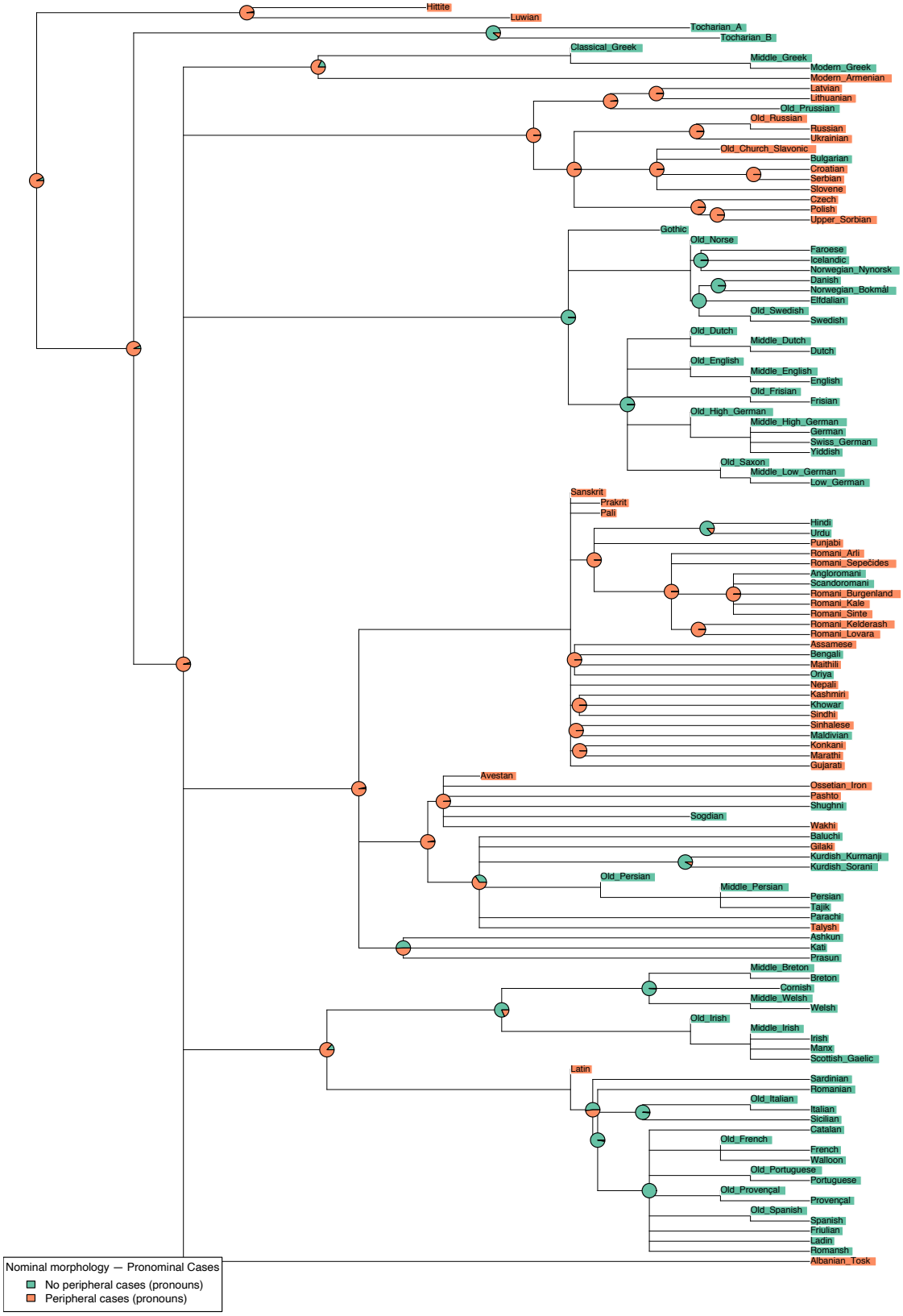

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Pronominal Cases
Fewer than 7 pronominal cases (pronouns)
More than 7 pronominal cases (pronouns)

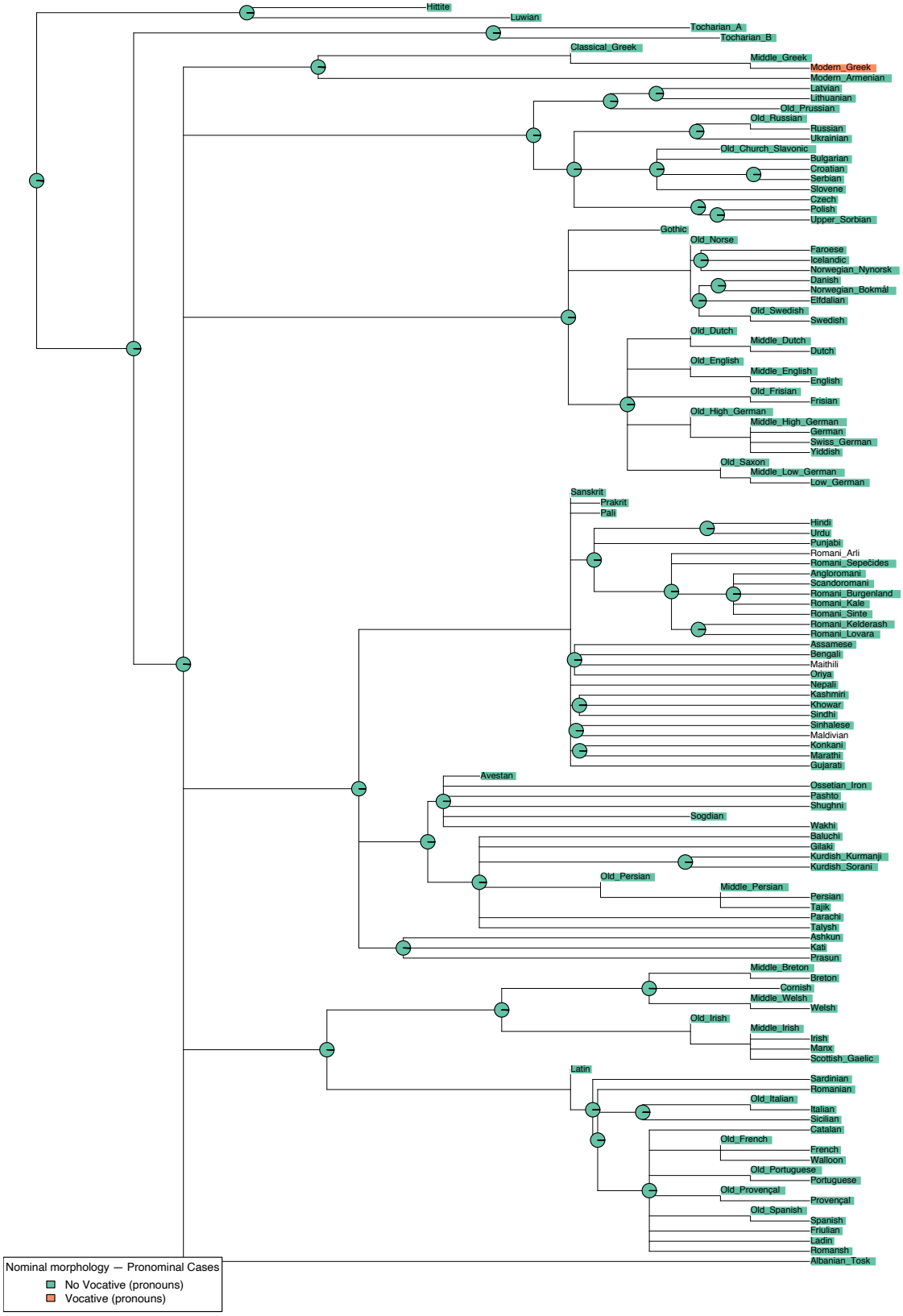

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Pronominal Cases
Agglutination for number (pronouns)
No agglutination for number (pronouns)

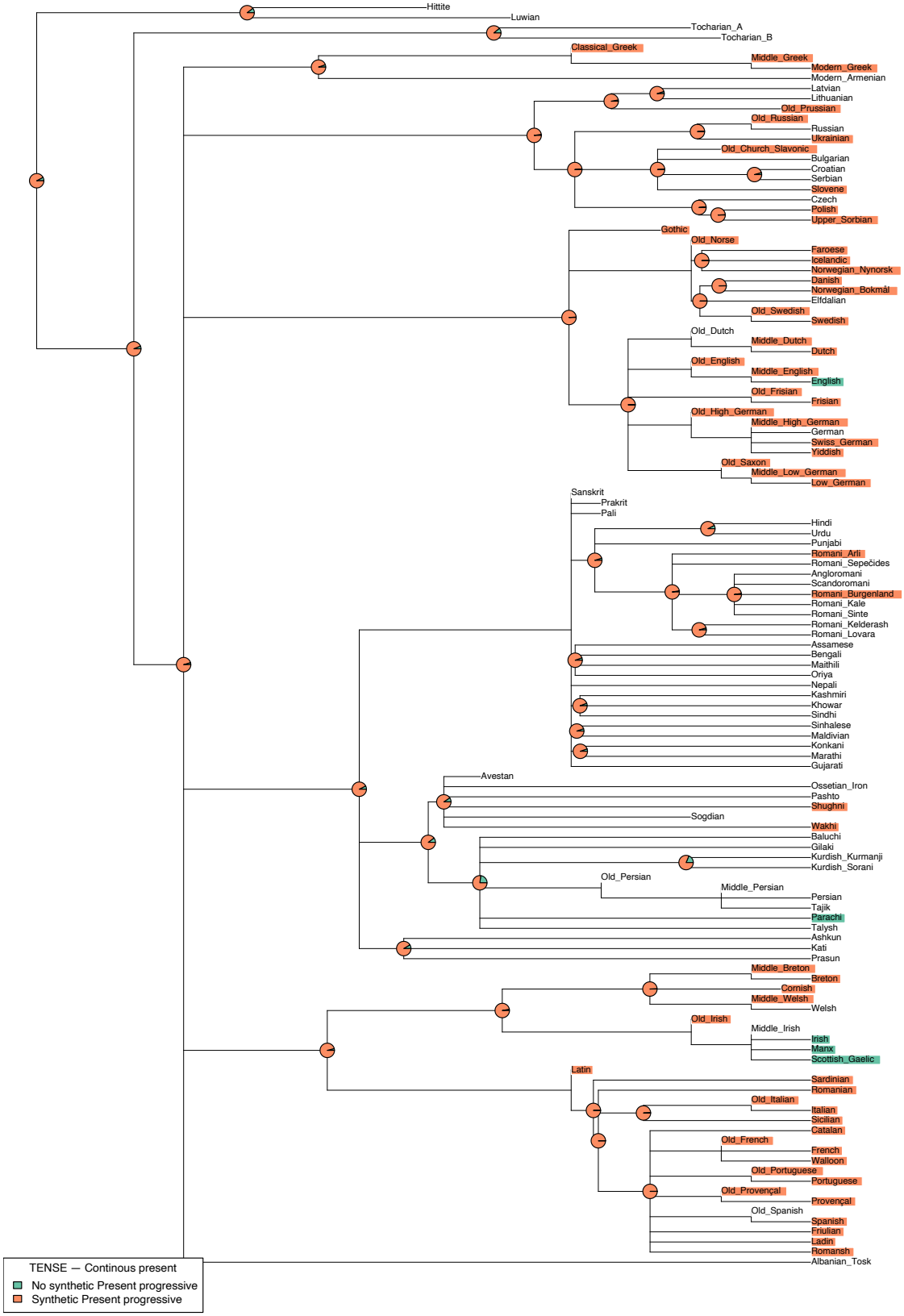

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Pronominal Cases
Agglutination for case (pronouns)
No agglutination for case (pronouns)

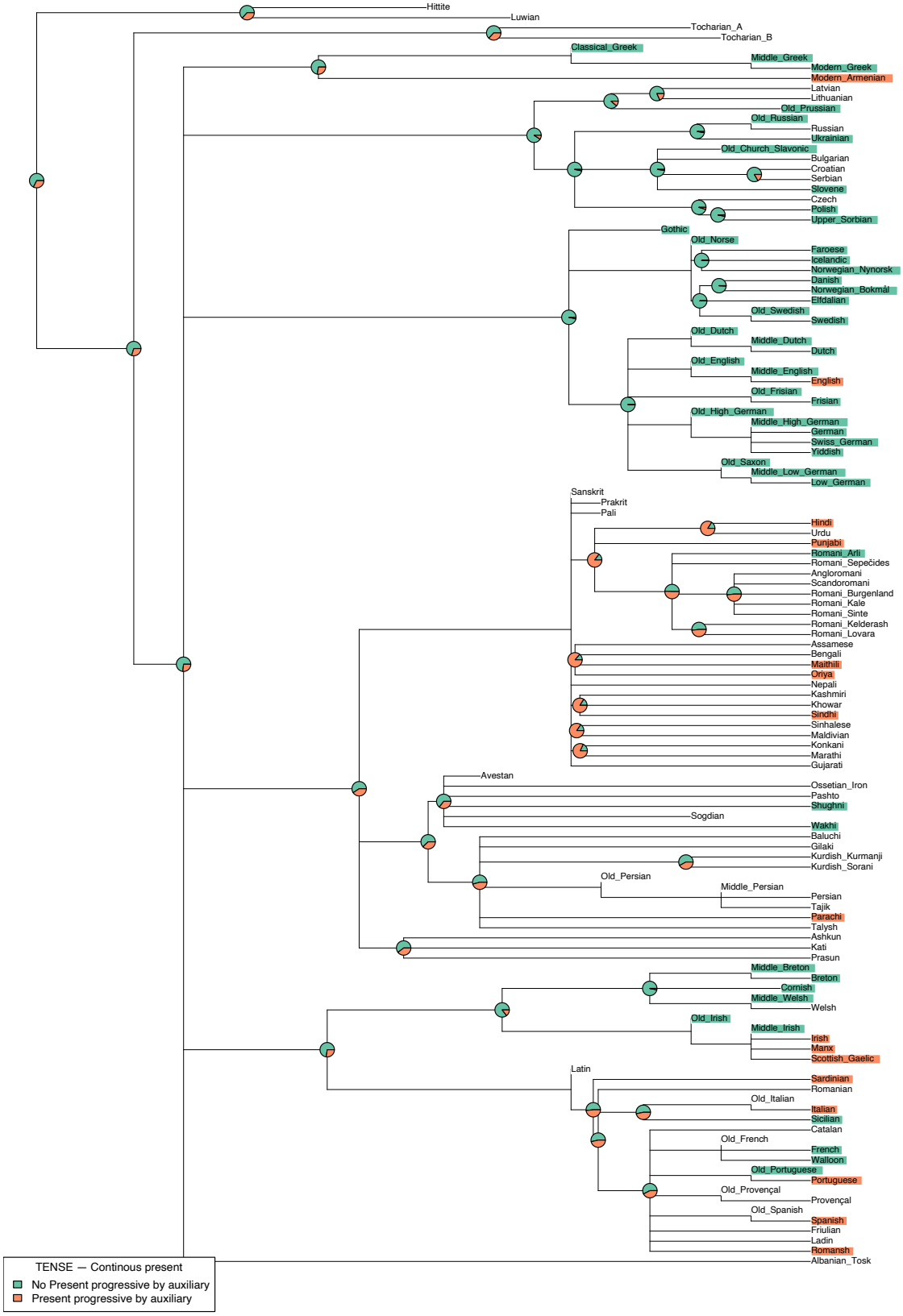

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Pronominal Cases
Difference A and O (pronouns)
No difference A and O (pronouns)

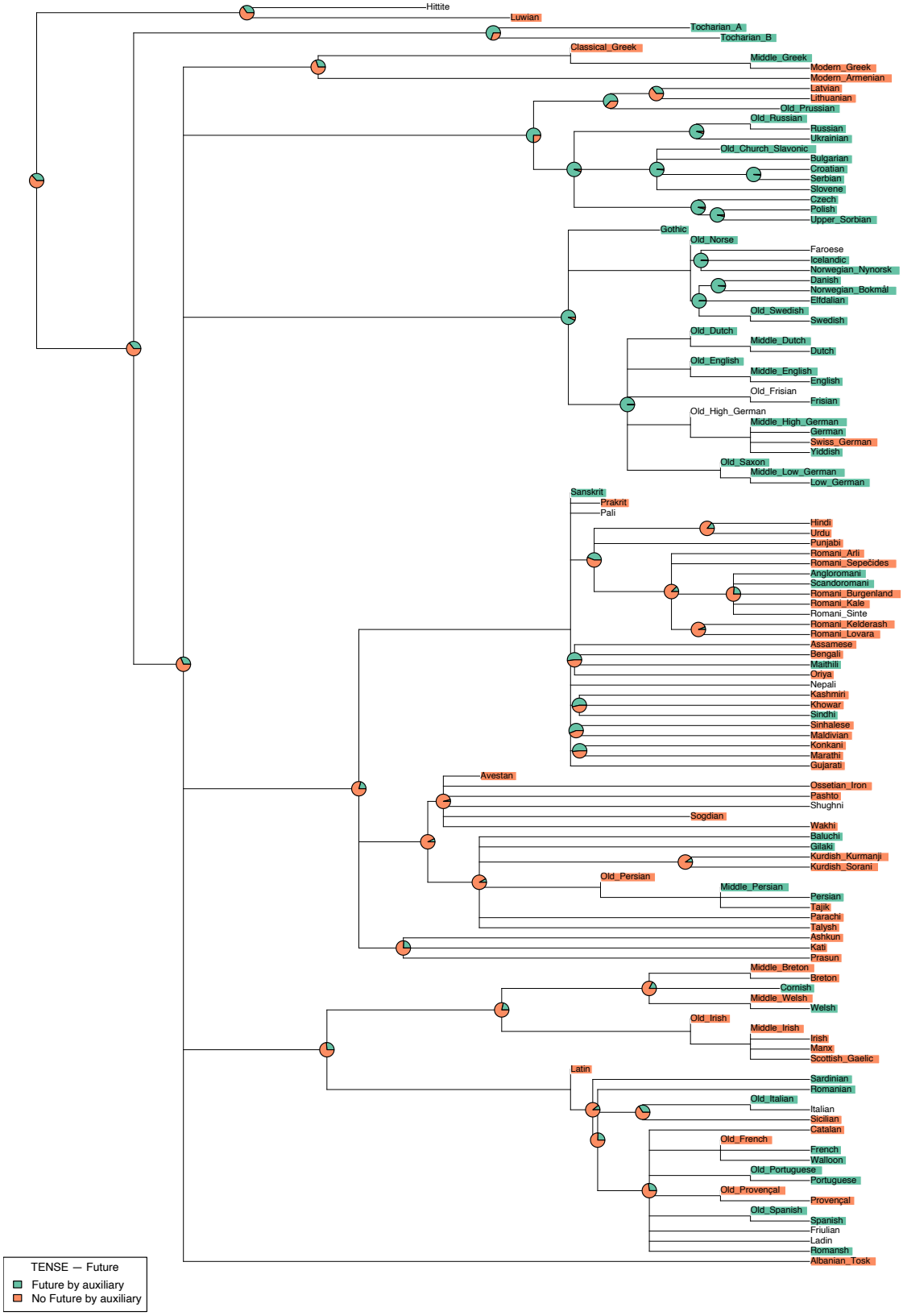

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Pronominal Cases
Difference, O and Dative (pronouns)
No difference O and Dative (pronouns)

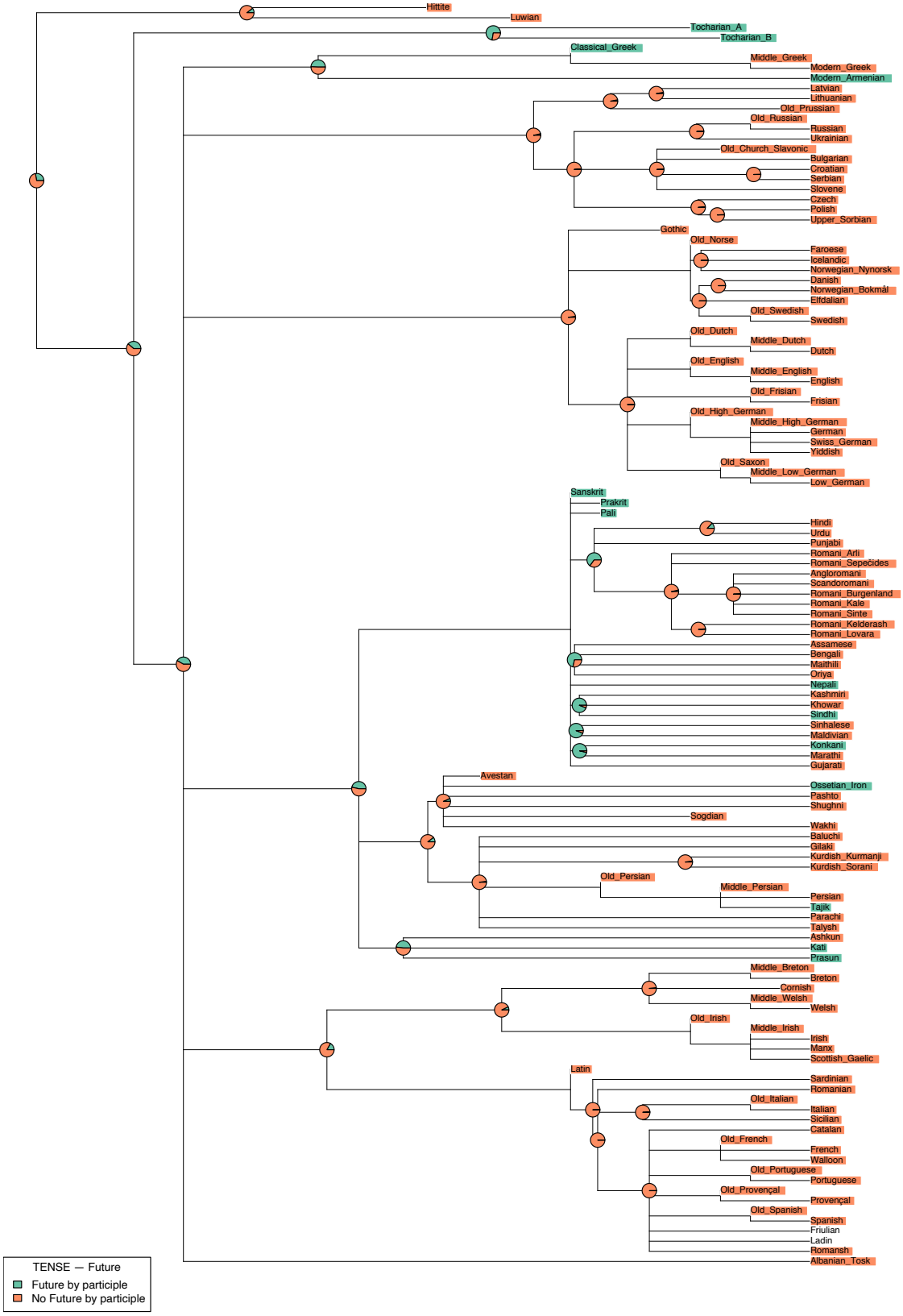

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Pronominal Cases
No peripheral cases (pronouns)
Peripheral cases (pronouns)


Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Nominal morphology — Pronominal Cases
No Vocative (pronouns)
Vocative (pronouns)

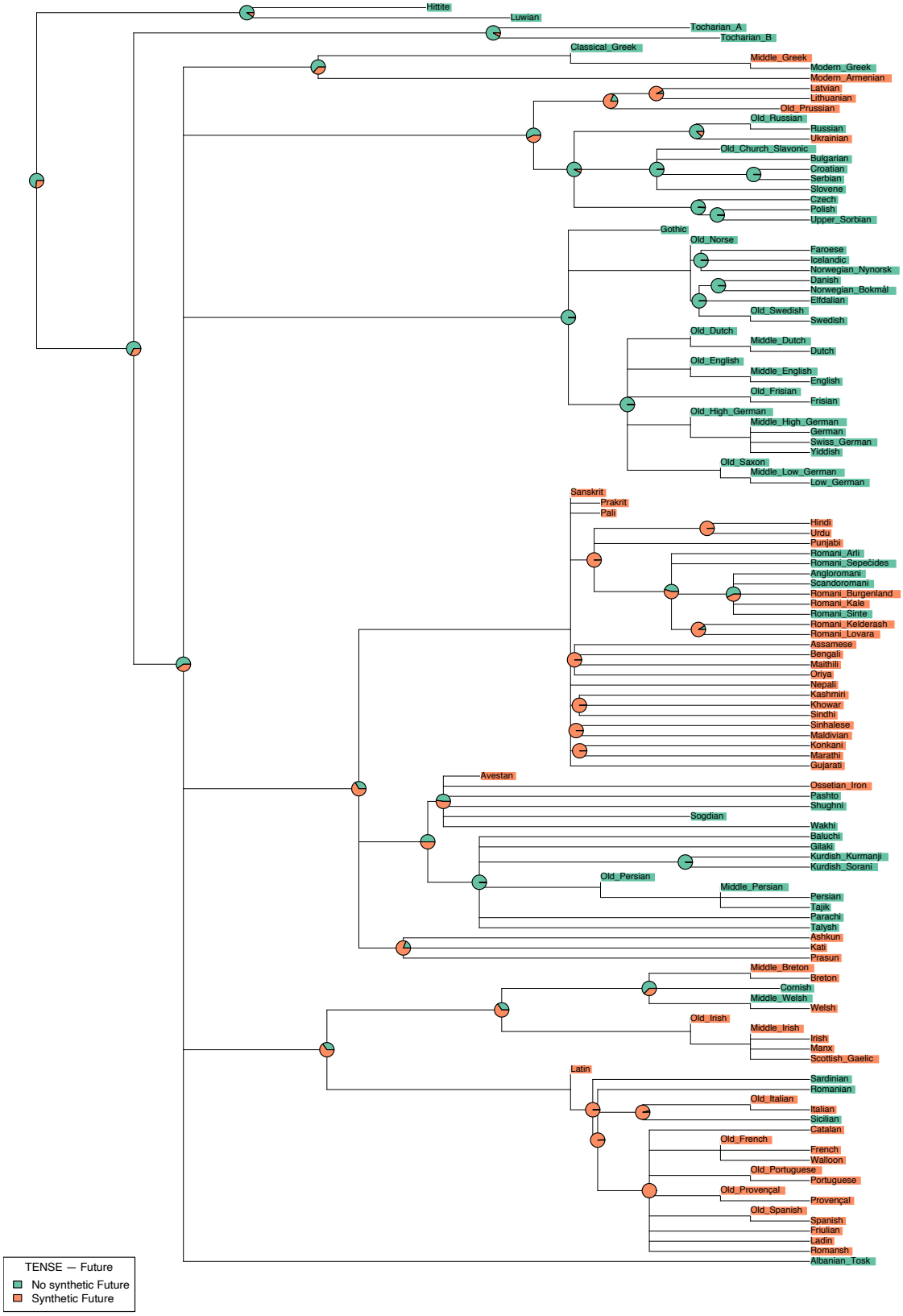

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
TENSE — Continous present
No synthetic Present progressive
Synthetic Present progressive


Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
TENSE — Continous present
No Present progressive by auxiliary
Present progressive by auxiliary

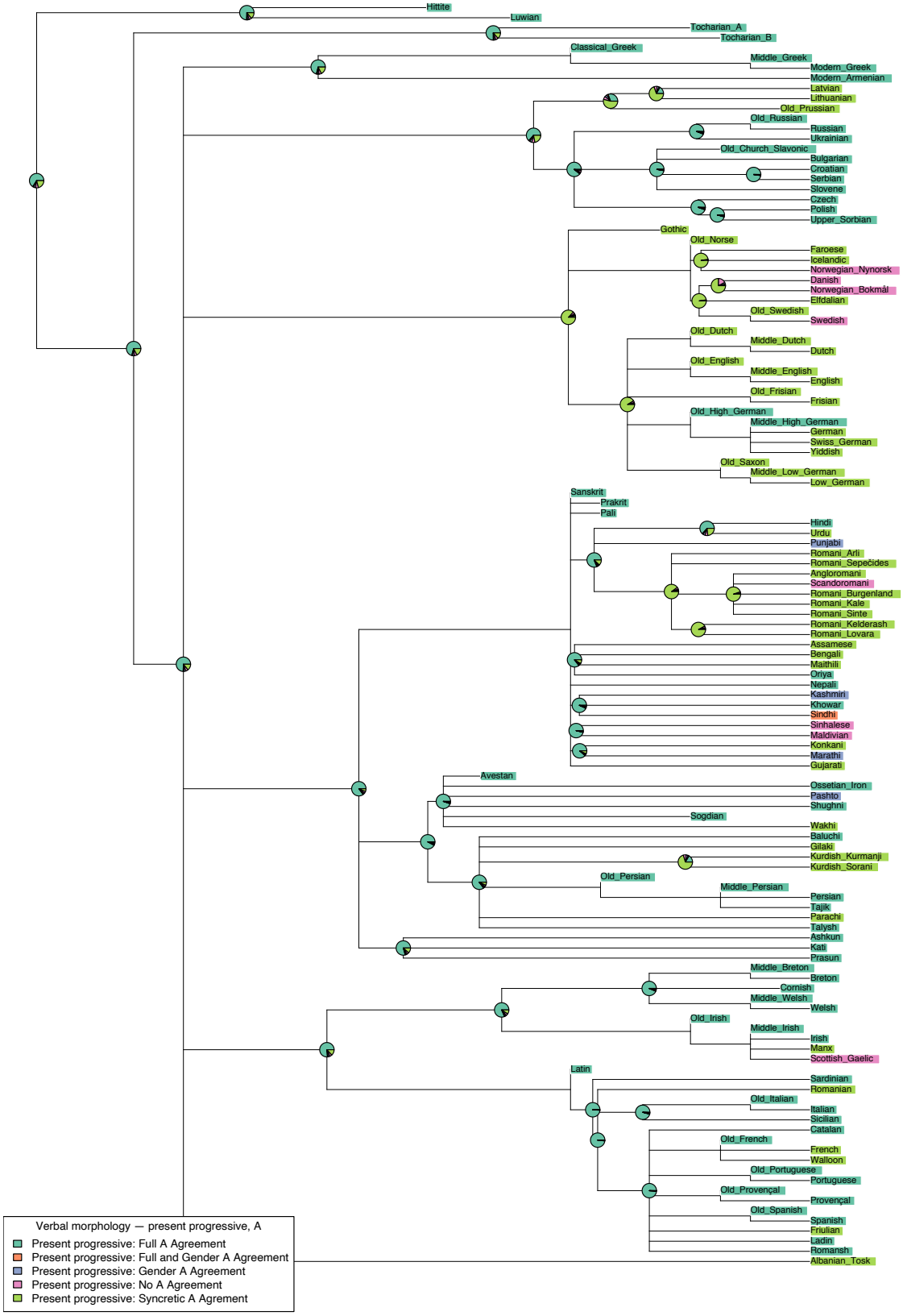

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
TENSE — Future
Future by auxiliary
No Future by auxiliary

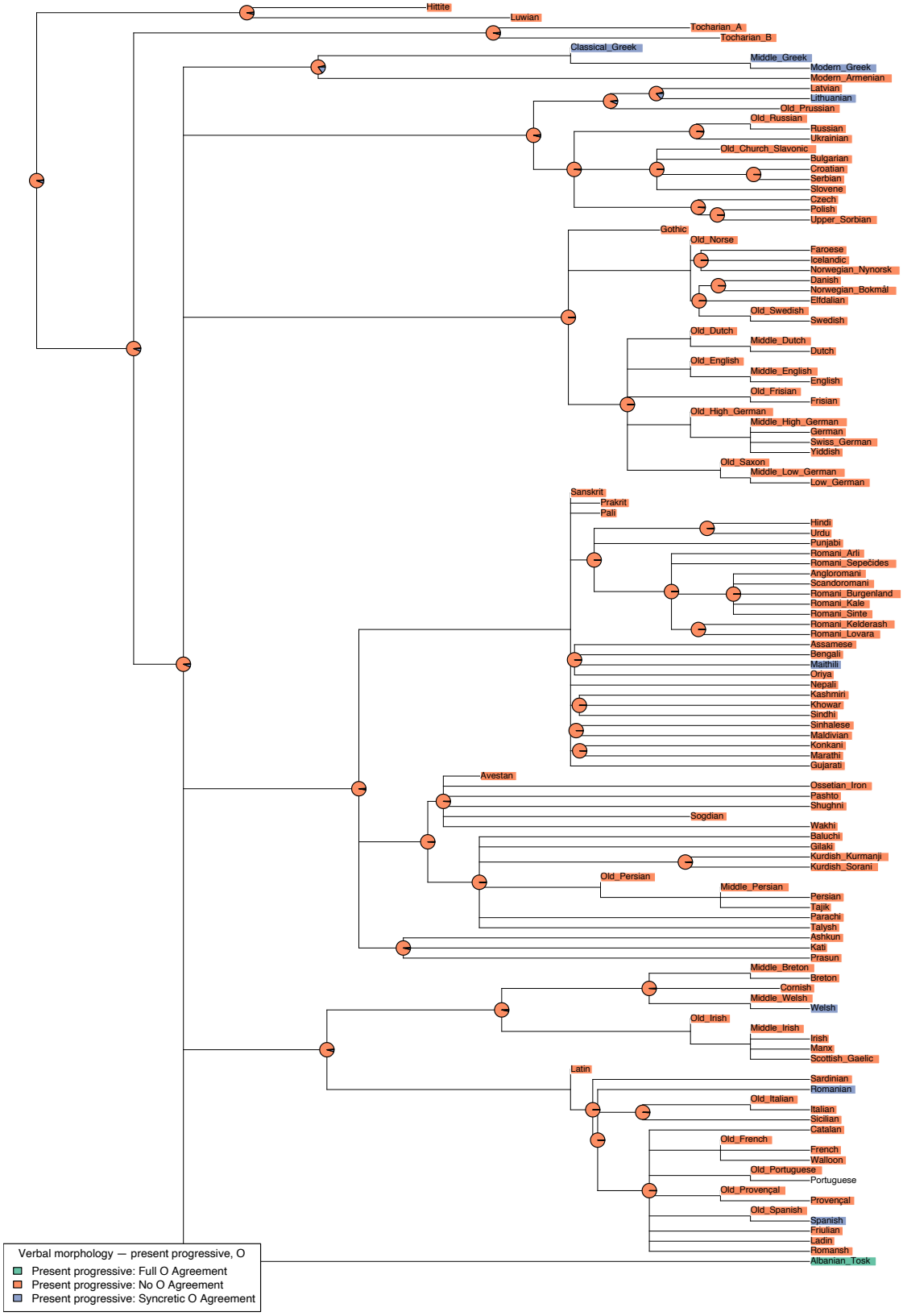

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
TENSE — Future
Future by participle
No Future by participle

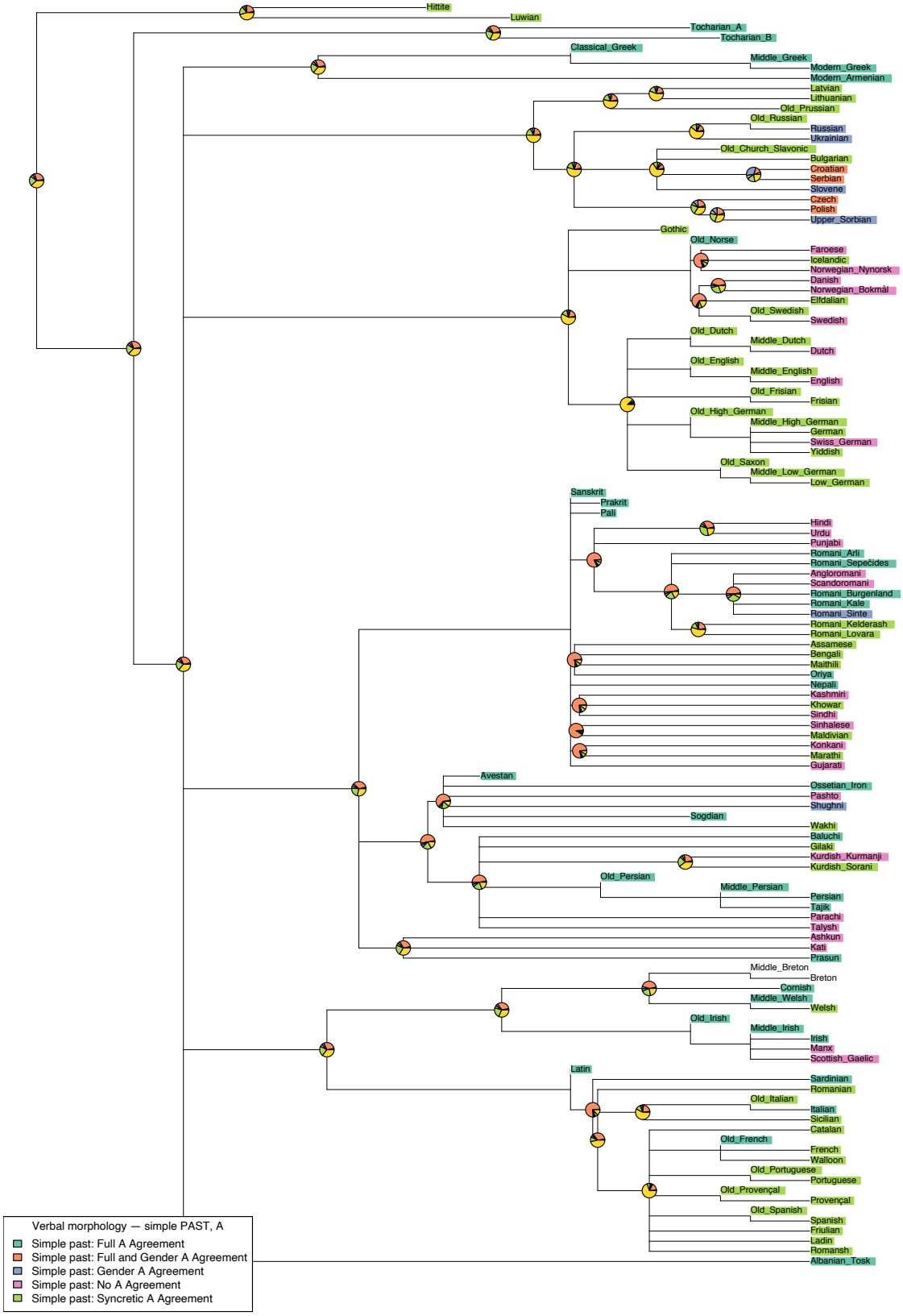

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
TENSE — Future
Future by particle
No Future by particle

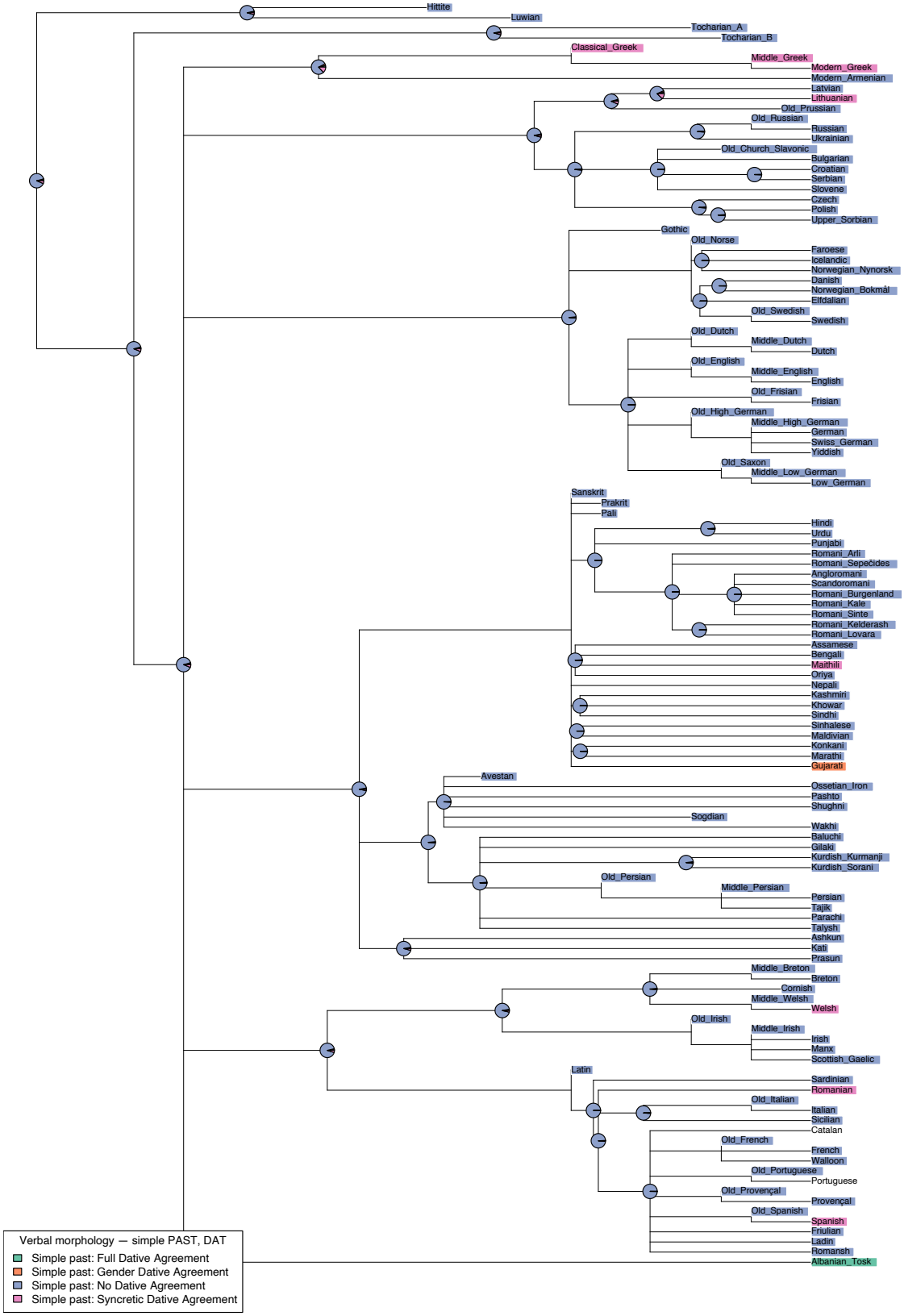

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
TENSE — Future
No synthetic Future
Synthetic Future

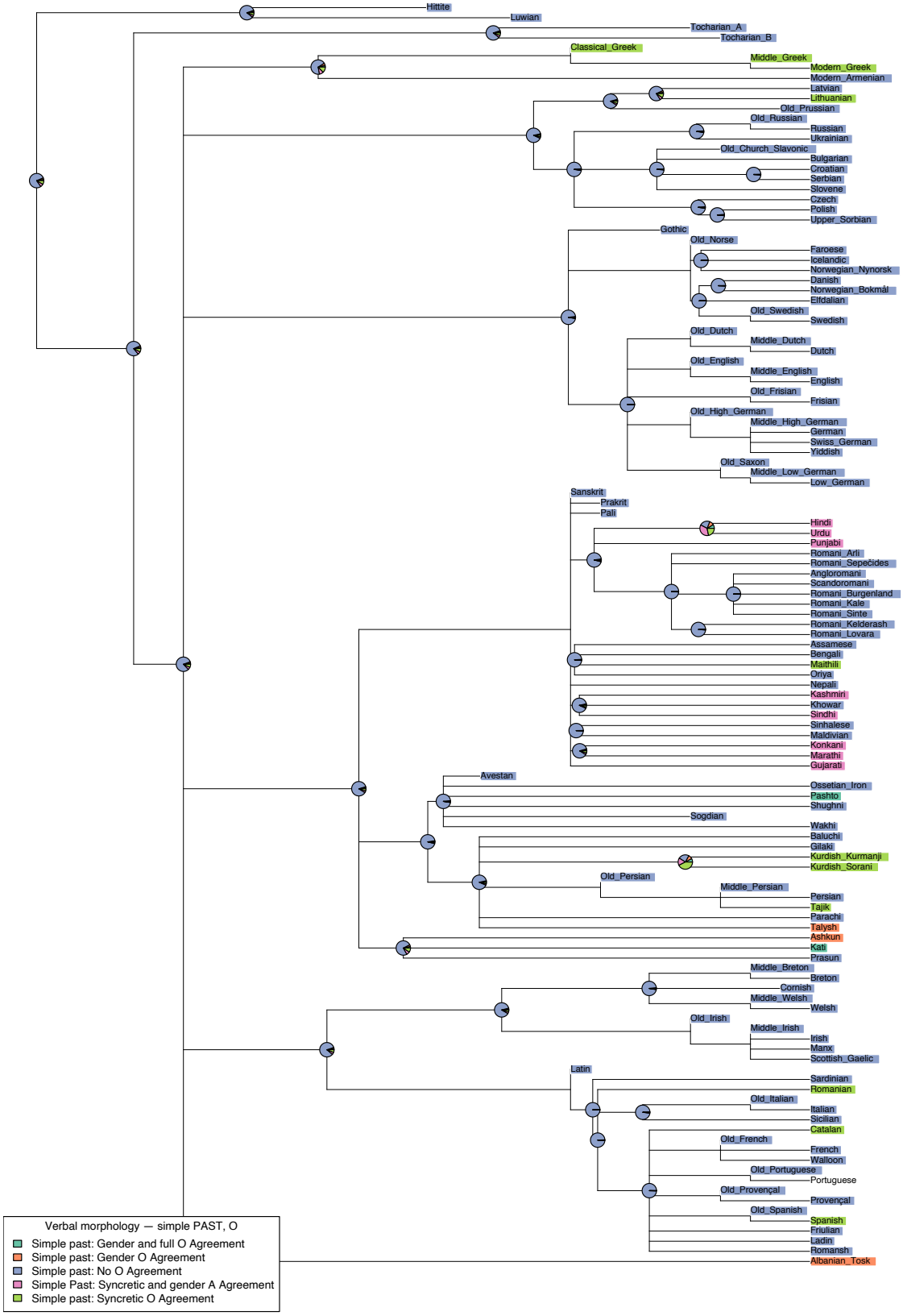

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
TENSE — Future
Future by perfect
No Future by perfect

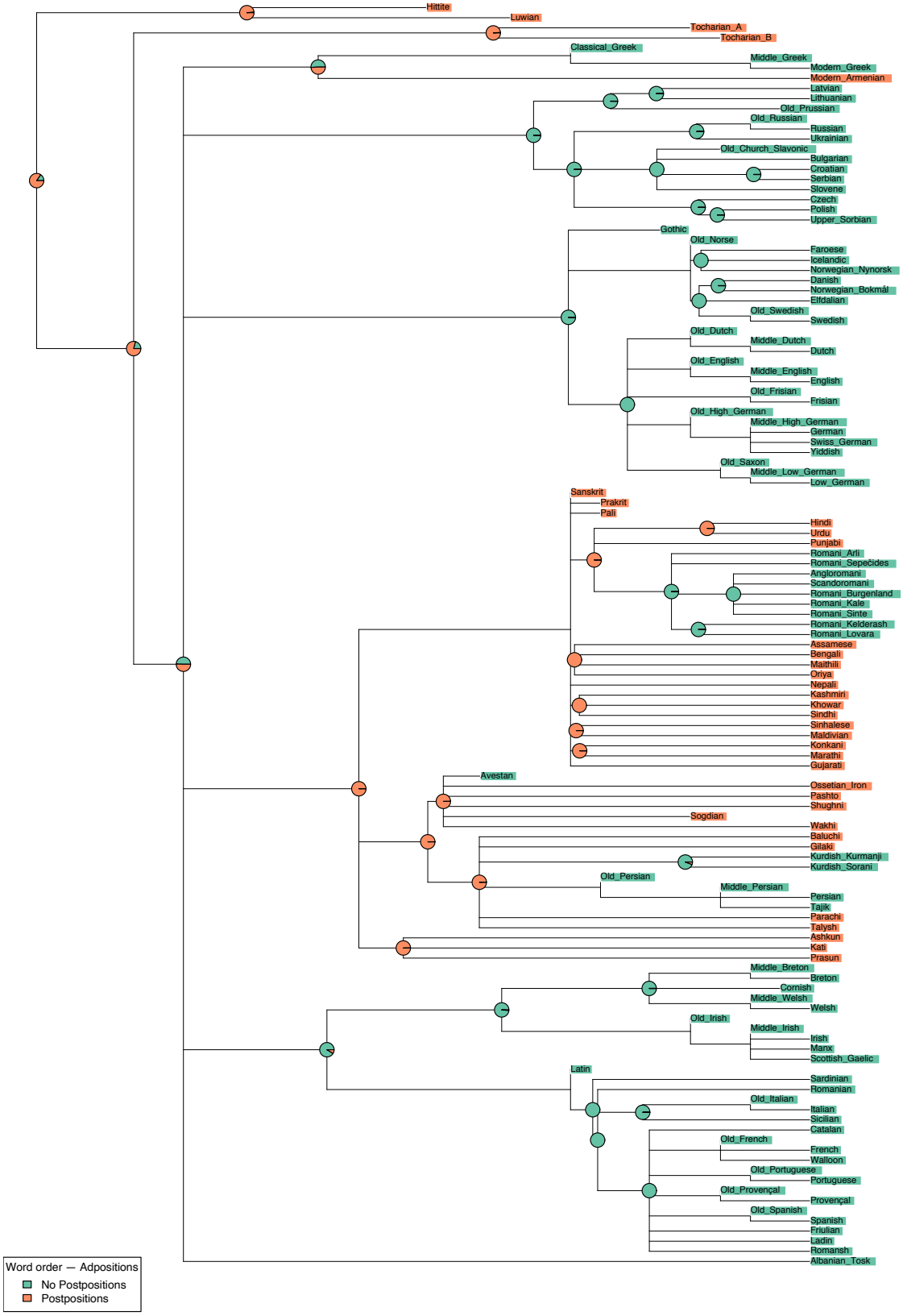

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Verbal morphology — present progressive, A
Present progressive: Full A Agreement
Present progressive: Full and Gender A Agreement
Present progressive: Gender A Agreement
Present progressive: No A Agreement
Present progressive: Syncretic A Agrement

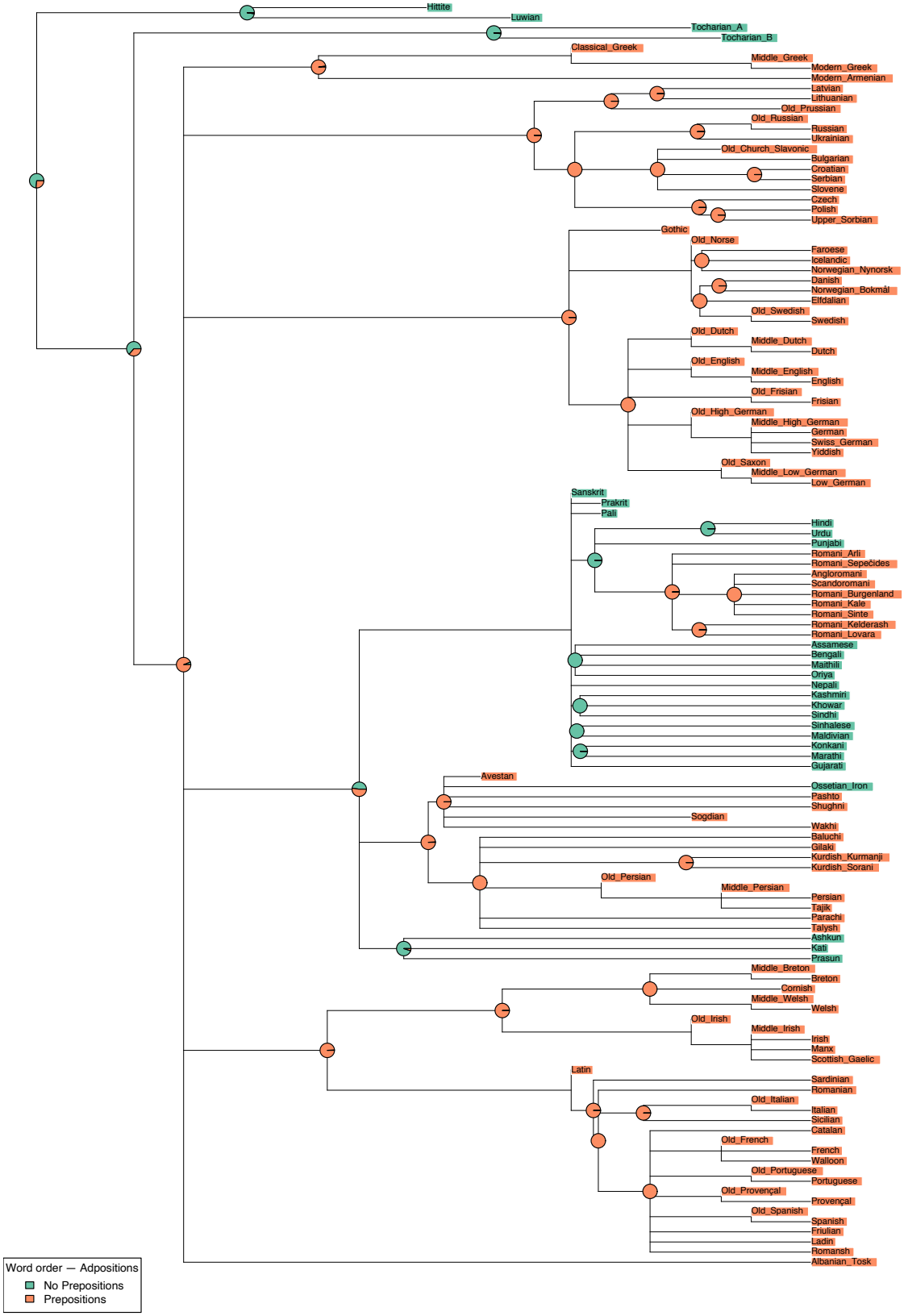

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Verbal morphology — present progressive, O
Present progressive: Full O Agreement
Present progressive: No O Agreement
Present progressive: Syncretic O Agreement

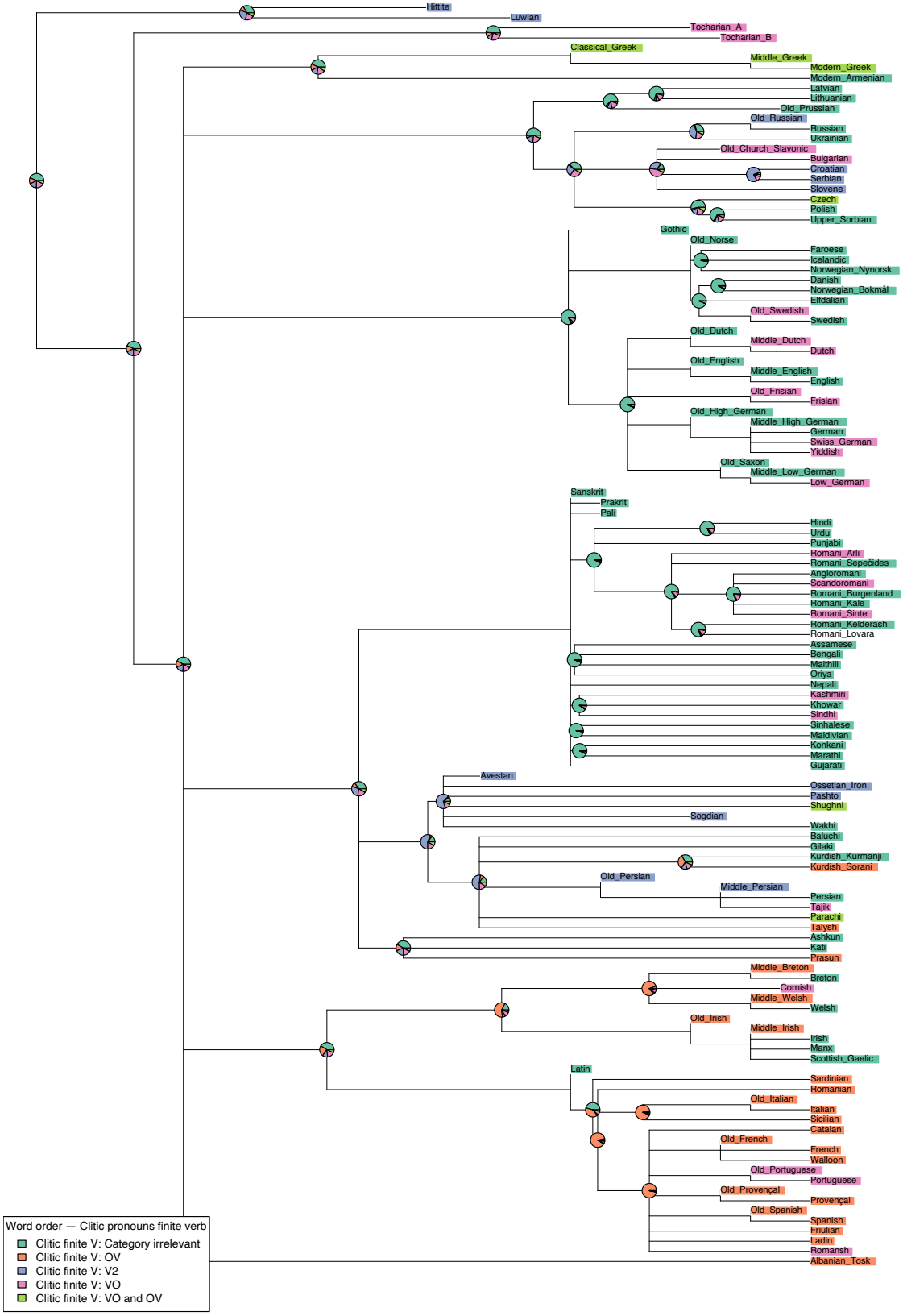

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Verbal morphology — simple PAST, A
Simple past: Full A Agreement
Simple past: Full and Gender A Agreement
Simple past: Gender A Agreement
Simple past: No A Agreement
Simple past: Syncretic A Agreement

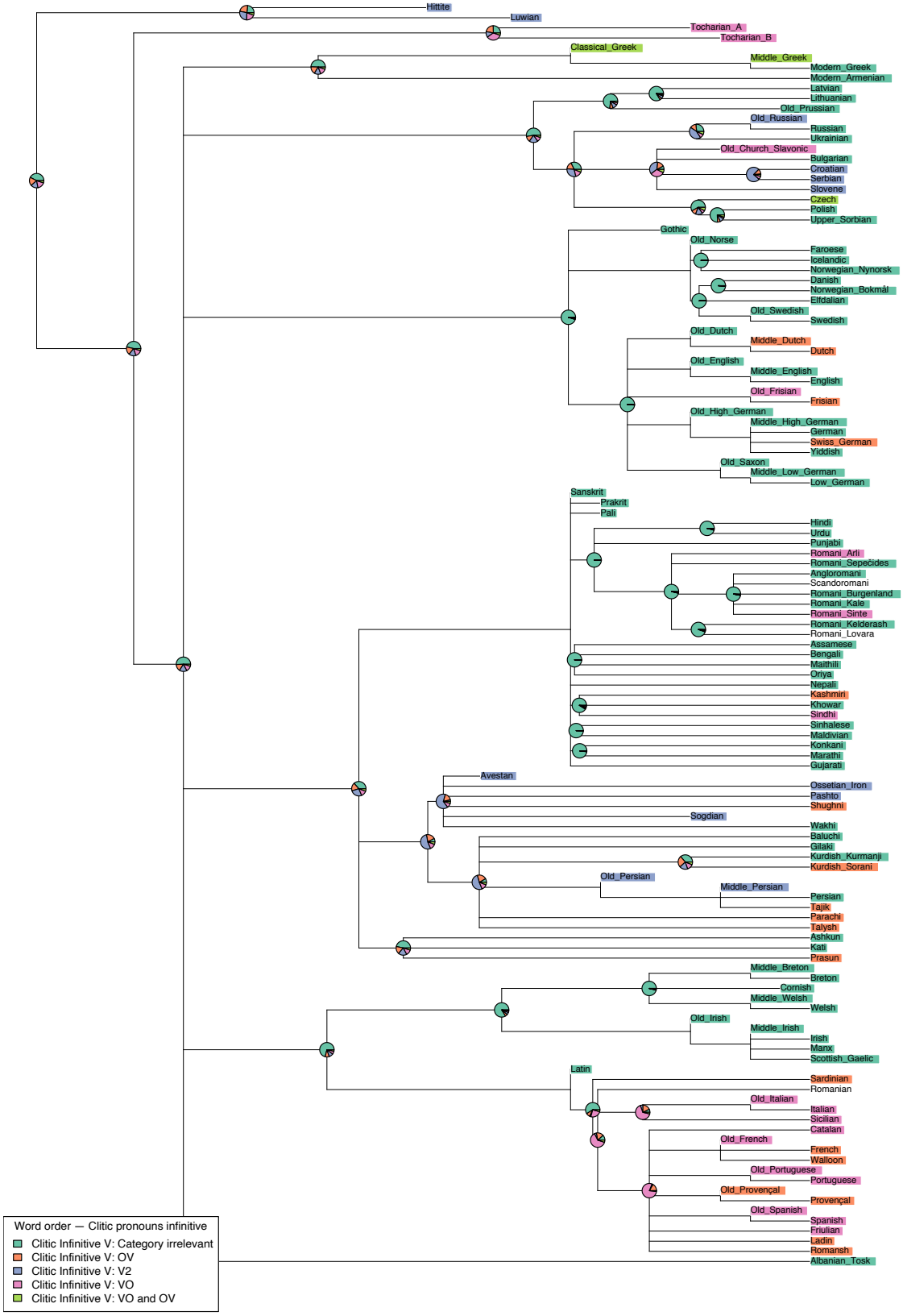

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Verbal morphology — simple PAST, DAT
Simple past: Full Dative Agreement
Simple past: Gender Dative Agreement
Simple past: No Dative Agreement
Simple past: Syncretic Dative Agreement

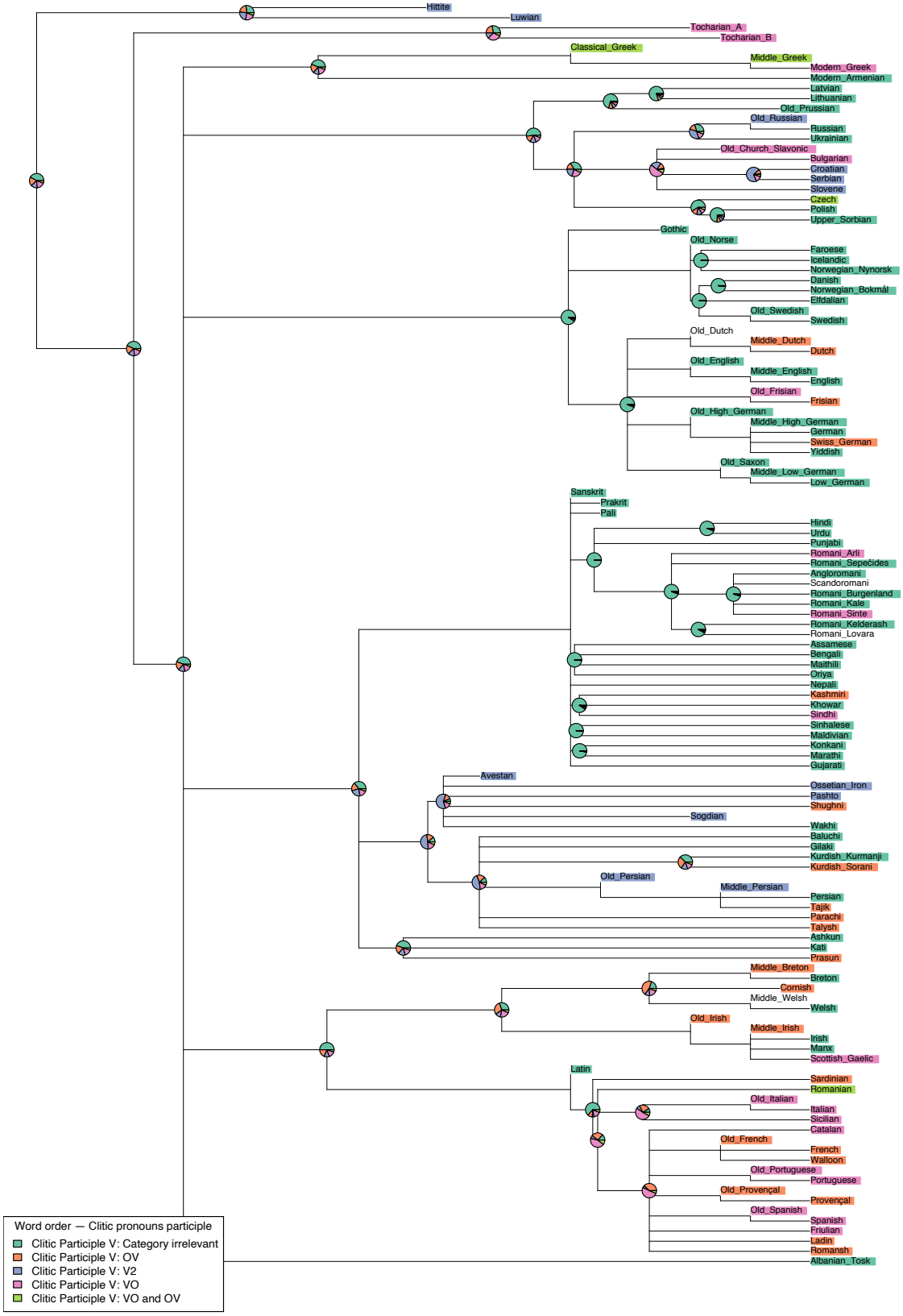

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Verbal morphology — simple PAST, O
Simple past: Gender and full O Agreement
Simple past: Gender O Agreement
Simple past: No O Agreement
Simple Past: Syncretic and gender A Agreement
Simple past: Syncretic O Agreement

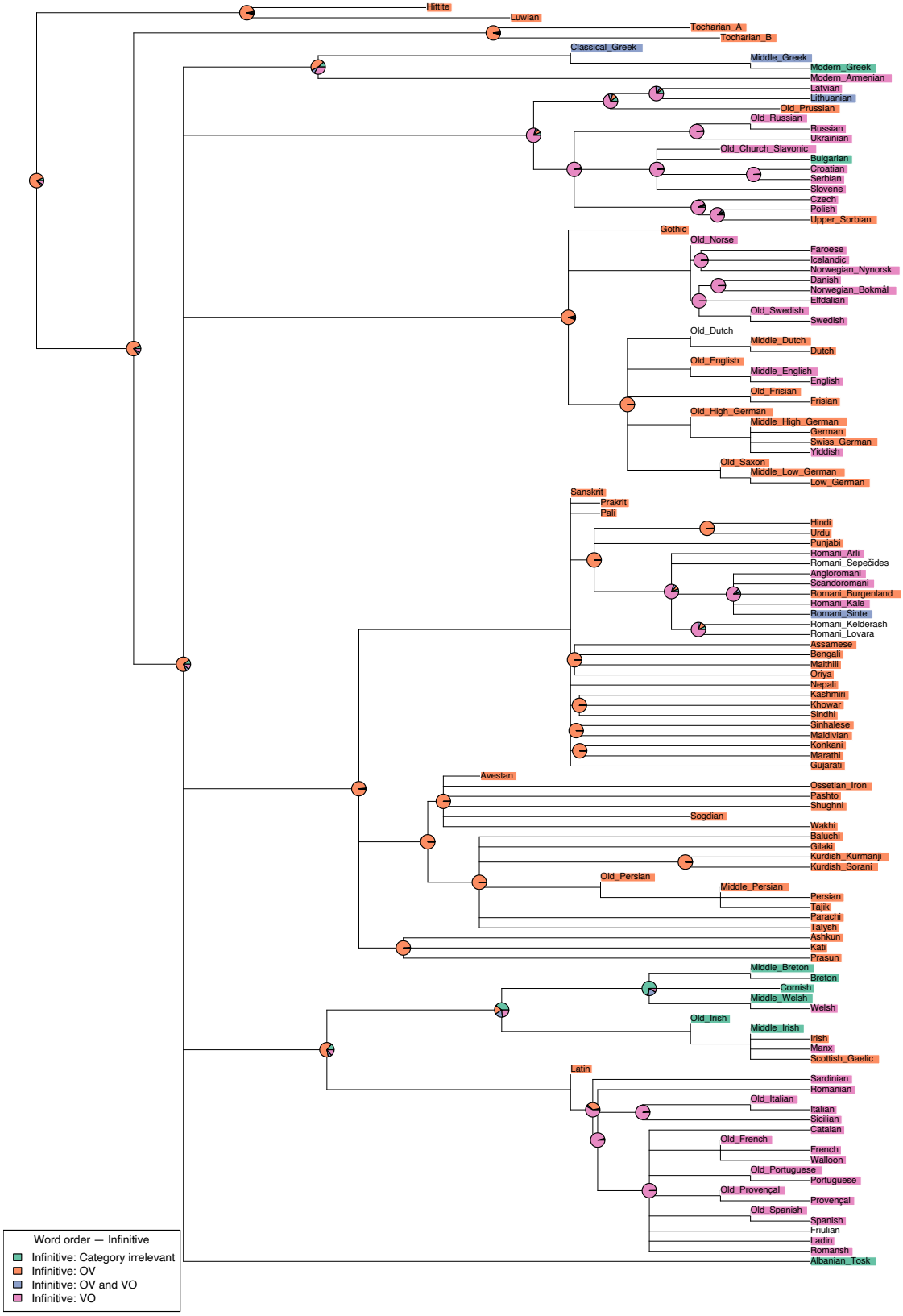

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talyshi
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Word order — Adpositions
No Postpositions
Postpositions

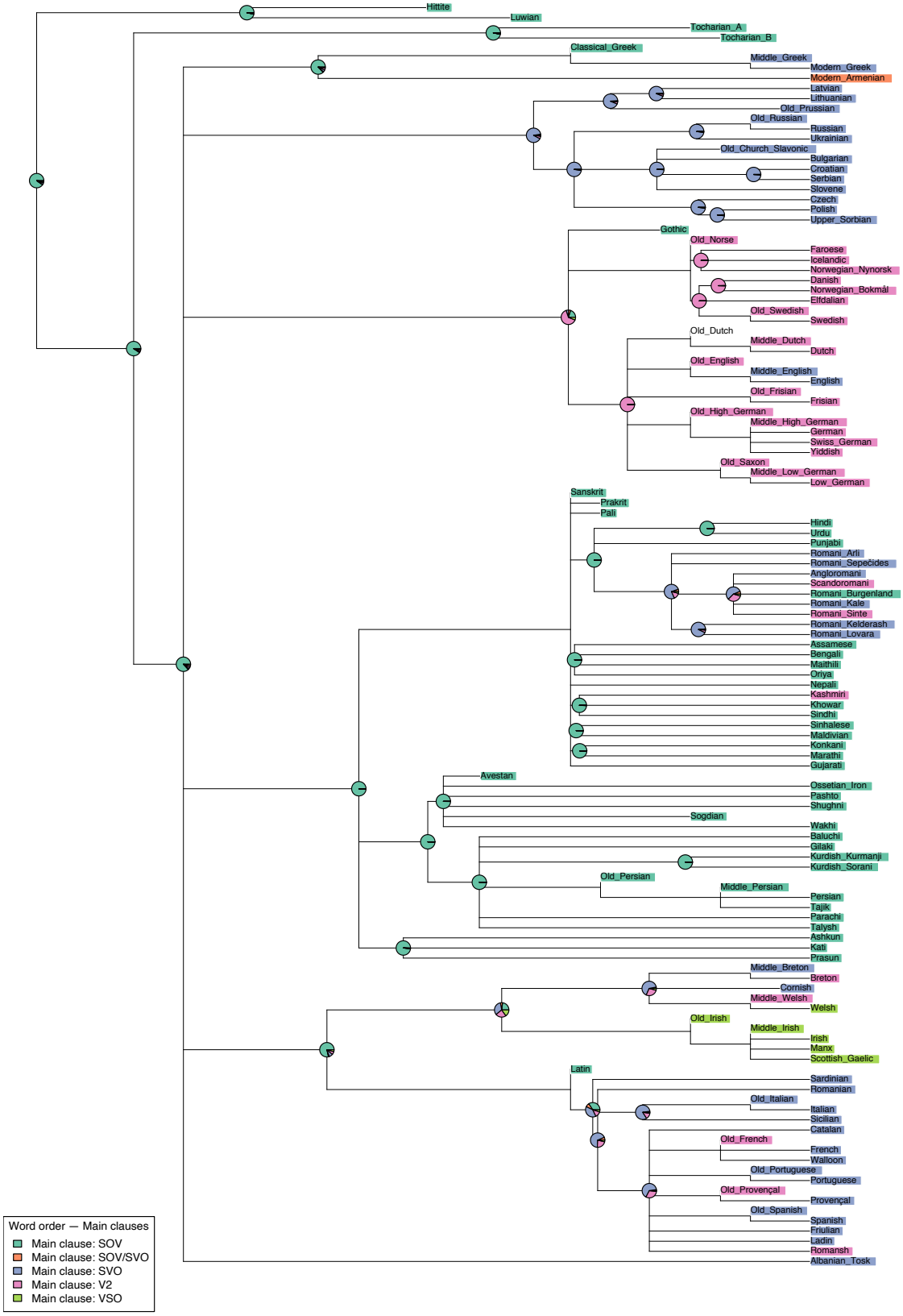

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Word order — Adpositions
No Prepositions
Prepositions

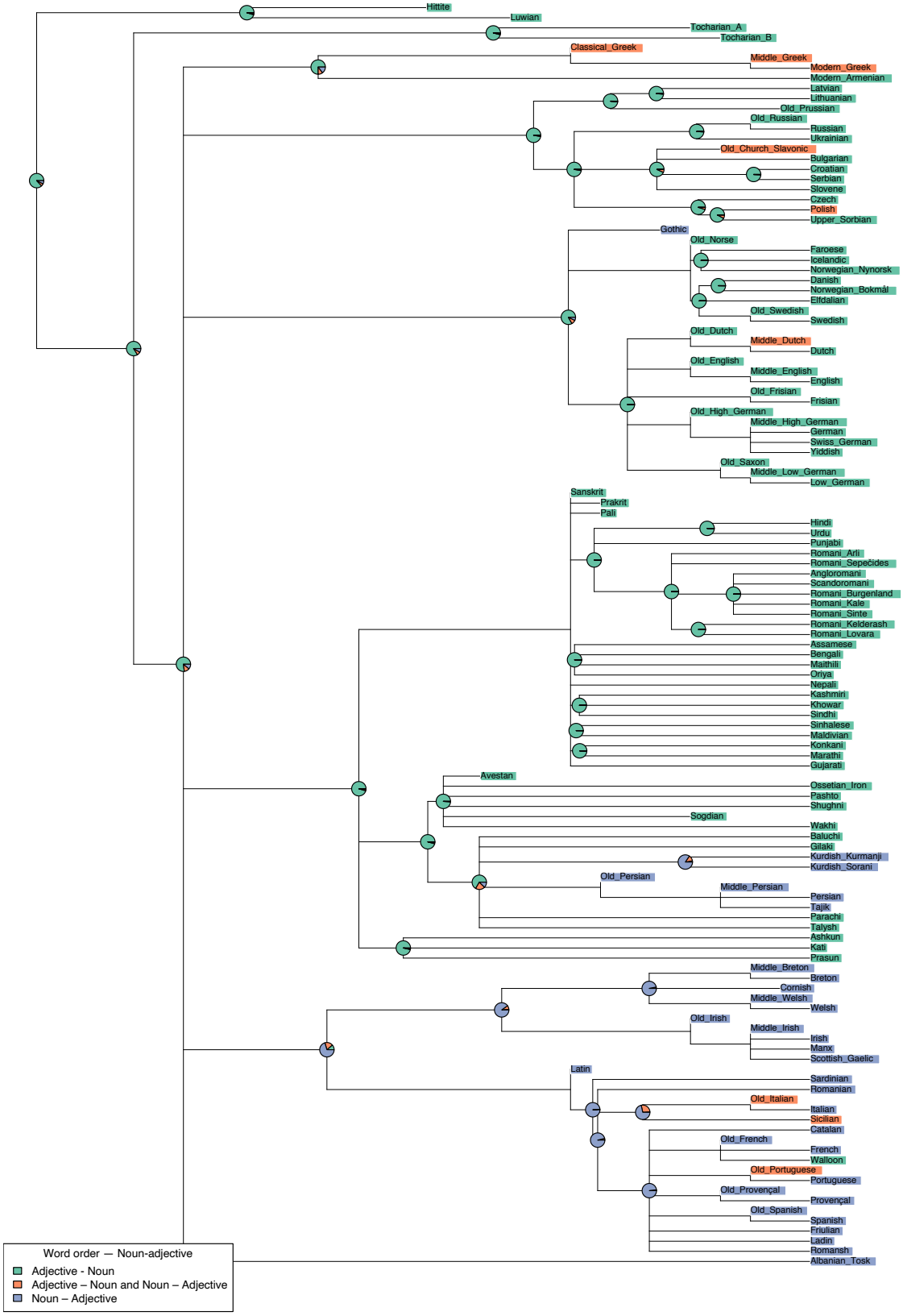

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Word order — Clitic pronouns finite verb
Clitic finite V: Category irrelevant
Clitic finite V: OV
Clitic finite V: V2
Clitic finite V: VO
Clitic finite V: VO and OV

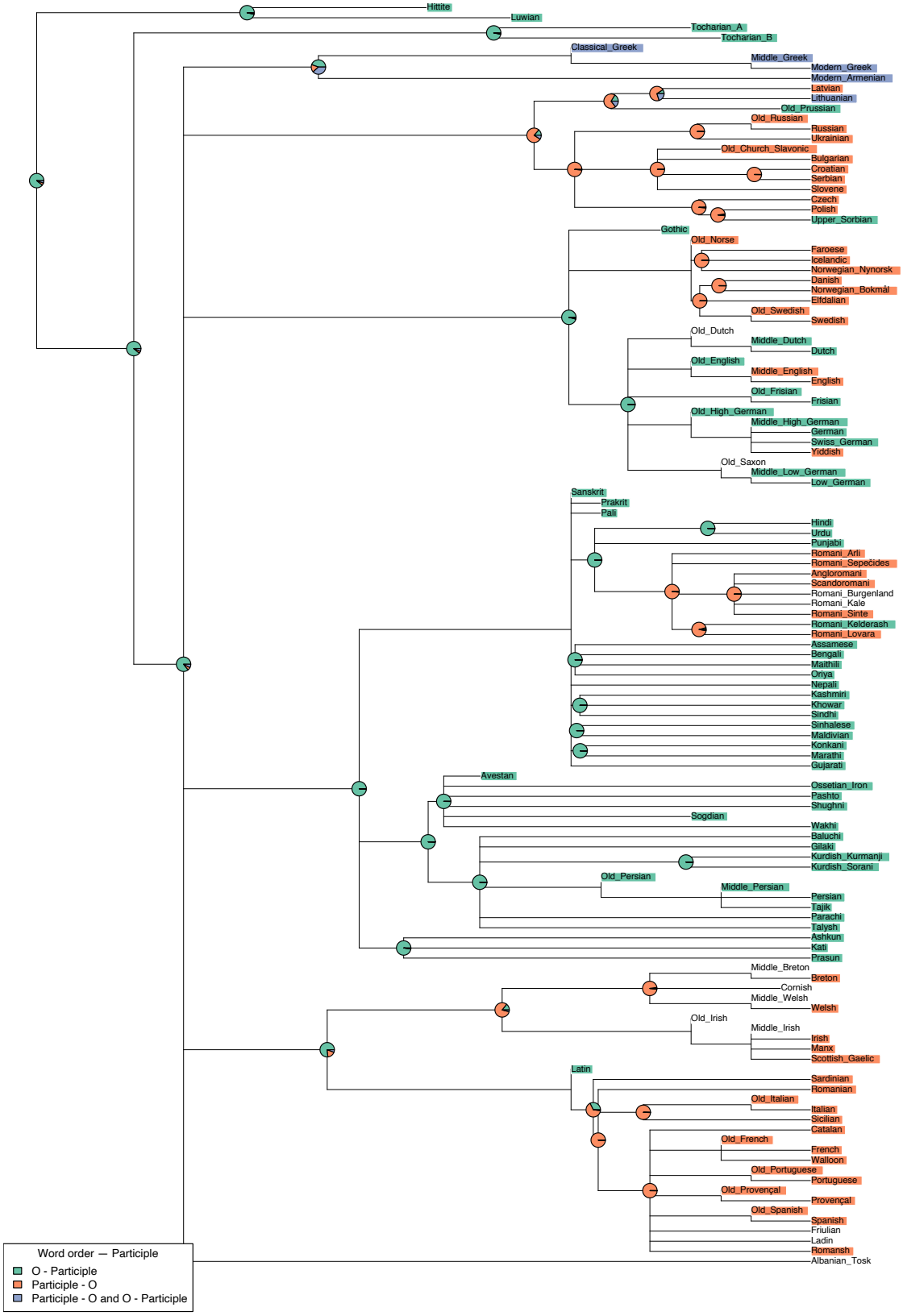

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Word order — Clitic pronouns infinitive
Clitic Infinitive V: Category irrelevant
Clitic Infinitive V: OV
Clitic Infinitive V: V2
Clitic Infinitive V: VO
Clitic Infinitive V: VO and OV

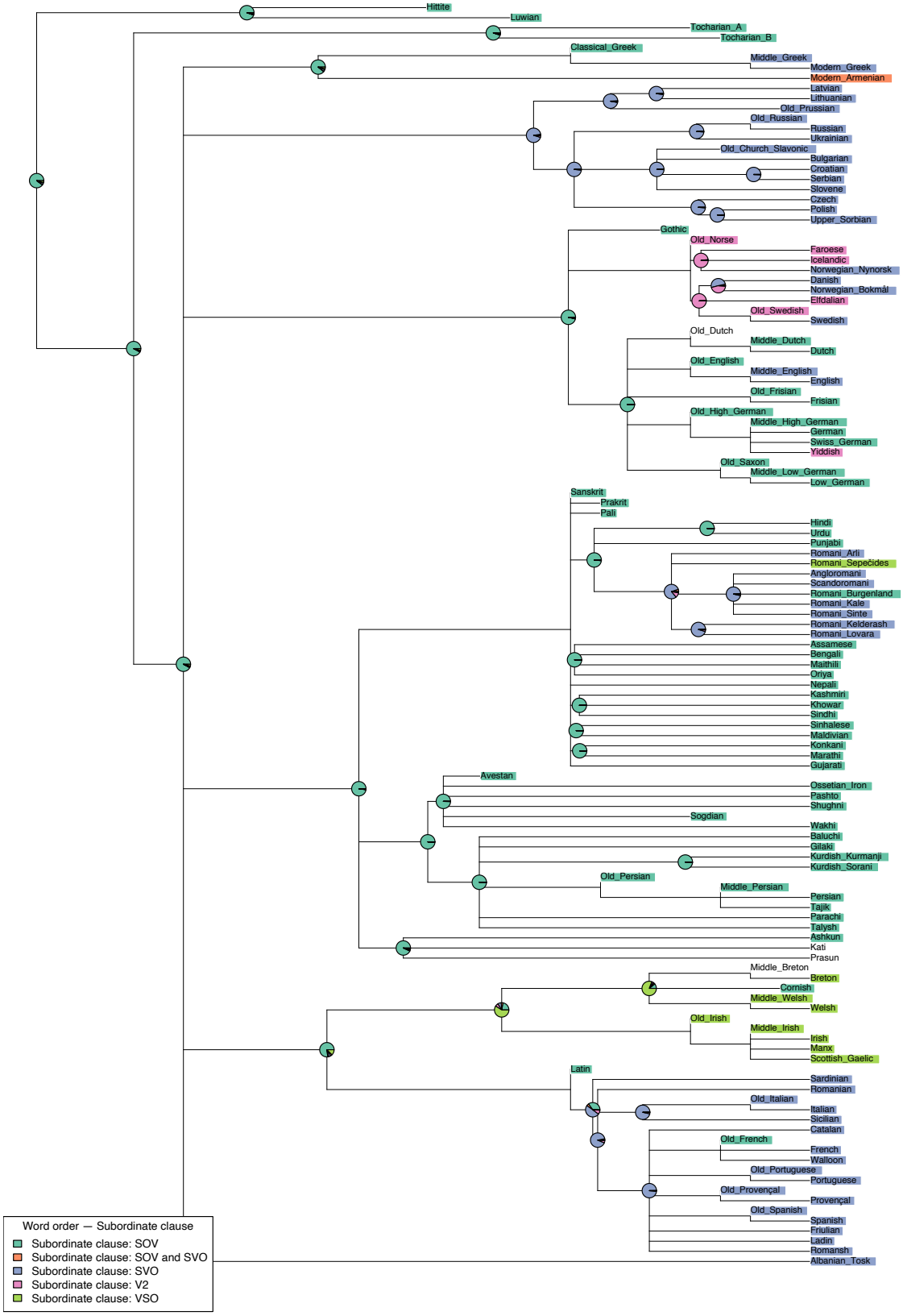

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Word order — Clitic pronouns participle
Clitic Participle V: Category irrelevant
Clitic Participle V: OV
Clitic Participle V: V2
Clitic Participle V: VO
Clitic Participle V: VO and OV

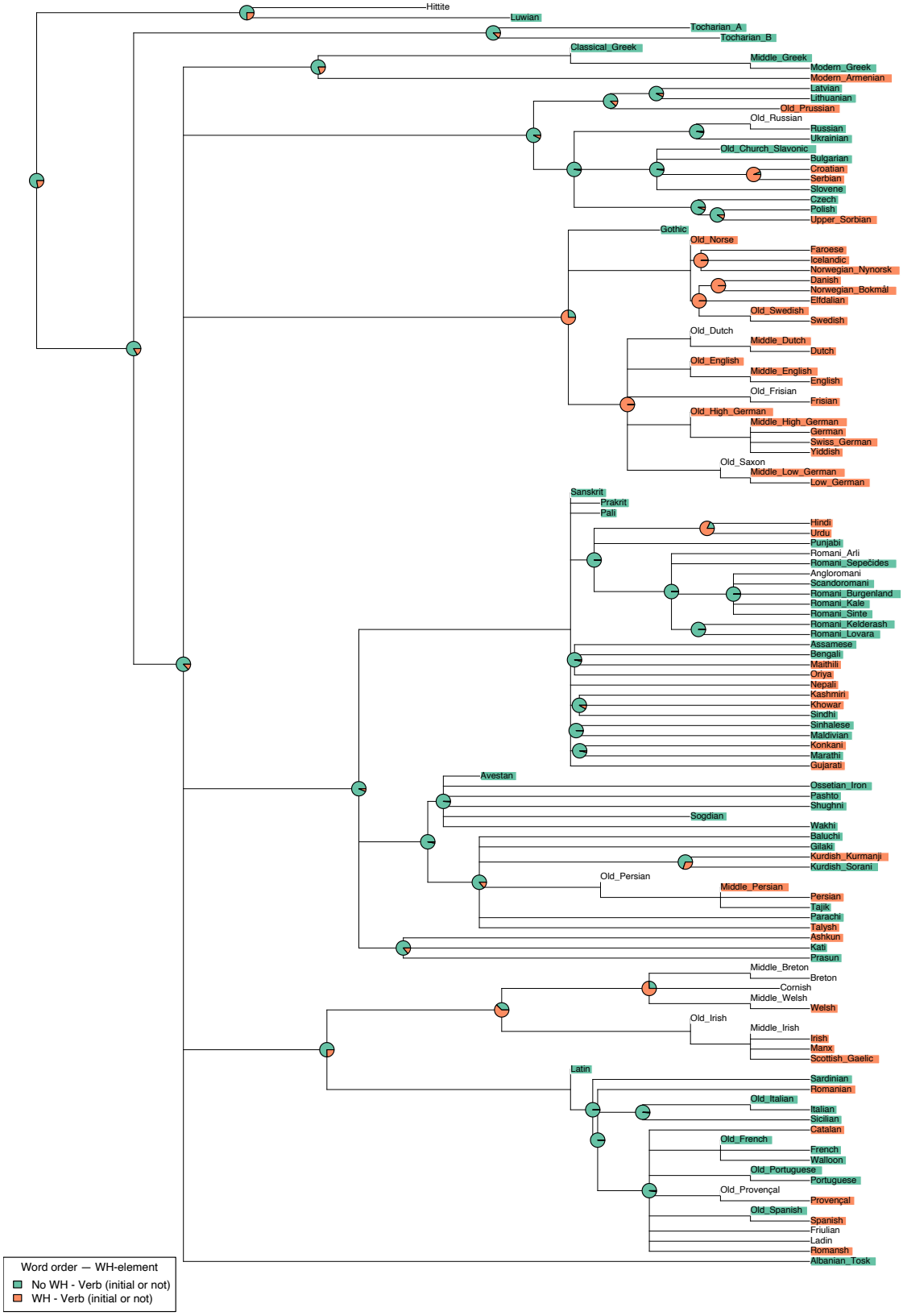

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Word order — Infinitive
Infinitive: Category irrelevant
Infinitive: OV
Infinitive: OV and VO
Infinitive: VO

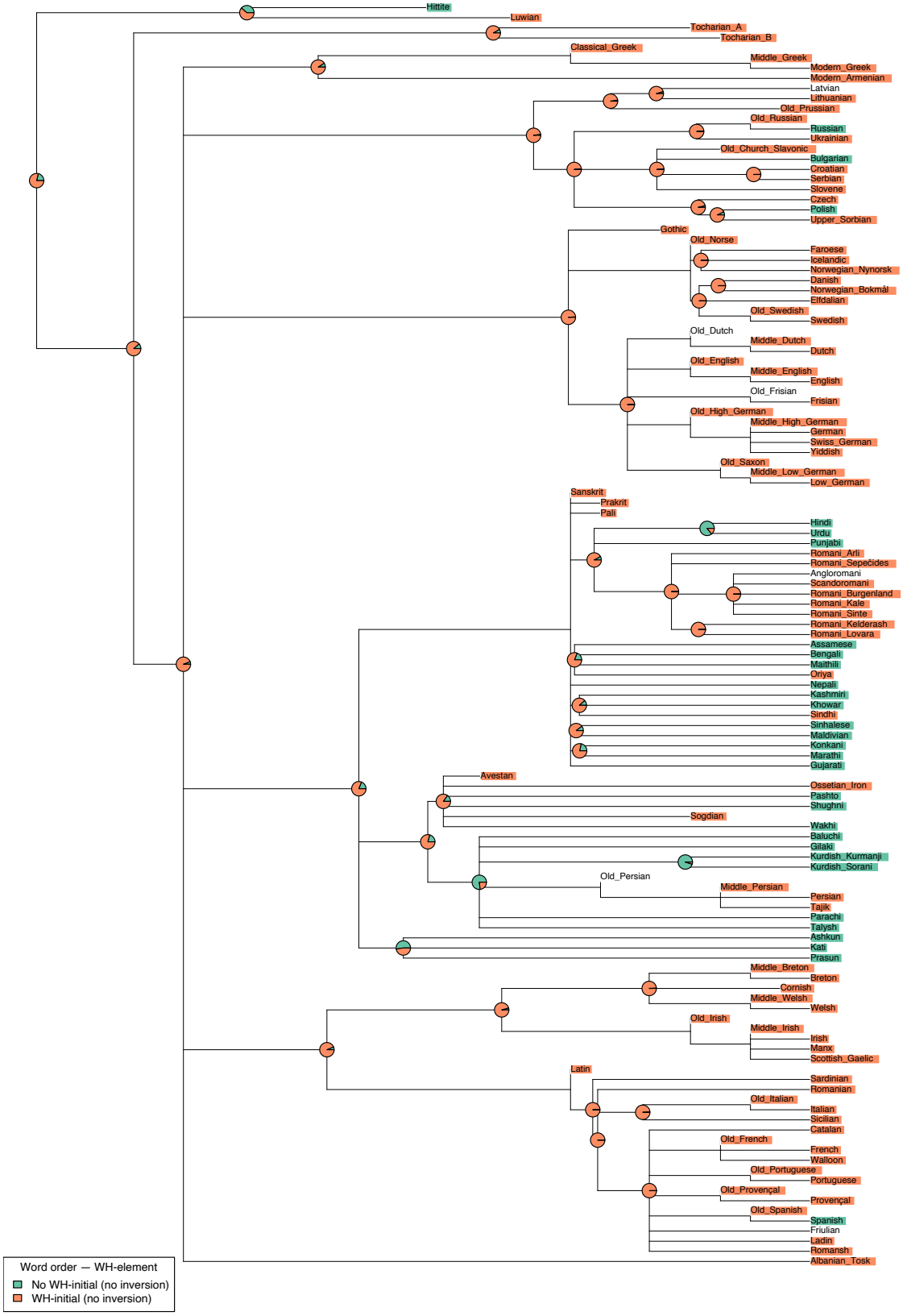

Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Word order — Main clauses
Main clause: SOV
Main clause: SOV/SVO
Main clause: SVO
Main clause: V2
Main clause: VSO


Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Word order — Noun-adjective
Adjective - Noun
Adjective – Noun and Noun – Adjective
Noun – Adjective


Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Word order — Participle
O - Participle
Participle - O
Participle - O and O - Participle


Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Word order — Subordinate clause
Subordinate clause: SOV
Subordinate clause: SOV and SVO
Subordinate clause: SVO
Subordinate clause: V2
Subordinate clause: VSO


Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Word order — WH-element
No WH - Verb (initial or not)
WH - Verb (initial or not)


Hittite
Luwian
Tocharian_A
Tocharian_B
Classical_Greek
Middle_Greek
Modern_Greek
Modern_Armenian
Latvian
Lithuanian
Old_Prussian
Old_Russian
Russian
Ukrainian
Old_Church_Slavonic
Bulgarian
Croatian
Serbian
Slovene
Czech
Polish
Upper_Sorbian
Gothic
Old_Norse
Faroese
Icelandic
Norwegian_Nynorsk
Danish
Norwegian_Bokmål
Elfdalian
Old_Swedish
Swedish
Old_Dutch
Middle_Dutch
Dutch
Old_English
Middle_English
English
Old_Frisian
Frisian
Old_High_German
Middle_High_German
German
Swiss_German
Yiddish
Old_Saxon
Middle_Low_German
Low_German
Sanskrit
Prakrit
Pali
Hindi
Urdu
Punjabi
Romani_Arli
Romani_Sepečides
Angloromani
Scandoromani
Romani_Burgenland
Romani_Kale
Romani_Sinte
Romani_Kelderash
Romani_Lovara
Assamese
Bengali
Maithili
Oriya
Nepali
Kashmiri
Khowar
Sindhi
Sinhalese
Maldivian
Konkani
Marathi
Gujarati
Avestan
Ossetian_Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Kurdish_Kurmanji
Kurdish_Sorani
Old_Persian
Middle_Persian
Persian
Tajik
Parachi
Talysh
Ashkun
Kati
Prasun
Middle_Breton
Breton
Cornish
Middle_Welsh
Welsh
Old_Irish
Middle_Irish
Irish
Manx
Scottish_Gaelic
Latin
Sardinian
Romanian
Old_Italian
Italian
Sicilian
Catalan
Old_French
French
Walloon
Old_Portuguese
Portuguese
Old_Provençal
Provençal
Old_Spanish
Spanish
Friulian
Ladin
Romansh
Albanian_Tosk
Word order — WH-element
No WH-initial (no inversion)
WH-initial (no inversion)

S9

Tree sample of Indo-European languages used for this paper's experiments.


Hittite
Luwian
Tocharian A
Tocharian B
Albanian Tosk
Modern Armenian
Classical Greek
Middle Greek
Modern Greek
Old Prussian
Latvian
Lithuanian
Ukrainian
Old Russian
Russian
Czech
Polish
Upper Sorbian
Old Church Slavonic
Bulgarian
Slovene
Croatian
Serbian
Gothic
Old Norse
Faroese
Icelandic
Norwegian Nynorsk
Elfdalian
Danish
Norwegian Bokmål
Old Swedish
Swedish
Old Frisian
Frisian
Old Dutch
Middle Dutch
Dutch
Old English
Middle English
English
Old Saxon
Middle Low German
Low German
Old High German
Middle High German
German
Swiss German
Yiddish
Cornish
Middle Breton
Breton
Middle Welsh
Welsh
Old Irish
Middle Irish
Irish
Manx
Scottish Gaelic
Latin
Sardinian
Romanian
Sicilian
Old Italian
Italian
Catalan
Friulian
Ladin
Romansh
Old Portuguese
Portuguese
Old Provençal
Provençal
Old Spanish
Spanish
Old French
French
Walloon
Ashkun
Kati
Prasun
Avestan
Ossetian Iron
Pashto
Shughni
Sogdian
Wakhi
Baluchi
Gilaki
Parachi
Talysh
Kurdish Kurmanji
Kurdish Sorani
Old Persian
Middle Persian
Persian
Tajik
Sanskrit
Prakrit
Pali
Nepali
Gujarati
Sinhalese
Maldivian
Konkani
Marathi
Kashmiri
Khowar
Sindhi
Assamese
Bengali
Maithili
Oriya
Punjabi
Hindi
Urdu
Romani Arli
Romani Sepečides
Romani Kelderash
Romani Lovara
Angloromani
Scandoromani
Romani Burgenland
Romani Kalé
Romani Sinte
7000
5000
3000
1000
0