

Divergence-time estimation in Indo-European

The case of Latin

David Goldstein

01 March 2024

Abstract

Divergence-time estimation is one of the most important endeavors in historical linguistics. Its importance is matched only by its difficulty. As Bayesian methods of divergence-time estimation have become more common over the past two decades, a number of critical issues have come to the fore, including model sensitivity, the dependence of root-age estimates on uncertain interior-node ages, and the relationship between ancient languages and their modern counterparts. This study addresses these issues in an investigation of a particularly fraught case within Indo-European, the diversification of Latin into the Romance languages. The results of this study support a gradualist account of their formation that most likely begins after 300 CE. They also bolster the view that Classical Latin is a sampled ancestor of the Romance languages (i.e., it lies along the branch leading to the Romance languages).

1 Introduction

The importance of divergence-time estimation for linguistic history cannot be overstated. Absolute chronologies of linguistic history are not only essential for estimating rates of change and diversification, but they are also crucial to our narratives of prehistory. To take one prominent example, divergence-time estimates of Proto-Indo-European play a central role in the debate over the Steppe and Anatolian Hypotheses (Gray and Atkinson [2003](#), Nicholls et al. [2008](#), Bouckaert et al. [2012](#), Chang et al. [2015](#)).

Archaeological evidence and linguistic palaeontology have traditionally been used to estimate divergence times. A classic example is the age and location of Proto-Oceanic, where both linguistic and archaeological evidence point to the Bismarck Archipelago in the late 2nd millennium BCE (Pawley 2007). Quantitative methods for estimating divergence times on the basis of linguistic data were developed in the second half of the 20th century within glottochronology, a now defunct intellectual program. More recently, Bayesian methods have come to play a dominant role in divergence-time estimation (e.g., Kitchen et al. 2009, Bouckaert et al. 2012, Bouckaert et al. 2013, Chang et al. 2015, Rama 2018, Ritchie et al. 2019, Sagart et al. 2019, Savelyev et al. 2020, Hartmann 2023). These methods are unquestionably the most sophisticated to date, but as their use has increased, three critical issues have come into focus.

1.1 Three issues

The first is the problem of model sensitivity (Bromham 2019, Heggarty 2021, Ringe 2022:60). Divergence-time estimates can be impacted not only by the choice of model, but even by its very parameterization (e.g., Gavryushkina et al. 2016, Warnock et al. 2020:24–25, Bromham 2022:7, Wright et al. 2022:2). In light of this situation, Heggarty (2021:382) rightly poses the question: “All are Bayesians now, perhaps, but can one just pick one’s assumption to get one’s desired answer?” I argue for a resounding “no” and show that with increased model exploration, analyses of model sensitivity, and tests of model adequacy, the problems that motivate Heggarty’s question can be mitigated at least to an extent.

The second issue is that divergence-time studies have focused overwhelmingly on root ages to the neglect of more recent events. While this focus on the root age is understandable given the urge to peer into the deepest past (e.g., Nichols 1994, Baxter et al. 2000, Greenberg 2002), one must bear in mind that estimates of root ages usually depend on constraints on the ages of interior clades (e.g., Gray and Jordan 2000, Gray and Atkinson 2003, Bouckaert et al. 2012, Bouckaert et al. 2013, Chang et al. 2015, Sagart et al. 2019, Koile et al. 2022). Such constraints typically take the form of a calibration density, but specifying a calibration density can be anything but straightforward, especially when the age of a clade is uncertain or disputed. More work at the level of interior clades is therefore necessary to put divergence-time estimates on more secure ground.

The final problem is perhaps the most contentious and concerns the relation-

ship between ancient languages (such as Latin, Old English, and Sanskrit) and extant languages. There is ongoing debate as to which (if any) ancient languages should be modeled as sampled ancestors, that is, as direct ancestors, of extant languages.¹ The ancestry question is of paramount importance because it can impact divergence-time estimates. For instance, Chang et al. (2015) impose ancestry constraints on eight ancient Indo-European languages to force them to be ancestral to extant languages. These constraints had a crucial effect on their divergence-time estimates, as it was the first Bayesian phylogenetic study to yield ages of Proto-Indo-European consistent with the Steppe Hypothesis (Rama 2018 subsequently found further support for this hypothesis with certain datasets and tree models). The use of such constraints has been contentious, however. More work at the level of the clade is required to establish which (if any) ancient languages should be modeled as direct ancestors of extant languages.

1.2 The diversification of Latin

The present study addresses these issues in an investigation of the diversification of Latin and fourteen Romance languages, whose geographic distribution is presented in Figure 1. I focus in particular on the following three debates. First, when does Latin begin to diversify into what will become the Romance languages? This is a question with a long history of research from the perspective of both Latin and Romance (e.g., Hall 1974, Dardel 1985, Adams 2007). It is intimately related to the second question: to what extent did the demise of the Western Roman Empire cause or accelerate the diversification of Latin? It has been claimed that catastrophic social events can accelerate linguistic change (see, e.g., Garrett 2006:142, Trudgill 2011:9–13) and some argue that the social fragmentation that followed the dissolution of the Western Roman Empire catalyzed the formation of the Romance languages (e.g., Varvaro 1991:48). Others contend that the diversification of Latin was a more gradual process that was already underway by the time the Western Empire fell. Finally, is Classical Latin a direct ancestor of the Romance

¹The term *sampled ancestor* is not used in historical linguistics. As Gavryushkina et al. (2016:59–60) explain, there are two types of sampling: fossil sampling and extant sampling. Suppose that a language is recorded at some point in the past. The discovery of the language and its representation in a dataset constitutes a fossil sampling event. When the properties of a contemporary language are encoded in a dataset, that is an example of extant sampling. When a sampled fossil belongs to a lineage from which another fossil or extant language is sampled, it is a sampled ancestor. So if Latin belongs to the ancestral language from which the Romance languages were sampled, it is a sampled ancestor.

languages? It is widely believed that the ancestor of the Romance languages is not written Classical Latin but rather the colloquial language or “Vulgar Latin.” So the question then arises of what the exact phylogenetic relationship between the written language and Vulgar Latin is. Almost fifty years ago, Hall (1976:9) noted the lack of consensus on this issue:

All competent scholars agree that the Romance languages have arisen, over something more than two millennia, by gradual differentiation from a common source, which was closely related to, but not identical with, the Latin used by authors from Plautus’ time to that of, say, Tertullian. Beyond this, however, there is little agreement concerning either the name to be given to this common source of the Romance languages, its chronology, or its exact relation to attested Latin.

Opinion on this final point remains sharply divided (e.g., Chang et al. 2015:206–207, Cathcart et al. 2018:11, Garrett 2018:33–35, Heggarty 2021:381).



Figure 1: Geographic distribution of the language sample

1.3 Main claims

The results of my study support the following answers to the questions above. First, the lexical diversification of Latin in all likelihood begins after 300 CE. Second, the dissolution of the Western Roman Empire did not accelerate the rate of diversification. Finally, the posterior probability that Classical Latin is a sampled ancestor of the Romance languages ranges from approximately 70% to 90% among the best performing models. (This is not to say, however, that Classical Latin is the most recent common ancestor of the Romance languages, i.e., Proto-Romance.) In contrast to previous phylogenetic studies of the Romance languages, the results of this study were obtained without the use of clade constraints, ancestry constraints, or node calibrations.

Since the central focus of this paper is divergence-time estimation, it will be helpful to say a word at the outset about my interpretation of divergence times. The view adopted here is purely mechanical. The methods used in this study estimate the average number of changes on a branch and then propose an age for that branch based on the rate of change from the clock model. This mechanical interpretation makes no commitment about the relationship between divergence times and population movement, the textual record, or the mutual intelligibility of the diversifying dialects.

1.4 Outline

The remainder of this paper is structured as follows. Section 2 provides an overview of traditional and computational work on the topology, divergence times, and ancestry of the Romance languages. Sections 3 and 4 introduce the methods and data used in this study, respectively. The results are presented in Section 5, which are then discussed in Section 6. Section 7 brings the paper to a close with brief summary remarks.

2 Previous research

2.1 Tree topology

Although the Romance languages are among the best investigated in the world, their tree topology is not well understood (Posner 1996:196, Stefenelli 1996, Hoinkes 2003:134). This situation may be due to some extent to a lack of in-

terest on the part of Romance linguists in tree graphs as models of linguistic history. In a recent handbook chapter on the formation of the Romance languages, Varvaro (2013:11) presents only one phylogenetic tree of Romance—from Schuchardt (1866:82)! In an article devoted to the transition from Latin to the Romance languages, Banniard (2013) says nothing about the phylogeny of Romance.

To the extent that there is consensus on the topology of Romance, it has emerged from the work of Robert Hall (Posner 1996:196, Klinkenberg 1999:136). Hall (1950:24) contends that applying the Comparative Method to the Romance languages results in a tree in which Southern Romance (Sardinian, Lucanian, Sicilian) split off first, followed by Eastern Romance (the dialects of Romanian), and finally Italo-Western Romance. A number of other scholars now subscribe to this view. A composite and slightly modified version of the trees in Hall 1976:14, 16 is presented in Figure 2. The evidence for this hypothesis comes primarily from developments among stressed vowels (e.g., Straka 1953, Straka 1956, Herman 2000:27–34).

The view that Sardinian and Romanian were the first clades to form is now widespread (e.g., Meier 1940:183, Straka 1956, Hall 1974:14–15, Hall 1976:16, Leonard 1980:39, Dardel 1985, Vallejo 2012:454–455, Buchi, González Martín, et al. 2015, Dworkin 2016a:12, 14, Weiss 2020:543) and has been adopted by the *Dictionnaire Étymologique Roman (DÉRom)* (Swiggers 2001). Not all Romanists subscribe to this tree, however. For instance, Agard (1984:64, 250–251) argues for a polytomy in which Proto-Romance splits into Southern, Eastern, and Italo-Western Romance. He does not identify his polytomy as hard or soft. (A hard polytomy is a true divergence of more than two languages at the same time, whereas a soft polytomy represents uncertainty in the tree topology.)



Figure 2: The phylogeny of Romance adapted from Hall 1974:14, 16. (Polytomies appear to represent dialect continua.) ⁷

2.1.1 Results from Bayesian phylogenetic studies

The topological results from Bayesian phylogenetic studies vary. In Gray and Atkinson 2003:437, Sardinian forms first, but is followed by Italian and only then Romanian (which strangely forms a clade with Ladin). In Bouckaert et al. 2013, Romanian forms first, followed in turn by Sardinian (Figure S1 in the Supplementary Material). The results of Rama (2018:199), which are based on a uniform time tree model (Ronquist et al. 2012), are similar in that Romanian is the first clade to form, followed by Sardinian. By contrast, in his fossilized birth-death (FBD) analyses, Sardinian forms first. Jäger (2019:170) infers a phylogeny of Romance from an impressive stock of Romance dialects, but his results bear little resemblance to the traditional topology (e.g., neither Romanian nor Sardinian are the first clades to form). He comments that “[t]hese results indicate that the data only contain a weak tree-like signal. This is unsurprising since the Romance languages and dialects form a dialect continuum where horizontal transfer of innovations is an important factor.”

2.2 Divergence times

Väänänen (1983) distinguishes two primary hypotheses regarding the timing of the diversification of Latin, which he refers to as the *thèse différencielle* (Väänänen 1983:490–494) and *thèse unitaire* (Väänänen 1983:486–490). I will refer to them as the *early hypothesis* and *late hypothesis*, respectively.

The early hypothesis, which has found favor especially among Romanists, is based mainly on evidence from the relative chronology of sound change and places the diversification of Latin in the first centuries CE (e.g., Selig 2008:16, Vallejo 2012:459, 462, Buchi, González Martín, et al. 2015:250–252). Not all Romanists agree on this time frame, however. Hall (1950:19), for instance, argues that Proto-Romance should be dated to the time of “the late Republic and the early Empire,” if not earlier (he suggests 250–200 BCE), and Hall (1974:17) dates Proto-Romance to between 100 BCE and 0 CE. Classicists, by contrast, tend to subscribe to the late hypothesis. Löfstedt (1959), for instance, argues that diversification begins only after 600 CE and Muller (1921) around 800 CE (which Väänänen 1983:488 considers an extreme position). Support for the late hypothesis comes primarily from the uniformity of written Latin.

Two pieces of historical evidence provide a *terminus ante quem* for the diversification of Latin. In 712 CE, Saint Boniface of England reports difficulty understanding

the spoken Latin of Pope Gregory II during an audience in Rome. The Council of Tours in 813 CE resolved that sermons be delivered in the vernacular because Latin was no longer understood by the laity (Cano González 2007:88). In interpreting these dates, it is crucial to bear in mind that mutual unintelligibility and linguistic speciation are distinct phenomena, which in all likelihood do not emerge simultaneously. In the case of the Romance languages, mutual unintelligibility may have developed only centuries after dialect formation (Palmer 1954:180, Löfstedt 1959:4).

2.2.1 Results from Bayesian phylogenetic studies

The results of computational investigations of Romance divergence times vary. For instance, Proto-Romance is dated to 1,700 BP (= 250 CE) in Gray and Atkinson 2003 and to 1,989.39 BP (= 39.39 BCE) in Bouckaert et al. 2012. The inferred ages are the product of node calibrations in both cases. The analysis of Chang et al. (2015) returned a much later date of around 1000 CE for Proto-Romance. The ages of Romance from the majority-rule consensus trees of the FBD analyses of Rama (2018) are presented in Table 1. These estimates are particularly interesting because they are not the result of node calibrations (Rama 2018:190). The dates from the Broad, Chang B1, and Chang B2 datasets are consistent with the early hypothesis, but interpreting the results solely on the basis of a point estimate is difficult. It is worth noting that Rama (2018:192) uses a single clock model (the Independent Gamma Rate model) for all his analyses. It is unclear how well this model performs compared to others (such as the ones described in section 3.1 below).

Dataset	Proto-Latin-Romance	Proto-Romance
Narrow	585.9155 BCE	383.3557 CE
Medium	631.092 BCE	324.6624 CE
Broad	654.7365 BCE	228.7998 CE
Chang B1	702.475 BCE	160.6884 CE
Chang B2	705.7255 BCE	127.5228 CE

Table 1: Inferred ages from the FBD analyses of Rama 2018

2.3 The ancestry question

As noted above in Section 1.1, questions of linguistic ancestry have played a central role in the debates over divergence times within Indo-European. Even in cases where an ancestry relationship may seem *prima facie* straightforward, it often is not. Consider for instance Old English and modern Standard English. Our Old English texts are primarily written in the Wessex dialect, but modern Standard English descends from the Mercian dialect (Janda et al. 2003:19, Finegan 2009:65). Frustratingly, the Mercian dialect is poorly attested during the Old English period and the Wessex dialect is only sparsely attested after the Norman Conquest. Although later English texts do not directly descend from Old English, it could be the case that Old English was only minimally different from the true ancestor of later English, in which case it could be considered ancestral to later English for all practical purposes.

The relationship between Classical Latin and the Romance languages has emerged as one of the more contentious cases of linguistic ancestry (e.g., Chang et al. 2015:206–208, Heggarty 2021:381). It is surprising that, of all corpus languages, Latin has engendered so much debate. For over five hundred years, it has been widely maintained that Classical Latin is not a direct ancestor of the Romance languages (e.g., Mańczak 1994b:365, 368, Mańczak 1994a:17, Eskhult 2018:203). Heggarty (2021:381) considers it a matter of linguistic orthodoxy that “High-status ancient written languages are almost by definition *not* the direct sources of the modern *spoken* languages” (emphasis in the original). There are two counterpoints to bear in mind, however. First, the view that Vulgar Latin is the ancestor of the Romance languages crystallized in an era when we had far less evidence for Vulgar Latin than we do today (e.g., in the form of papyri and inscriptional remains). Second, it has not always been clear what exactly “direct ancestry” is supposed to mean.

2.3.1 What is direct ancestry?

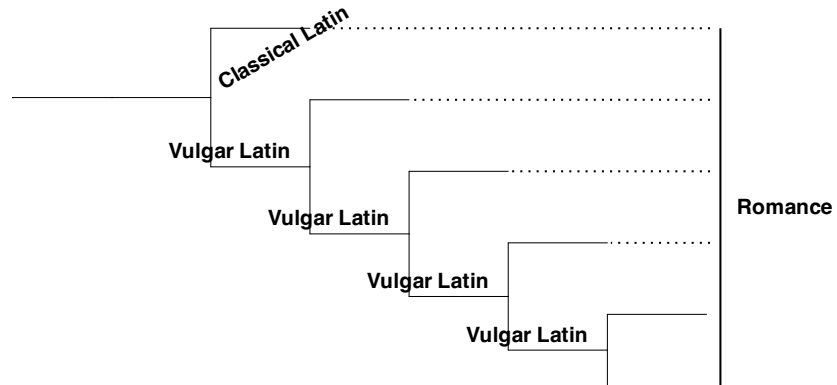
To elaborate on this latter point, consider the three hypotheses in Figure 3. In the first, Classical Latin is the most recent common ancestor (MRCA) of the Romance languages; i.e., it is tantamount to Proto-Romance, and each branch leading to a Romance language is a distinct form of Vulgar Latin. In the second, Proto-Romance is the MRCA of the Romance languages and Latin is a sampled ancestor that lies on the branch leading to the Romance languages. Note that

Classical Latin and Vulgar Latin are not distinct taxa under this hypothesis.² In the third, Classical and Vulgar Latin are siblings, with the latter a sampled ancestor of the Romance languages.

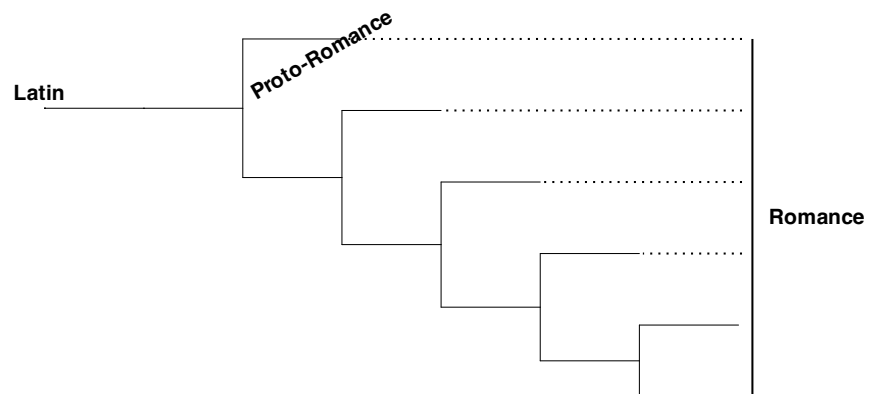
What one considers a direct ancestor will depend on how exactly direct ancestry is defined. One option is to equate it with the most recent common ancestor. On this approach, Classical Latin is the direct ancestor of the Romance languages in the first tree of Figure 3, but not in the second or third. Another possibility is to equate direct ancestry with a position on the branch leading to the most recent common ancestor. Under this definition, Classical Latin is the direct ancestor of the Romance languages in the first tree, Latin (both Vulgar and Classical) in the second, and Vulgar Latin in the third. To answer the question of whether or not Latin (Classical or Vulgar) is the direct ancestor of the Romance languages, direct ancestry must be explicitly defined.

²A reviewer questions whether there is any meaningful difference between the MRCA and sampled-ancestor hypotheses. He argues that since the former is a special case of the latter, distinguishing them only fosters confusion. Although the MRCA hypothesis is indeed a special case of the sampled-ancestor hypothesis, they nevertheless make different predictions. Under the MRCA hypothesis, ancestral Proto-Romance forms inferred on the basis of the comparative method are predicted to be identical to Classical Latin forms. Under the sampled-ancestor hypothesis, they need not be.

The most-recent common ancestor (MRCA) hypothesis



The sampled-ancestor hypothesis



The sibling hypothesis

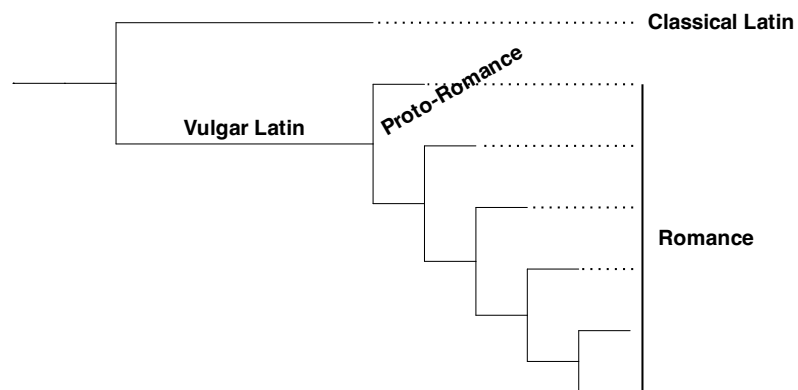


Figure 3: Three phylogenetic hypotheses for the relationship between Classical Latin, Vulgar Latin, and the Romance languages

2.3.2 The sibling hypothesis

Of the three hypotheses in Figure 3, the sibling hypothesis enjoys by far the most support (e.g., Väänänen 1983:483, Hall 1950:19, Coseriu 1954:29, Hall 1974:14, Mańczak 1977:13, Vallejo 2012:458). It has also featured prominently in some Bayesian phylogenetic studies, such as Gray and Atkinson 2003, Bouckaert et al. 2012, and Bouckaert et al. 2013. It is essential to bear in mind, however, that these studies use coalescent tree models, which do not allow sampled ancestors, so Latin was bound a priori to be a sibling to Proto-Romance. In other words, the sibling hypothesis is an assumption of these studies, not a result. By contrast, in the FBD analyses of Rama 2018, Latin could have been sampled as an ancestor but was not.

The sibling hypothesis relies crucially on the view that Vulgar Latin and Classical Latin are distinct enough that they should be modeled as separate taxa. There is one grammar and lexicon for Vulgar Latin and one for Classical Latin. Although they share many properties on account of their common ancestry, they are ultimately distinct entities (as opposed to a single entity with variation, e.g., between high- and low-register features). Garrett (2018:34) estimates that, according to the results of Bouckaert et al. 2012, Classical and Vulgar Latin would differ in just over ten percent of their basic vocabulary, which is comparable to the difference between French and Italian. Given that the field still awaits an adequate definition of Vulgar Latin,³ it is unclear how one would even demonstrate such lexical disparity.⁴

Despite wide acceptance, empirical support for the sibling hypothesis has been difficult to come by, as even some of its advocates acknowledge (Murray et al. 1994:371). The crucial evidence for the sibling hypothesis is innovations in Classical Latin that are absent in Vulgar Latin and Romance. While a full review of

³There is an overwhelming literature on what the term *Vulgar Latin* denotes or should denote. See, e.g., Tovar 1964, Mańczak 1974:218–220, Lloyd 1979, Väänänen 1983:483, Coseriu and Meisterfeld 2003:149–171, Stefenelli 2003:530, Coseriu 2008:29–32, 108–114, 147–167, Wright 2011:63, Adams 2013:3–12, Eskhult 2018, Weiss 2020:542 n. 6, Versteegh 2022.

⁴A reviewer wonders what a definition of Vulgar Latin would amount to and why it is so important. An adequate definition of Vulgar Latin would enable one to identify its empirical basis—that is, to determine what is and is not Vulgar Latin. As it stands, the identification of Vulgar Latin traits (in particular, lexical items) is anything but clear. Without knowing the lexical inventory of Vulgar Latin, one cannot even claim that Vulgar Latin and Classical Latin differ in their basic vocabulary, let alone assert that the former is a sampled ancestor of the Romance languages.

this evidence cannot be undertaken here (see Mańczak 1977:71–95 for discussion of some of it), I am aware of no unambiguous example of such a change. For instance, Hall (1950:19) argues that the invariant relative pronoun *cuius* ‘whose’ of Classical Latin is an innovation that sets it apart from Vulgar Latin and Romance, where we find an inflected relative adjective (e.g., Spanish *cuyo*, *cuya*). Hall’s description of the facts is inaccurate, however, because the relative adjective is attested in Classical authors such as Cicero (e.g., *Ver.* 2.127, 3.68) and the invariant relative pronoun is found in Vulgar texts from various periods (e.g., Petron. 29.8, *Itin. Eger.* 21.1, *Anon. Val. II* 56). Under the sibling hypothesis, invariant *cuius* should not occur in Vulgar texts and the relative adjective should not occur in Classical texts. These predictions are countered by the evidence.

In a similar vein, Murray et al. (1994) argue that the diachrony of the syllabification of *muta cum liquida* clusters (i.e., plosive plus liquid, such as *-gr-*, *-dr-*, *-pl-*, *-kl-*) provides evidence for the sibling hypothesis. They contend that these clusters were heterosyllabic in Vulgar Latin and Romance, but tautosyllabic in Classical Latin. This description of the facts is not entirely accurate, however (see Weiss 2020:76–78). Leaving this problem aside, their conclusion is still unwarranted, because they fail to take into account a competing hypothesis. All forms of Latin could have passed through a stage in which *muta cum liquida* clusters were tautosyllabic and in the first centuries CE became heterosyllabic (for more on such an analysis, see Loporcaro 2005:422). Under such a scenario, a covert form of Latin with heterosyllabic *muta cum liquida* clusters would be otiose. In short, uncontested evidence for the sibling hypothesis has yet to emerge. In Section 6.4 below, I show that lexical evidence likewise argues against it.

2.3.3 The most-recent common ancestor hypothesis

Far fewer scholars have come out in favor of the MRCA hypothesis. Its most prominent proponent is Witold Mańczak, who argued for decades on its behalf (Mańczak 1974, Mańczak 1978, Mańczak 1977, Mańczak 1980, Wittoch 1984, Mańczak 1987a, Mańczak 1987b, Mańczak 1991, Mańczak 1994b, Mańczak 1994c, Mańczak 1998, Mańczak 2006, Mańczak 2007, Mańczak 2013). Despite his determination, he persuaded few if any. One of the problems with this hypothesis is that Classical Latin often does not agree with reconstructions of Proto-Romance via the comparative method (e.g., Pulgram 1958:146–147, Hall 1974:12, Agard 1984:36–38, Coseriu 2008:111). If Latin is the MRCA of the Romance languages, such discrepancies are difficult to explain. Although Mańczak’s hypothesis has

not found much favor, his work did highlight an extremely important point, namely that evidence supporting the sibling hypothesis is difficult to come by (Murray et al. 1994:371–372).

2.3.4 The sampled-ancestor hypothesis

The sampled-ancestor hypothesis is the least discussed of the three hypotheses. Chang et al. (2015) argue extensively for it and other scholars have expressed views consistent with it (e.g., Janson 1979:13, Varvaro 1991:47, Stefenelli 2003:530, Adams 2013:10, 819, Banniard 2013:66, 68). The central feature of this hypothesis is that it embraces variation within Latin. In contrast to the sibling hypothesis, it does not assign low-register or low-sociolect features to another taxon. Instead the variation is a property of a single language, Latin *tout court* (Meyer-Lübke 1920:§99). The results of my experiment support this hypothesis (see Section 5.4 below) and further evidence in its favor is presented in Section 6.4 below.

3 Methods

Bayesian divergence-time analyses of languages standardly consist of three components. The first is the clock model, which relates the amount of linguistic change to time. The second is the character model, which characterizes the relative rates of change among character states. The third is the tree model, which is a stochastic model of linguistic speciation and extinction. Each of these components is described in greater detail in the following sections.

3.1 Clock models

Three clock models are compared in this study. The simplest is the strict clock, which imposes a single rate of change on the entire tree. This rate is sampled from a gamma prior distribution with a shape parameter of 2.0 and a scale parameter of 4.0, which is visualized in Figure 4. I selected this prior because it places more density on lower rates of change without entirely excluding the possibility of a higher rate. Although they boast simplicity, strict clocks are generally considered linguistically unrealistic.

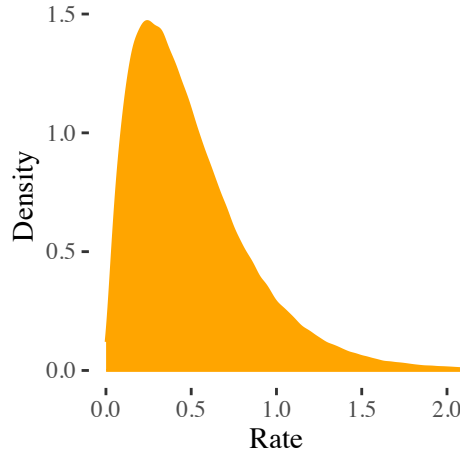


Figure 4: Gamma prior distribution for the strict clock model

The second two are relaxed clock models, in which a rate of change for each branch is independently drawn from a prior distribution. Relaxed clock models thus allow the rate of linguistic change to vary from branch to branch. Since the rates for each branch are drawn independently, they can vary considerably across the tree.

The first of the two relaxed clock models is an uncorrelated exponential clock model (UCE) and the second is an uncorrelated log-normal (UCLN) clock model. The central difference between these two models is how much rate variation they allow. With both clocks, the rate of linguistic change is generally going to be slow, but the log-normal model allows for higher rates of change. The prior distributions of the relaxed-clock models are visualized in Figure 5, where it can be seen that the UCLN clock assigns more density to a far wider range of rates compared to the UCE.

The branch rates for the UCE model are drawn from an exponential distribution with a parameter of $1/x$, where x is the mean value of the branch rates and is itself drawn from an exponential distribution with a parameter of 5.0. The branch rates of the UCLN model are drawn from a log-normal distribution controlled by two parameters—the location parameter μ and the standard deviation σ . The location parameter is calculated as follows:

$$\mu = \ln(\text{mean}) - (\sigma^2 * 0.5)$$

The value for the mean is drawn from an exponential distribution with a parameter of 2.0 and sigma from an exponential distribution with a parameter of 3.0.

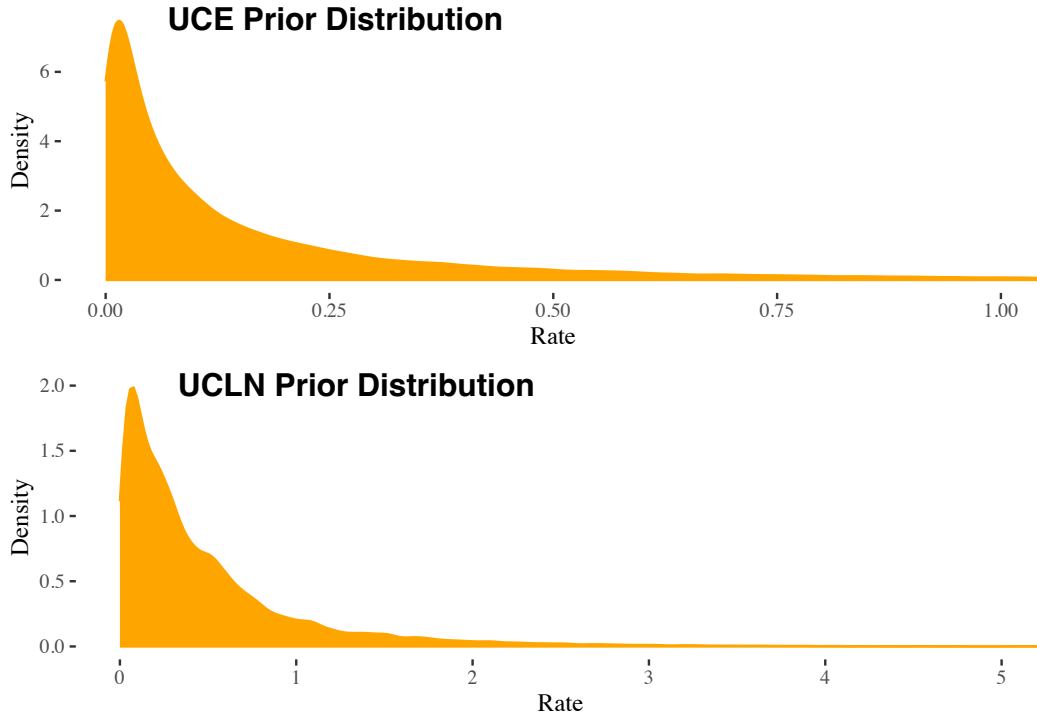


Figure 5: Prior distributions for the relaxed clock models

3.2 Character models

In Bayesian phylogenetics, linguistic change is standardly modeled as a continuous-time Markov chain (CTMC; for an overview, see Cathcart 2018, Jäger 2019:166–168). Since the values in the dataset are arbitrary and lack a consistent meaning from one root-meaning character to the next, the choice of character-change models is limited. In this study, the Mk (Lewis 2001) and F81 mixture model are used. The Mk model is the simpler of the two, as it imposes symmetric relative rates of change. Transitions from 0 to 1 happen at the same rate as transitions from 1 to 0:

$$Q = \begin{pmatrix} - & 1.0 \\ 1.0 & - \end{pmatrix}$$

The values on the diagonal are equal to the negated sum of the values in each row (in this case -1.0). The Mk model assumes that all characters change at the same rate and that the stationary frequencies of the character states are equal.

Both of these assumptions can be relaxed. To relax the first, rates are drawn from a gamma distribution whose shape and rate parameters are constrained to be equal. The distribution is then partitioned into four rate classes, with the mean value from each partition serving as the rate for each class. The likelihood of a character is then a weighted average over the four rate classes. The prior distribution for the shape parameter α is drawn from a log-normal distribution with a mean of $\ln(5.0)$ and a standard deviation of 0.587405 (Höhna, Landis, and Heath 2017:6–7). To relax the assumption of equal stationary frequencies among characters, an F81 mixture model is used, whose stationary frequencies are drawn from a beta distribution discretized into two categories. In short, Mk+ Γ and F81+ Γ (where F81 represents an F81 mixture model) character models are used in this study.

These models introduce a number of assumptions about the process of lexical change. The first is the Markov property, which means that the probability of a particular character state at a particular time depends only on its immediately preceding state. Earlier states are irrelevant. Second, the relative rates of change between character states are constant over time. Finally, character histories are independent: how one character evolves has no impact on any other.

3.3 Tree models

The tree models used in this study are based on stochastic models of lineage diversification and extinction known as birth-death processes (Kendall 1948). Under the standard birth-death model (Yang et al. 1997), the probability density of a tree is conditioned on a speciation-rate parameter λ , an extinction-rate parameter μ , a sampling probability parameter ρ , and the age of the tree. The parameters λ and μ are constant over time, which is to say that at every moment in time each language has the same speciation and extinction rates. (The assumption of rate constancy is relaxed under episodic birth-death models, which are introduced in Section 3.3.2 below.) I follow the parameterization in Gavryushkina et al. 2016:60, according to which the estimated parameters are not λ and μ , but the net diversification rate d and relative extinction rate v :

$$d = \lambda - \mu$$

$$\nu = \frac{\mu}{\lambda}$$

One advantage of these parameters is that specifying prior distributions is more straightforward. With the net diversification rate, for instance, the prior distribution for d is centered on the number of languages at the present.

The parameter ρ is the probability of sampling a language at the present time and is calculated as the ratio of languages in the study group to the total number of extant languages in the family. It thus takes into account incomplete sampling of extant languages. In the present study, for instance, there are fourteen contemporary languages out of a total of 76 extant Romance languages recognized by Glottolog 4.6 (Hammarström et al. 2021). (Glottolog identifies 78 Romance languages, from which I excluded Mozarabic and Dalmatian since they are extinct.)

There are two significant drawbacks to the standard birth-death model. The first is that it can only model trees with extant taxa (i.e., languages). Ancient languages cannot be positioned on the tree, whether as tips or as sampled ancestors. The second is that the standard birth-death model relies crucially on node-calibration densities, which are distributions over the ages of clades specified by the researcher. Although node calibrations have been used in a number of high-profile studies (e.g., Gray and Atkinson 2003, Chang et al. 2015, Bouckaert et al. 2012, Bouckaert et al. 2013), they bring with them pitfalls (Ronquist et al. 2012:973–974). For one, node calibrations can involve guesswork. In some cases, the temporal information for the age of a clade can be limited to the dates of the earliest texts. While such dates do provide a *terminus ante quem*, calibrating the age of a node solely from the textual record is anything but straightforward. For another, in clades with more than a single fossil only the earliest is used (Ronquist et al. 2012:973–974). Rama (2018:187) illustrates this point with Germanic, which features a number of ancient languages, such as Gothic, Old English, and Old High German. Under a node-calibration approach, only Gothic is relevant, since it is the oldest of the ancient Germanic languages and as such provides a *terminus ante quem* for the clade. For estimating rates of change, all of the taxa offer crucial information, however.

3.3.1 Fossilized birth-death models

Stadler (2010) introduced an innovation to the birth-death process that incorporates information from non-extant taxa. This innovation was a major advance, since it created a coherent framework for the calibration of divergence-time estimates (Heath et al. 2014). This is a particularly important development for linguistic phylogenetics where ancient languages play a crucial role in the estimation of both tree topology and divergence times (Rama 2018:189). In addition to the parameters mentioned above for birth-death models, FBD models also have a parameter ψ , the fossilization rate. This is the rate at which lineages produce fossils and encompasses a number of different aspects of linguistic history, from the actual extinction of a language to its recovery by scholars. The following parameterization from Gavryushkina et al. 2016:60 was used:

$$s = \frac{\psi}{\psi + \mu}$$

s is the fossil sampling proportion, ψ the fossilization rate, and μ the extinction rate. In informal terms, s represents the probability that a language undergoes fossilization before extinction. An exponential prior distribution with a parameter of 5.0 was placed over s in this study.

All of the analyses presented in this study are conditioned on the origin age of the process, which means that they start with a single lineage. The alternative is to condition on the root age, in which case the process starts with two lineages (i.e., the two that emerge from the first split). If one were to condition the process on the root age, then Latin could not be sampled as an ancestor of the Romance languages, since the branch leading to the Romance languages (i.e., the stem) would not be part of the process. It is thus crucial to condition on the origin age so as not to rule out a priori the possibility of Latin as a sampled ancestor of the Romance languages.

The origin age is drawn from a uniform distribution with a lower bound of 2.2 kya and an upper bound of 2.6 kya (from the year 2000, so 600 BCE), which represent the approximate dates of pre-Classical Latin. The lower bound of 2.2 kya coincides with the beginning of Classical period and the upper bound of 2.6 kya is the approximate age of the earliest Latin inscriptions.

In a Bayesian-Markov chain Monte Carlo (MCMC) framework, trees and param-

eter values are sampled from the posterior distribution. For the FBD process, the posterior distribution conditioned on time t is calculated as follows (Stadler 2010:402):

$$f[\mathcal{T}, \lambda, \mu, \rho, \psi, \theta | \mathbf{X}] \propto f[\mathbf{X} | \mathcal{T}, \theta] f[\mathcal{T} | \lambda, \mu, \rho, \psi, t] f[t | \lambda, \mu, \rho, \psi] f[\lambda, \mu, \rho, \psi] f[\theta]$$

\mathcal{T} is the tree, θ is a metavariable over the parameters of the character-transition model (presented in Section 3.2 above), and \mathbf{X} the character data. $f[\mathcal{T} | \lambda, \mu, \rho, \psi, t]$ is calculated with equation (3) in Stadler 2010:400.

Stadler et al. (2018) introduce another innovation to the birth-death process, which incorporates the stratigraphic ranges of fossils, that is, the span of time in which they are believed to have existed in the geological record. This process is known as the fossilized birth-death range (FBDR) process (see in particular equation 10 from Stadler et al. 2018:50). From a linguistic perspective, stratigraphic ranges can be compared to the first and last occurrences of texts in a given language. Compared to the standard FBD process, the FBDR process more accurately represents information from ancient languages, since it takes into account that (at least in the case of Latin) it comes from a dense textual record and not a single specimen (i.e., a single text). I elaborate on this point in Section 4.1 below.

3.3.2 Episodic models

As noted above, the parameters λ , μ , and ψ in the standard FBD and FBDR models are time-homogeneous (i.e., they are constant over time), but this assumption can be relaxed with episodic models. With such models the values of these parameters are estimated per epoch. Table 2 specifies the epochs. The year 476 CE (the traditional date of the fall of the Western Roman Empire) partitions into two the time from the maximum value of the origin age (2.6 kya) to the present. Only two epochs were used on account of the limited number of languages in the dataset. With a larger dataset, more time intervals could be included.

Name	Start	End
Epoch 2	1.52	0.00
Epoch 1	2.60	1.52

Table 2: Epoch ranges (kya)

3.4 Model comparison

The components introduced in the preceding sections can be combined into different model configurations. To compare the performance of different models, I use the log Bayes factor, which is the natural logarithm of the ratio of the marginal likelihoods under competing models:

$$BF_{10} = \ln \left[\frac{p(\mathbf{X}|M_1)}{p(\mathbf{X}|M_0)} \right]$$

BF_{10} is a measure of the extent to which the data support model M_1 over M_0 and is interpreted according to the categories in Table 3. Bayes Factors only measure the *relative* fit of a model to data and not the absolute fit. Model adequacy is assessed with posterior predictive simulation, which is introduced in the next section.

Strength of Evidence	Log-BF(M_1, M_0)
Negative (supports M_0)	<0
Barely worth mentioning	0 to 1.16
Substantial	1.16 to 2.3
Strong	2.3 to 4.6
Decisive	> 4.6

Table 3: Discrete categories of log Bayes Factors (Höhna, Landis, and Heath 2017:27)

3.5 Posterior predictive simulation

A fundamental (but neglected) question in all model-based investigations of linguistic history is the models adequately represent the process of linguistic change. This question can be addressed with posterior predictive simulation (PPS). PPS works by simulating a dataset on the basis of parameter values sampled from the posterior distribution. Repetition of this process yields multiple simulated datasets, whose similarity to the original dataset is measured through test statistics.

I use the test statistics of May et al. 2021, namely the total parsimony score (the minimum number of transitions for each character on a given tree) and the

variance in parsimony scores among characters. In addition, I also calculate the proportions of 1 values in the simulated datasets. The total parsimony score and proportion of 1 values reflect the ability of a model to adequately describe the overall rate of character change, whereas the variance in parsimony scores characterizes how well a model captures differences in the rate of change among characters.

A posterior predictive p -value is used to compare the test statistics of the simulated datasets to that of the original data. The posterior predictive p -value is the proportion of simulated test statistics that are greater than the observed test statistic:

$$p = \frac{1}{n} \sum_{i=1}^n T(\mathbf{X}_i^{\text{sim}}) > T(\mathbf{X}^{\text{obs}})$$

N represents the total number of simulated datasets, $T(\mathbf{X})$ a test statistic, $\mathbf{X}_i^{\text{sim}}$ the i th simulated dataset, and \mathbf{X}^{obs} the observed data.

To illustrate how this formula works, assume 100 datasets simulated from the parameter estimates of a model. The test statistics described above are calculated for the observed dataset as well as the 100 new datasets. The number of times a given test statistic of the simulated datasets is greater than that of the observed dataset is then summed up and divided by 100. For instance, if 67 of the 100 simulated datasets have a total parsimony score greater than that of the original dataset, the posterior predictive p -value is 0.67. For a significance level of $\alpha = 0.05$, a p -value greater than 0.975 or less than 0.025 is interpreted as model inadequacy. The guiding idea behind posterior predictive simulation is that adequate models generate datasets resembling the original data.

3.6 Software

All analyses were performed with RevBayes 1.1.1 (Höhna, Landis, Heath, et al. 2016, Höhna, Landis, and Heath 2017). Six independent MCMC chains were run for 500,000 cycles for the time-homogeneous analyses and for 800,000 cycles for the episodic analyses. In both cases, the warm-up phase was 1,000 cycles. Every fifth sample was recorded and the first twenty percent of the samples were discarded as burn-in. Convergence was confirmed with the measures in the R package `convenience` (Fabreti et al. 2021). Marginal likelihoods were calculated

with both the path-sampling and stepping-stone methods. Analyses with 128 stones were run for 10,000 generations after a warm-up phase of 1,000 generations. Every tenth generation was sampled.

4 Data

This study uses a modified version of the broad dataset of Chang et al. 2015, which is ultimately based on the IELex dataset created in 2011 by Michael Dunn and colleagues at the Max Planck Institute for Psycholinguistics in Nijmegen. The complete broad dataset spans 94 Indo-European languages and contains 5,694 binary characters from 197 meaning classes. The characters are root-meaning traits (Chang et al. 2015:201) and loanwords were not excluded. 4,142 characters are uniformly 0 in Latin and the Romance languages, since most of the root-meaning traits present elsewhere in Indo-European are absent in Latin and Romance. After these characters are removed, 1,552 are left, of which 56 are uniformly 1. A dataset containing characters that can be uniformly 1 but not uniformly 0 will inflate the overall rate of change, so the 56 characters with uniformly 1 values are removed as well. The result is that the dataset only contains characters that vary, as a result of which it is necessary to correct for ascertainment bias. (Ascertainment bias refers to the absence of characters that either vary minimally or not at all.) Correcting for ascertainment bias in *RevBayes* also requires the removal of characters with missing values (i.e., ? values). Once these are removed, 643 characters remain, of which 283 are parsimony informative (parsimony-informative characters are those with at least two 0 values and at least two 1 values).

4.1 The age of Classical Latin

In addition to the primary linguistic data, the analyses in this study also take into account the age of Latin, which is presented in Table 4. A range of 1.9–2.2 kya, which corresponds to 200 BCE–100 CE, is used, since the data represent the Latin of the Classical period (on the periodization of Latin, see, e.g., Adamik 2015, Vincent 2016:5).

Although both the FBD and FBDR models use the same age range of 1.9–2.2 kya, its treatment differs crucially between the two. For the FBDR models, the range encompasses the first and last occurrences of Classical Latin, i.e., a stratigraphic range. By contrast, for the FBD models, the interval represents uncertainty: Latin

Language	Min	Max
Latin	1.9	2.2

Table 4: The age range of Latin (kya)

is treated as a single fossil, whose age is sampled from a uniform distribution bound by the minimum and maximum occurrence times. Representing Latin in this way is awkward because the Classical Latin dataset in this study is not based on a single text dated between 100 BCE and 200 CE. The character states are instead extrapolated from a dense textual history attested throughout this period. As stratigraphic ranges more adequately represent the nature of linguistic data from ancient languages, this is one reason to prefer them.

Before moving on to the results, I want to note that the FBDR process is also able to take into account the ages of extant languages. Unfortunately this experiment is unable to take advantage of this feature, since it is unclear how far back the root-meaning traits in the dataset date. Ideally investigation of the divergence times of the Romance would include not only stratigraphic ranges for contemporary languages but also datasets and age ranges for the medieval languages. I reserve this endeavor for future work. For the moment it is important to understand simply that the full potential of the FBDR process is not being exploited.

5 Results

5.1 Model sensitivity

To illustrate the sensitivity of the phylogenetic inferences to the specification of the model, this section presents lineage-through-time (LTT) and multidimensional-scaling (MDS) plots. The LTT plot in Figure 6 displays the average number of languages over time in the posterior distributions of trees from the FBD and FBDR F81 models. The models exhibit a range of trajectories, with the process of diversification beginning earlier under the FBD F81 UCE model and later under the FBDR F81 Strict model. Differences such as these exemplify the model sensitivity that Heggarty (2021:382) and Ringe (2022:60) highlight.

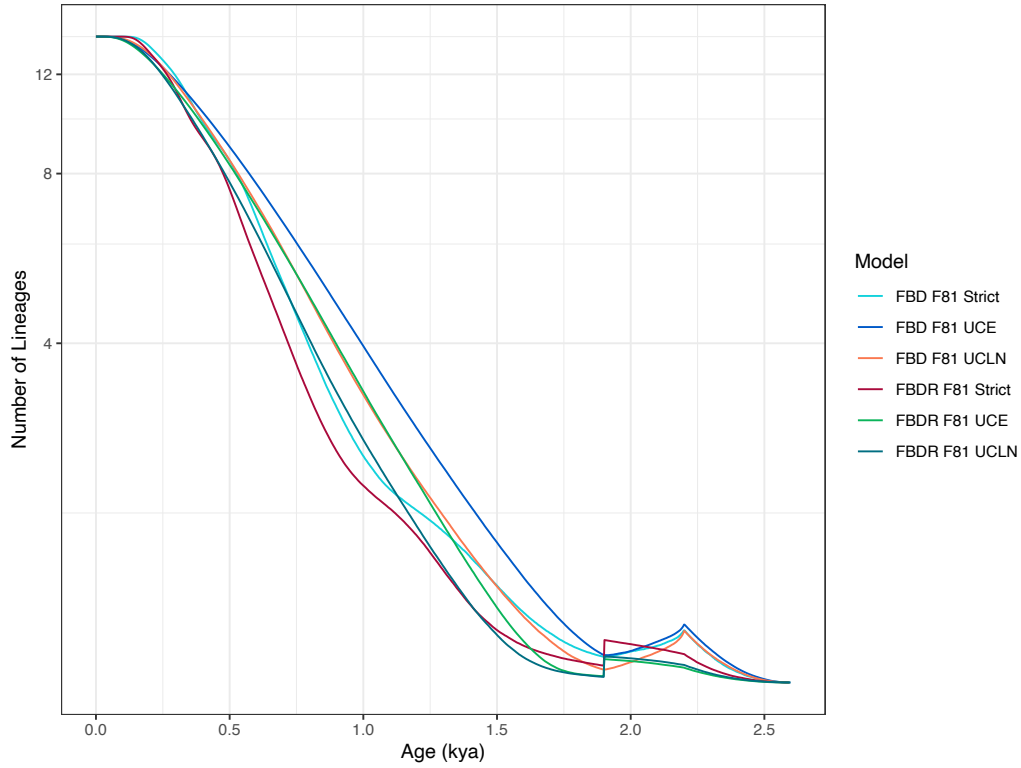


Figure 6: Lineages through time (LTT) plot

While useful, lineage-through-time plots are nonetheless blinkered, since they only compare the average number of languages. They neglect important differences between the posterior distributions, such as tree topology and branch lengths. To assess the sensitivity of these aspects of the model, I use the Kühner-Felsenstein distance metric (Kühner et al. 1994), which calculates the distance between every pair of trees within and between each posterior distribution, taking into account both tree topology and branch lengths. Multidimensional scaling (MDS) plots then reduce these distances to a lower-dimensional space (Hillis et al. 2005, Huang et al. 2016). Figure 7 reveals a profound difference between the strict and relaxed clock models. The latter are sampling extensive areas of the tree space, whereas the former cluster together near the center.

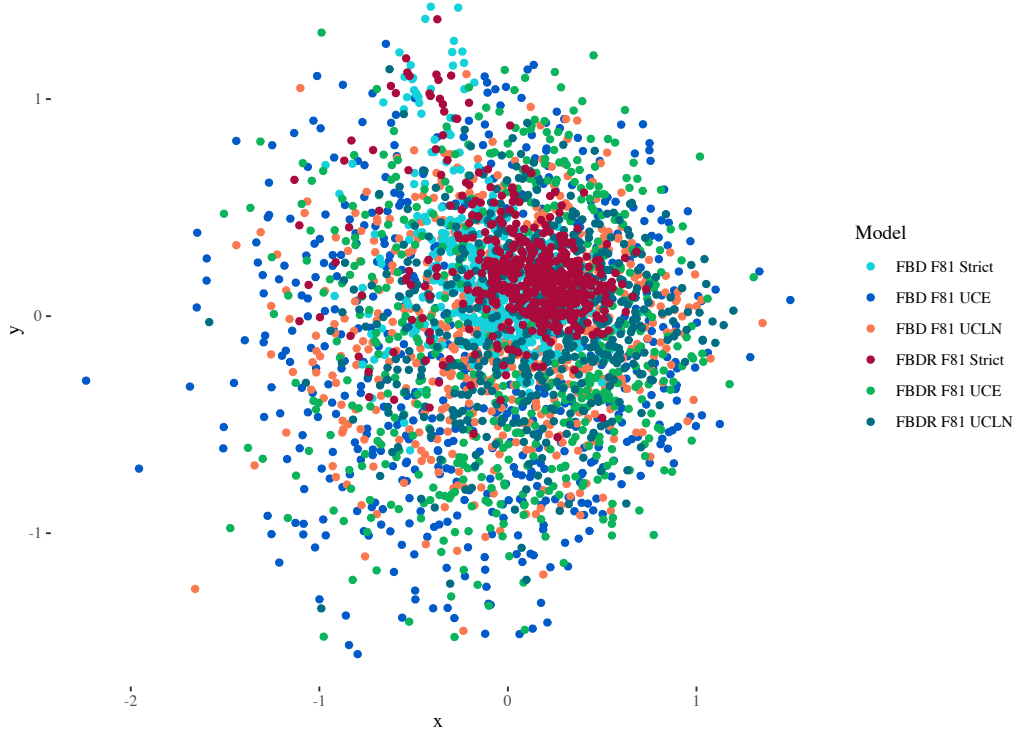


Figure 7: Multidimensional Scaling (MDS) plot of the Kühner-Felsenstein distances

5.2 Model comparison

The log marginal likelihood of each model is presented in Figure 8, which reveals three key insights. The first is that the F81 mixture model outperforms the Mk model across the board. The second is that the relaxed clock models (UCE and UCLN) without exception improve on the strict clock models. The third is that neither the more complex FBDR models nor the episodic models improve the marginal likelihoods. The best models are thus the time-homogeneous F81 relaxed clock models.

Although the marginal likelihoods of the FBD F81 UCE and UCLN models are similar (the log Bayes factor in favor of the former is 0.06, which falls into the category “barely worth mentioning”), there is reason to prefer the simpler UCE clock model. As noted above in Section 3.1, the UCLN clock model assigns more density to a wider range of rates. The 95% HPD interval of the standard deviation

of FBD UCLN clock model is $[0.37, 0.92]$, with a median of 0.623. Such low values suggest that the rate heterogeneity is not substantial enough to warrant the greater flexibility of the log-normal distribution. I therefore devote especial attention to the FBD F81 UCE model in the discussion below.

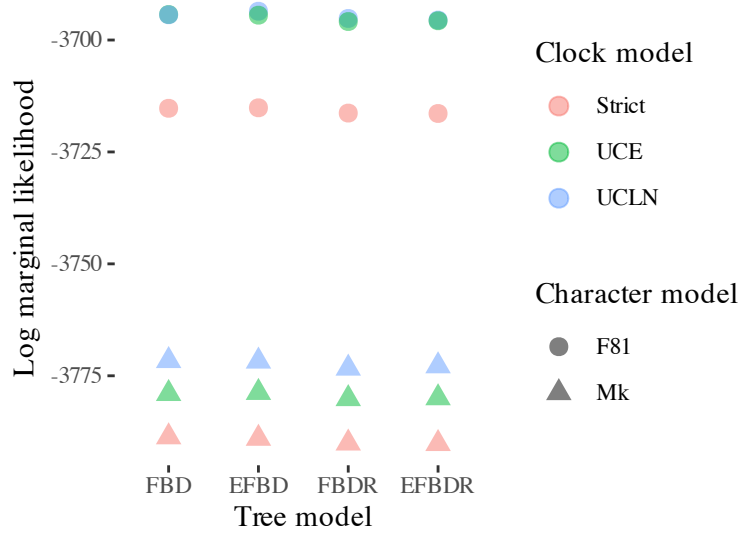


Figure 8: Log marginal likelihoods

5.3 Divergence times

The posterior distributions of the age of Proto-Romance under the four F81 models with relaxed clocks are presented in Figure 9. Two properties of these distributions stand out. First, they are remarkably similar. With the exception of the FBDR F81 UCLN model, which is centered around a slightly earlier date, the distributions cluster around 400–600 CE. So by and large the inferred age of Proto-Romance is robust across the four models. Second, the distributions are wide—in each case, they span a millennium—and thus reflect considerable uncertainty.

Figure 10 homes in on the posterior distribution of the FBD F81 UCE model. The black lines demarcate the 89% credible interval, while the green line represents the maximum a posteriori (MAP) estimate. The results of this model (as well as the other three in Figure 9) support the late hypothesis of the diversification of the Romance languages, as only 22% of the posterior density lies between 0 and 300 CE.

Although the results support the late hypothesis, they do not offer a clear answer to the question of whether the diversification of Romance begins before the fall of the Western Roman Empire. About 45% of the posterior density lies between 0 and 476 CE, whereas 44% lies between 476 CE and 1000 CE. It is true that the maximum a posteriori estimate (1.568 kya or 432 CE) predates the fall of the Western Roman Empire, but given the width of the distribution it would be unwise to make too much of this value.

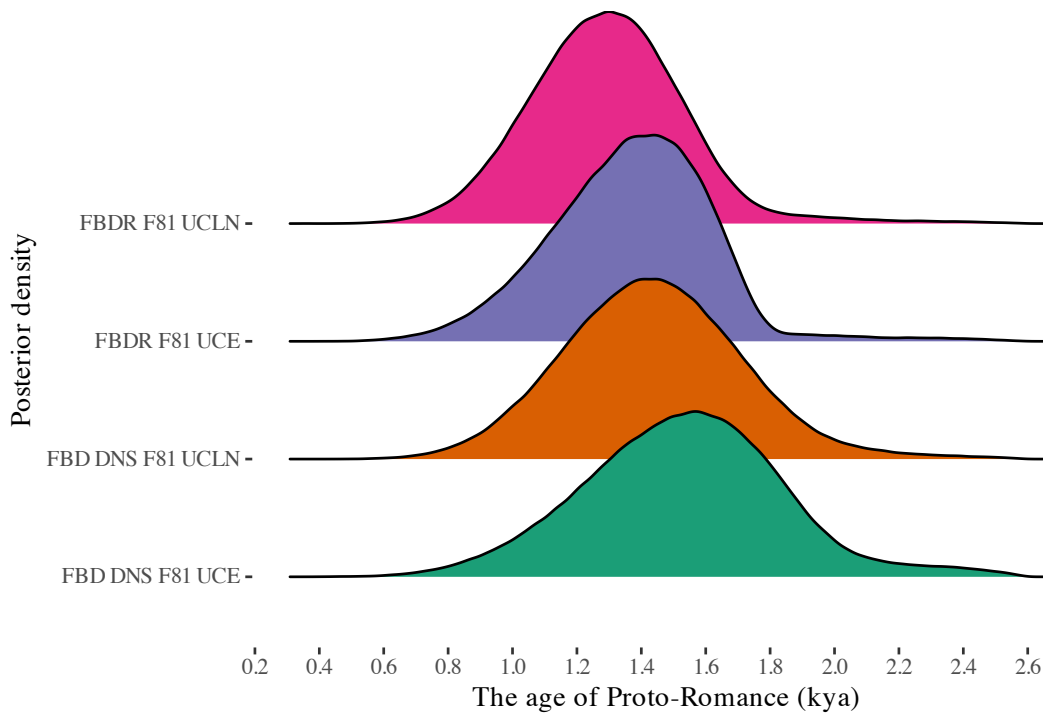


Figure 9: The age of Proto-Romance

5.4 Tree topology

The maximum clade credibility (MCC) trees for the FBD and FBDR F81 analyses with relaxed clock models are presented in Figure 11. The width of the node bars represents the 95% credible interval of the age of that clade. The color of the node bars corresponds to the posterior probability that the clade exists or that Latin is a sampled ancestor. The tree topologies are identical: the diversification of Latin begins at the eastern edge of the Empire with the formation of the Romanian

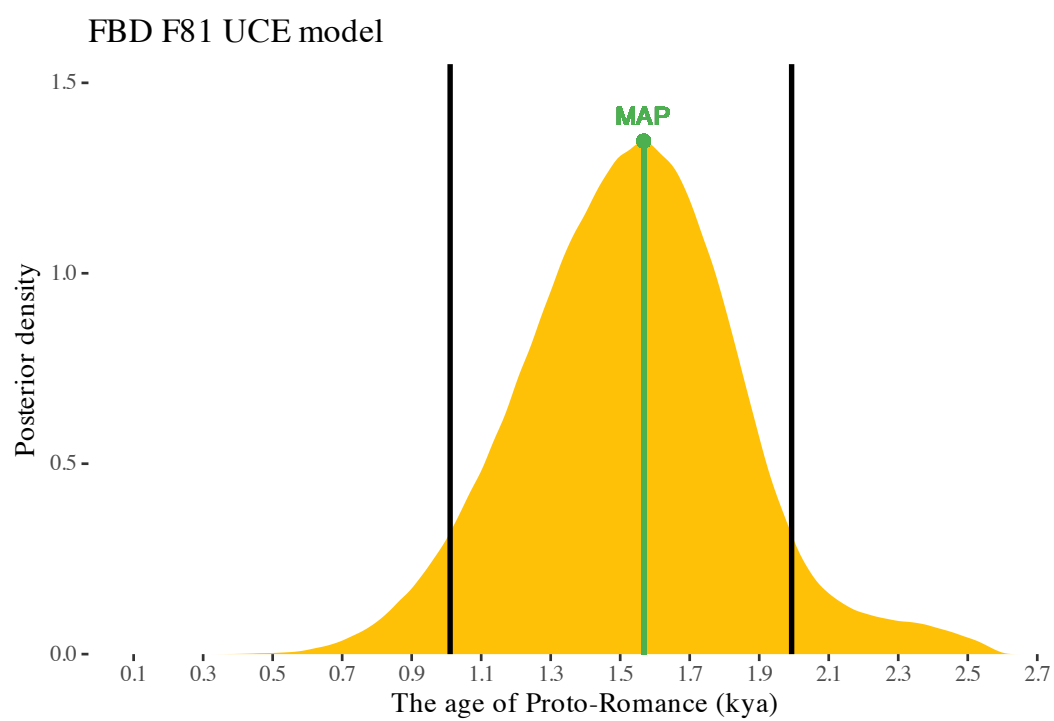


Figure 10: The age of Proto-Romance under the FBD F81 UCE model

dialects, which is in turn followed by the development of the Sardinian dialects. The Italo-Western and Western clades then emerge in the wake of Sardinian.

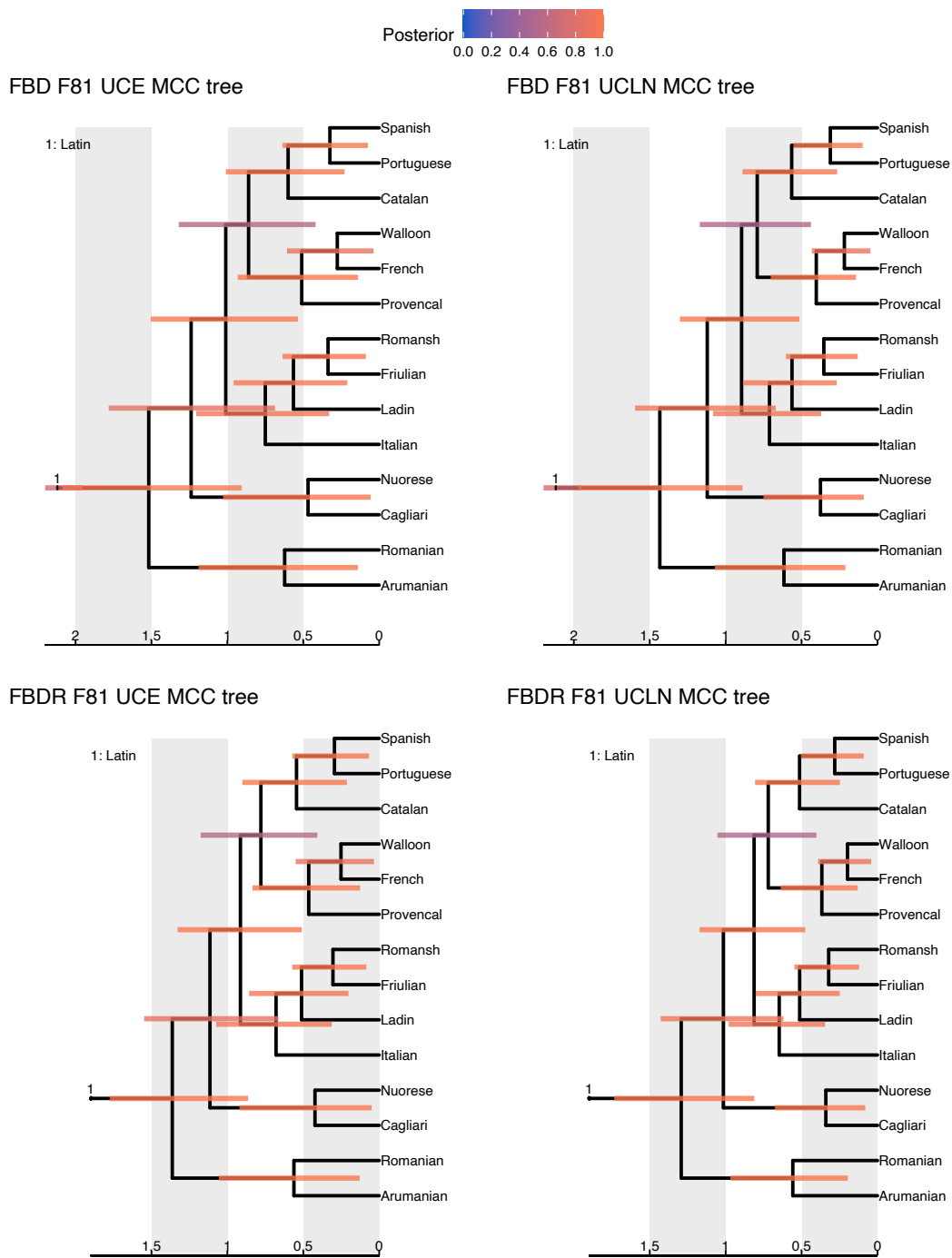


Figure 11: Maximum clade credibility (MCC) trees

In all of the trees in Figure 11, Latin is positioned on the branch leading to the Romance languages. It is, in other words, a sampled ancestor of the Romance languages. The posterior probabilities of Latin as a sampled ancestor in each of the six F81 models are presented in Figure 12, where two properties stand out. First, there is little difference in posterior probability between the UCE and UCLN clock models. Second, the tree model has a substantial impact on the posterior probabilities. For the FBD relaxed clock models, the posterior probability is around 70%, whereas for the FBDR models, it is around 90%.

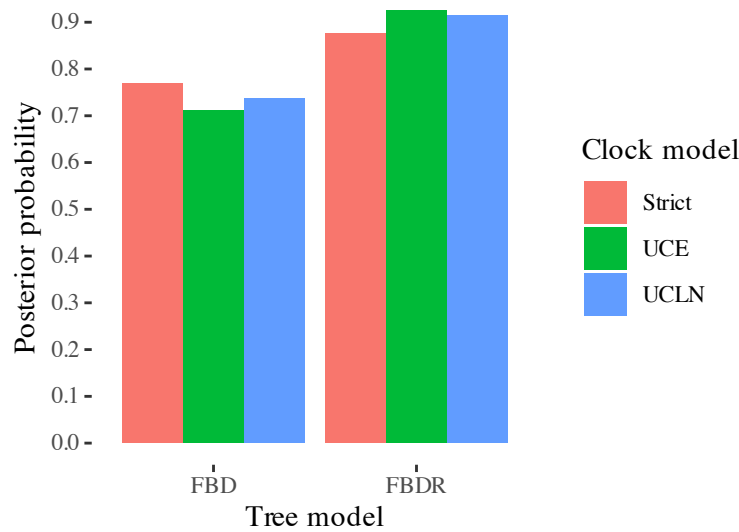


Figure 12: Posterior probabilities of Latin as a sampled ancestor

5.5 Model adequacy

The results of the posterior predictive simulations are presented in Figure 13 for the FBD and FBDR F81 relaxed-clock analyses. The p -values of all four models are greater than 0.025 and less than 0.975, from which I conclude that they are all adequate.

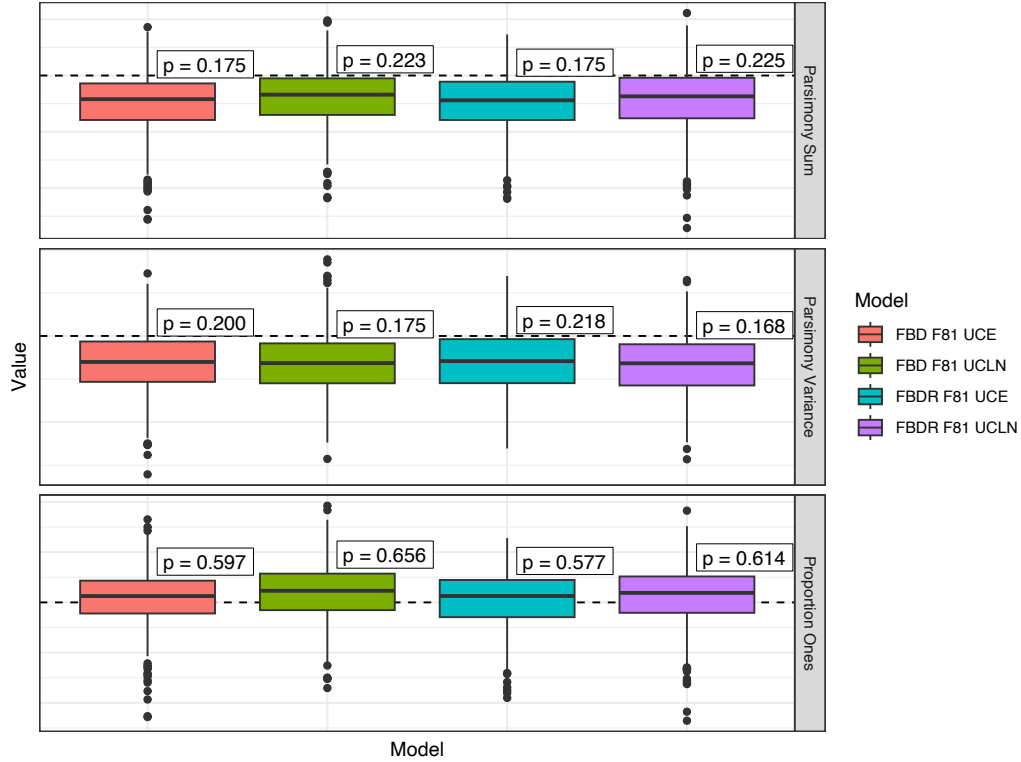


Figure 13: Posterior-predictive distributions of the test statistics

6 Discussion

6.1 Tree topology

As noted above in Section 2.1, many if not most Romance linguists now adhere to the view that the Sardinian dialects formed first and Romanian second. The position of Romanian in the trees in Figure 11 is at odds with this view. There are at least three possible causes of this discrepancy. First, recall from Section 2.1 above that developments in the vowel systems of Sardinian and Romanian play a crucial role in traditional phylogenetic analyses of Romance. Such data are absent in my experiment and it could be the case that the lexical data point to a different tree topology. Second, it is possible that the inferred position of Romanian is sensitive to the selection of lexical characters. Rama (2018), for instance, runs five different FBD analyses with different datasets from Chang

et al. 2015. Of these five, Romanian forms before Sardinian in the consensus tree of three; in two (which are based on the medium and narrow datasets of Chang et al. 2015), Sardinian forms first. Finally, the position of Romanian may be due to inaccuracies in the coding of the data (Heggarty 2021:385).

6.2 Divergence times

In Section 2.2 above, I introduced two competing hypotheses concerning the diversification of Latin. The first maintains that the process of diversification begins in the first centuries CE and is supported above all by evidence from sound change. The second contends that diversification begins later and finds its support in textual evidence. The results of the analyses presented in Section 5.3 above support the late hypothesis. They do not, however, warrant the rejection of the early hypothesis.

The early and late hypotheses are presented in the literature as competing hypotheses (e.g., Adams 2007:684). On the one hand, this presentation seems *prima facie* unobjectionable. There are two hypotheses concerning the diversification of Latin and the goal is to evaluate the evidence in favor of each and decide between them. On the other hand, these two hypotheses may not necessarily be in competition with one another. It could instead be the case that they capture two different phenomena, namely two different aspects of the diversification of Latin. If linguistic diversification begins with low-level phonetic change and only later reaches basic vocabulary, then divergence-time estimates from sound change and lexical data are bound to differ. The history of American and British English (RP) illustrates such a scenario, inasmuch as the two dialects are now distinct phonologically, but differ minimally in basic vocabulary. On the basis of the IE-CoR dataset (Findell et al. 2023), for instance, it appears that the only difference is between American *vomit* and British *be sick*.

Coming back to Latin, the process of diversification could well have started with sound changes in the first centuries CE and only later affected basic vocabulary. Under such a scenario, divergence times inferred from basic vocabulary could underestimate the timing of speciation events and one could maintain that evidence from sound change should therefore be preferred. Although divergence times inferred from sound changes are arguably more important, those from basic vocabulary are not irrelevant. Different components of language potentially diversify at different times and at different rates. Our understanding of linguistic

diversification will be incomplete if we do not attend to those differences.

An awareness of these differences is also important at a practical level. As noted above in Section 3.3, studies of divergence-time estimation often rely on node calibrations. These node calibrations are sometimes based on sound change, even when the study itself estimates parameters from basic vocabulary (e.g., Gray and Atkinson 2003, Bouckaert et al. 2012). This blending of evidence from sound change and lexical change may result in erroneous parameter estimates, since it disregards potential discrepancies in the timing and rate of change between the two.

6.3 The diversification of Latin and Roman history

The relationship between the diversification of Latin and the dissolution of the Western Roman Empire has long been debated (Becker 2014). The following two questions are central to this debate. First, did the fall of the Roman Empire precede the diversification of Latin? Second, did the fall of the Roman Empire accelerate the process of diversification? The results presented in Sections 5.2 and 5.3 do not offer an unequivocal answer to the question of whether or not the diversification of Latin was already underway by 476 CE. They do, however, offer some support for the view that the speciation rate of the Romance languages did not spike in the period between 476 and 1000 CE, since the episodic models fail to improve the marginal likelihoods. Under the constant-rate models, the speciation rate is the same throughout time and consistent with a view in which it does not change in the wake of the destruction of the Western Empire. Overall, the results of this study agree with a gradualist view of the formation of Romance (Becker 2014:261) and the recommendation of Adams (2007:725) that we get “away from the idea that Latin was monolithic until a very late date, when some catastrophic event caused it to ‘split up.’” It may accordingly be the case that the decentralization of the Roman Empire (Väänänen 1983:486–487, Coseriu 2008:85) was a more important factor in the diversification of Latin than the collapse of the Empire. Further examination of this question lies beyond the remit of this study.

6.4 The ancestry question redux

This study is the first to show strong support for Classical Latin as a sampled ancestor of the Romance languages. The posterior probability that Classical Latin is a sampled ancestor is about 70% among the relaxed-clock FBD models and

around 90% among the relaxed-clock FBDR models. Since these results counter the entrenched attitude that Classical Latin is not ancestral to the Romance languages, I offer in this section two arguments to buttress my conclusion. First, basic vocabulary items of Latin with Romance descendants are not always restricted to low-register or colloquial contexts, which contradicts the sibling hypothesis. Second, there appear to be no archaisms shared between the Romance languages and an Indo-European language other than Latin. This absence is predicted under the sampled-ancestry hypothesis, but puzzling under the sibling hypothesis.

6.4.1 The distribution of basic vocabulary in Latin

Chang et al. (2015:206) argue that the phylogeny of Bouckaert et al. (2012, 2013), in which Latin is a sister to the Romance languages, entails either that Classical and Vulgar Latin were highly diglossic or that an undocumented regional dialect that exhibited many differences in its basic vocabulary gave rise to the Romance languages. Of these two possibilities, the first is more plausible and has accordingly garnered more support. It predicts a polarized distribution of basic vocabulary items, in which Latin lexemes with Romance descendants are restricted to either low-register or colloquial contexts (depending on how one defines “Vulgar Latin”; see footnote 3 above). By contrast, Latin lexemes that do not survive in Romance should only be found in high-register or written contexts. Fortunately, the remains of Latin are substantial enough to allow us to test this prediction empirically, since private letters, contracts, graffiti, and the metalinguistic comments of grammarians afford us insight into lower registers of the language. Chang et al. (2015:207) analyze each case in their dataset in which there is a discrepancy in basic vocabulary between Latin and at least twelve of their fourteen Romance languages. Under the sibling hypothesis, one expects to find precursors of the Romance words in—and *only* in—colloquial Latin sources. Not a single example in their dataset bears out this expectation (Chang et al. 2015:207).

The example of ‘mouth’ illustrates the failure of the sibling hypothesis. The Classical Latin word is *os*, which survives nowhere in Romance. Romance lexemes for ‘mouth’ descend from a few ancestral Latin forms, foremost among which is *bucca* (for a conspectus, see Adams 2013:782–783). According to the sibling hypothesis, *os* is the Classical Latin word for ‘mouth’ and *bucca* the Vulgar Latin word. The textual distribution of these lexemes does not support this hypothesis, however, as *os* is attested across a range of styles and registers, including colloquial and low-register sources (such as Plautus, Pompeian inscriptions, Anthimus, and

the *Regula Benedicti*). This distribution contradicts the sibling hypothesis because *os* is not simply the default term for ‘mouth’ in written Classical Latin, but rather the default term across a variety of registers in the Classical period and early centuries CE, if not later (*pace* Heggarty et al. 2023:SM 91).

The distribution of *bucca* is more complex. Its earliest meaning appears to be ‘lower part of the cheek, jaw’ and it is used as an expressive synonym for the Classical Latin word *gena* ‘cheek’ (Desnitskaja 1982:237). The meaning ‘mouth’ develops from this earlier sense (Blank 1997:239–240). The use of *bucca* to denote the mouth appears to be restricted to the colloquial register (*DELL*:77, André 1991:38, 57, Ferri et al. 2010:29, Adams 2016:194), which *prima facie* could be interpreted as support for the sibling hypothesis. Heggarty et al. (2023:SM 91), for instance, contend that *bucca* is “more likely the default form in the spoken variety of Latin.” The problem with this interpretation is that *bucca* was not a neutral term for ‘mouth’ (Palmer 1954:170, *DELL*:77, Väänänen 1981:78) and stylistic neutrality is one of the hallmark properties of basic vocabulary items (Kassian et al. 2010:48, Heggarty 2021:390, Heggarty et al. 2023:SM 32). By this criterion, *bucca* cannot be a basic word (or default term) for ‘mouth’ through at least the early centuries CE. At some point it will attain this status by replacing *os* (Stefenelli 1962:77), but exactly when is unclear (Jud 1917:32). Since the Romanian descendant of *bucca*, *bucă*, means ‘cheek’, Proto-Romance /*bʊkk-a*/ had to have had both this sense and ‘mouth’ (Blank 1997:239–240, DÉRoM: s.v.).

In sum, the distribution of *os* and *bucca* provides no evidence for the sibling hypothesis. Chang et al. (2015:207) therefore conclude that “[l]iterary and colloquial Latin had the same basic vocabulary.” Their conclusion is consistent with that of Adams (2013:10), who writes that “[w]hat must be avoided ... is the notion that this language of the *uulgus* was a separate language system completely discrete from that of higher social groups.” By divorcing Classical and Vulgar Latin, the sibling hypothesis strips both of lexical variation (since variants are assigned to either Classical or Vulgar Latin). This is a critical problem, because such variation is not only intrinsic to language but also the antecedent to lexical change (Weinreich et al. 1968).

Finally, it is worth noting that the lexical variation observable in Latin during the Classical period is not tantamount to the diglossia that Ferguson (1959) described for Arabic in his seminal article. In the following passage, he illustrates the polarized distribution of high- and low-register lexical items:

[A] striking feature of diglossia is the existence of many paired items, one *H* [high, DG] one *L* [low, DG], referring to fairly common concepts frequently used in both *H* and *L*, where the range of meaning of the two items is roughly the same, and the use of one or the other immediately stamps the utterance or written sequence as *H* or *L*. For example, in Arabic the *H* word for ‘see’ is *ra’ā*, the *L* word is *šāf*. The word *ra’ā* never occurs in ordinary conversation and *šāf* is not used in normal written Arabic. (Ferguson 1959:334; italics in the original)

Ferguson (1959:334) draws a contrast between Arabic diglossia and lexical alternations in English, such as *purchase* and *buy*. He contends that such alternations do not qualify as diglossia because “both words may be written and both may be used in ordinary conversation.” Chang et al. (2015:207) make the selfsame point in regard to *os* and *bucca*. The sibling hypothesis predicts for Vulgar and Classical Latin diglossia characteristic of, e.g., Egyptian and Classical Arabic, but the lexical distributions necessary to motivate this view are absent, as Clackson (2016:6) also concludes: “There is indeed little to support the view that the language situation during either the Roman republic or empire was one of diglossia.”

6.4.2 Shared archaisms in Romance

Since Latin is one of the more archaic Indo-European languages, it shares many cognates with other ancient languages in the family. The word *os* ‘mouth’ introduced above is a case in point. It has cognates elsewhere in Indo-European (e.g., Sanskrit *as*, Old Irish *á*), but none in Romance. The complementary pattern, whereby an archaism is shared between Romance and an Indo-European language other than Latin appears not to exist. This apparent absence is important because the sibling hypothesis (but not the sampled-ancestor hypothesis) allows for precisely this possibility.

The *Dictionnaire Étymologique Roman* (*DÉRom*) contains about 500 lexical items reconstructed to Proto-Romance on the basis of the comparative method (Buchi and Schweickard 2015, Buchi and Schweickard 2016, Buchi and Schweickard 2020), most of which have correlates in written Latin. Some, however, do not (see, e.g., Dworkin 2016b:580–581 for a sample). Among the reconstructed forms lacking a Latin correlate, none is cognate with a word in another Indo-European language. For instance, */baß-a/ ‘salive visqueuse qui s’échappe de la bouche d’une personne ou de la gueule d’un animal’ (‘viscous saliva that escapes from a person’s mouth or an animal’s mouth’) can be reconstructed to Proto-Romance

and lacks a Latin correlate, but this absence is due to incomplete sampling. Given the Latin personal name *Baba* (e.g., *TLL*:s.v., *OLD*:s.v.) and the derived adjectives *babulus* ‘garrulous’ and *bavosus* ‘stupid’, the absence of the word in the Latin textual record is simply an accident (*DÉRom*:s.v. */baβ-a/).

Elsewhere, the absence of a Latin correlate of a reconstructed Proto-Romance form is probably real. For instance, beside the adjective */nɪtɪd-u/ ‘qui étant sans tache réfléchit la lumière; dont la surface n’a pas d’aspérités’ (‘that which being spotless reflects the light; whose surface lacks roughness’) there is also a syncopated variant */nɪtt-u/ that can be reconstructed to Proto-Romance. The unsyncopated lexical item has a Latin correspondent (*nitidus*), but the syncopated form does not. Crucially the syncopated form lacks cognates elsewhere in Indo-European.

If we consider a case of clear siblinghood elsewhere in Indo-European, the Tocharian languages, the data are markedly different from what we find in Latin and Romance. Tocharian A and Tocharian B each exhibit lexemes that are cognate with words in other Indo-European languages but not in the other Tocharian language. For instance, Tocharian A *wir* ‘young’ has cognates elsewhere in Indo-European (e.g., Latin *vir*), but not in Tocharian B. In the same vein, Tocharian B *maiwe* ‘young’ is cognate with, e.g., Old Norse *mjór* ‘small, thin’, but not with any words in Tocharian A. I am aware of no such pattern of cognate distribution within Latin and Romance, however. (Were such a cognate to exist, the sibling hypothesis would in fact have to be correct.) This absence further motivates the view that at least in terms of its lexicon Classical Latin is a sampled ancestor of the Romance languages.

7 Conclusion

This study makes a number of contributions to the history of Latin and the Romance languages and the investigation of divergence-time estimation more generally. In the domain of the former, I want to highlight three key results. First, this study offers divergence-time estimates for Proto-Romance that are consistent with the late hypothesis, in which the diversification of Latin most likely begins after 300 CE. No firm conclusions can be drawn as to whether this process begins before or after 476 CE. Second, my results support the view that the fall of the Roman Empire did not accelerate the diversification of the Romance languages, but it will be important for future work to examine this question with more data. Finally, my investigation is the first to infer Classical Latin as a sampled ancestor

of the Romance languages, which counters a widely and deeply held contrary view.

More generally, the results of this investigation make contributions to the following areas. First, my study offers an answer to the question of Heggarty (2021:382) quoted in Section 1 above: “All are Bayesians now, perhaps, but can one just pick one’s assumption to get one’s desired answer?” While the problems of model sensitivity are real, they can at least to some extent be mitigated with greater attention to model sensitivity, model adequacy, and model comparison.

Second, although the fossilized birth-death range models did not improve the marginal likelihood scores in this study, there may still be reasons to prefer them, since FBDR models treat data from ancient languages in a more accurate way than their FBD counterparts. In my experiment, the FBDR models also had a significant impact on the posterior probabilities of Latin as a sampled ancestor, so these models may also have a role to play in ongoing debates over linguistic ancestry.

Finally, the evidence from the Romance languages suggests that we have to come to terms with the fact that linguistic diversification can happen at different times in different components of language. It may well be the case that divergence-time estimates from sound change will not agree with those inferred from lexical data. If this is the case, it may be time to abandon the practice of inferring divergence-time estimates from basic vocabulary with node ages calibrated on the basis of sound change (or other non-lexical evidence).

Divergence-time estimation remains a challenging problem. It is important to bear in mind that there are still crucial aspects of linguistic history that our models cannot yet incorporate. First, current phylogenetic models do not adequately handle polymorphism. In the case of Latin and the Romance languages in particular, we know lexical polymorphism existed (e.g., the variation between *os* and *bucca* in the sense ‘mouth’). The inability to model linguistic variation could well affect divergence-time estimates (specifically, by pushing them farther back into the past). This is an aspect of linguistic history that future models will hopefully be able to incorporate. Second, although this study has considered a broad spectrum of models, it has certainly not explored all modeling possibilities. This is especially true for the clock models. Exploration of a wider array of clock models will reveal how robust the estimated divergence times presented in this paper are. Ultimately it will only be through the scrupulous avoidance of model

myopia (Wagenmakers et al. 2022) that more accurate absolute chronologies of the world's languages will be attained.

The data and code used for this paper are archived at [10.5281/zenodo.8060376](https://doi.org/10.5281/zenodo.8060376).

Acknowledgements

This study grew out of the virtual workshop on divergence-time estimation organized by Carl Rothfels and Mike May. I would like to extend my deepest thanks to Carl and Mike, as well as Ixchel González Ramírez, for generously answering my endless stream of questions. This study would never have come to fruition without their assistance. I am grateful to audiences at UCLA, Cornell, and the University of Cambridge for exceptionally insightful feedback on the oral version of this paper. Andrew Garrett and Nicolas Bertrand offered comments on an earlier draft that improved both its quality and readability. Éva Buchi and Steven Dworkin saved me from numerous blunders and Ron Kim kindly clarified some Tocharian facts for me. I could not have asked for a better research assistant than Elisa Migliaretti. Last but certainly not least, the suggestions of two reviewers greatly improved the paper. Fault for all remaining errors lies solely with me. This work used computational and storage services associated with the Hoffman2 Shared Cluster provided by UCLA Institute for Digital Research and Education's Research Technology Group.

Résumé

L'estimation de l'âge des langues est l'une des entreprises les plus importantes en linguistique historique, mais aussi l'une des plus difficiles. Au cours des vingt dernières années, les méthodes bayésiennes ont gagné en importance dans cette estimation. Cependant, plusieurs problématiques cruciales ont émergé, telles que la sensibilité des modèles, la dépendance des estimations de l'âge des racines à l'incertitude de l'estimation de nœuds internes, ainsi que la relation entre les langues anciennes et contemporaines. Cette étude aborde ces problématiques en se concentrant sur un cas particulièrement complexe au sein de la famille indo-européenne : l'évolution du latin vers les langues romanes. Les résultats obtenus vont dans le sens d'une approche gradualiste de la formation des langues romanes, en suggérant un début probable après l'an 300 de notre ère. De plus,

ils confirment clairement l'idée que le latin classique est un ancêtre direct des langues romanes, étant positionné sur la branche menant à ces langues.

Zusammenfassung

Die Bestimmung des Alters von Sprachen ist nicht nur eines der zentralen Anliegen der historischen Linguistik, sondern auch eines der anspruchsvollsten. In den letzten beiden Jahrzehnten haben sich bayesianische Methoden zur Altersschätzung von Sprachen immer weiter verbreitet. Mit dieser zunehmenden Verbreitung sind eine Reihe kritischer Fragen in den Vordergrund getreten. Hierzu zählen die Sensitivität der Modelle, die Abhängigkeit der Altersschätzungen der Grundsprachen von unsicheren Altersangaben der Innenknoten (d.h., der Untergruppen) sowie die Beziehung zwischen alten und modernen Sprachen. Diese Studie untersucht diese Fragen anhand eines besonders herausfordernden Falls innerhalb des Indogermanischen, nämlich der Aufspaltung des Lateinischen in die romanischen Sprachen. Die Ergebnisse dieser Untersuchung unterstützen eine allmähliche Entstehung der romanischen Sprachen, die höchstwahrscheinlich nach dem Jahr 300 n. Chr. begann. Sie liefern auch starke Belege dafür, dass das klassische Latein ein direkter Vorläufer der romanischen Sprachen ist (d.h., es liegt auf dem Zweig, der zu den romanischen Sprachen führt).

Author's address

David Goldstein
Department of Linguistics
University of California, Los Angeles
3125 Campbell Hall
Box 951543
Los Angeles, CA 90095-1543

dgoldstein@humnet.ucla.edu

References

- Adamik, Béla (2015). The periodization of Latin. An old question revisited. *Latin linguistics in the early 21st century. Acts of the 16th International Colloquium on Latin Linguistics, Uppsala, June 6th–11th, 2011*. Ed. by Gerd V. M. Haverling. Uppsala: Acta Universitatis Upsaliensis, 640–652.
- Adams, James N. (2007). *The regional diversification of Latin 200 BC–AD 600*. Cambridge: Cambridge University Press. DOI: [10.1017/CBO9780511482977](https://doi.org/10.1017/CBO9780511482977).
- (2013). *Social variation and the Latin language*. Cambridge: Cambridge University Press. DOI: [10.1017/CBO9780511843433](https://doi.org/10.1017/CBO9780511843433).
- (2016). *An anthology of informal Latin, 200 BC–AD 900. Fifty texts with translations and linguistic commentary*. Cambridge: Cambridge University Press. DOI: [10.1017/9781139626446](https://doi.org/10.1017/9781139626446).
- Agard, Frederick B. (1984). *A course in Romance linguistics. A diachronic view*. Vol. 2. Washington, D. C.: Georgetown University Press.
- André, Jacques (1991). *Le vocabulaire latin de l'anatomie*. Paris: Belles Lettres.
- Banniard, Michel (2013). The transition from Latin to the Romance languages. *The Cambridge history of the Romance languages. Contexts*. Ed. by Martin Maiden, John Charles Smith, and Adam Ledgeway. Vol. 2. Cambridge: Cambridge University Press, 6–56. DOI: [10.1017/CHO9781139019996.003](https://doi.org/10.1017/CHO9781139019996.003).
- Baxter, William H. and Alexis Manaster Ramer (2000). Beyond lumping and splitting. *Time depth in historical linguistics*. Ed. by April M. S. McMahon and R. Larry Trask. Vol. 1. Cambridge: McDonald Institute for Archaeological Research, 167–188.
- Becker, Lidia (2014). La protohistoire médiévale des langues romanes. *Manuel des langues romanes*. Ed. by Andre Klump, Johannes Kramer, and Aline Willems. Berlin: de Gruyter, 261–286. DOI: [10.1515/9783110302585.261](https://doi.org/10.1515/9783110302585.261).
- Blank, Andreas (1997). *Prinzipien des lexikalischen Bedeutungswandels am Beispiel der romanischen Sprachen*. Tübingen: Niemeyer.
- Bouckaert, Remco R., Philippe Lemey, Michael Dunn, Simon J. Greenhill, Alexander V. Alekseyenko, Alexei J. Drummond, Russell D. Gray, Marc A. Suchard, and Quentin D. Atkinson (2012). Mapping the origins and expansion of the Indo-European language family. *Science* 337.6097 (Aug. 2012), 957–960. DOI: [10.1126/science.1219669](https://doi.org/10.1126/science.1219669).
- (2013). Corrections and clarifications. *Science* 342.6165 (Dec. 2013), 957–960. DOI: [10.1126/science.342.6165.1446-a](https://doi.org/10.1126/science.342.6165.1446-a).

- Bromham, Lindell (2019). Six impossible things before breakfast. Assumptions, models, and belief in molecular dating. *Trends in Ecology & Evolution* 34.5 (May 2019), 474–486. DOI: [10.1016/j.tree.2019.01.017](https://doi.org/10.1016/j.tree.2019.01.017).
- (2022). Meaning and purpose. Using phylogenies to investigate human history and cultural evolution. *Biological Theory* 18, 284–302. DOI: [10.1007/s13752-022-00401-5](https://doi.org/10.1007/s13752-022-00401-5).
- Buchi, Éva, Carmen González Martín, Bianca Mertens, and Claire Schlienger (2015). L'étymologie de FAIM et de FAMINE revue dans le cadre du DÉRom. *Le français moderne* 83, 248–263.
- Buchi, Éva and Wolfgang Schweickard (2008–). *Dictionnaire Étymologique Roman*. Nancy: Analyse et Traitement Informatique de la Langue Française (ATILF).
- eds. (2015). *Dictionnaire Étymologique Roman. Genèse, méthodes et résultats*. Vol. 1. Beihefte zur Zeitschrift für romanische Philologie 381. Berlin: de Gruyter. DOI: [10.1515/9783110313482](https://doi.org/10.1515/9783110313482).
 - eds. (2016). *Dictionnaire Étymologique Roman. Pratique lexicographique et réflexions théoriques*. Vol. 2. Beihefte zur Zeitschrift für romanische Philologie 402. Berlin: de Gruyter. DOI: [10.1515/9783110453614](https://doi.org/10.1515/9783110453614).
 - eds. (2020). *Dictionnaire Étymologique Roman. Entre idioroman et protoroman*. Vol. 3. Beihefte zur Zeitschrift für romanische Philologie 443. Berlin: de Gruyter. DOI: [10.1515/9783110654264](https://doi.org/10.1515/9783110654264).
- Cano González, Ana María (2007). Del latín vulgar a los primeros romances. Aparición del romance en las escrituras. *Manual de lingüística románica*. Ed. by José Enrique Gargallo Gil and Maria Reina Bastardas i Rufat. Barcelona: Ariel, 81–119.
- Cathcart, Chundra Aroor (2018). Modeling linguistic evolution. A look under the hood. *Linguistics Vanguard* 4.1, 1–11. DOI: [10.1515/lingvan-2017-0043](https://doi.org/10.1515/lingvan-2017-0043).
- Cathcart, Chundra Aroor, Gerd Carling, Filip Larsson, Niklas Johansson, and Erich R. Round (2018). Areal pressure in grammatical evolution. An Indo-European case study. *Diachronica* 35.1, 1–34. DOI: [10.1075/dia.16035.cat](https://doi.org/10.1075/dia.16035.cat).
- Chang, Will, Chundra Aroor Cathcart, David P. Hall, and Andrew J. Garrett (2015). Ancestry-constrained phylogenetic analysis supports the Indo-European steppe hypothesis. *Language* 91.1 (Mar. 2015), 194–244. DOI: [10.1353/lan.2015.0005](https://doi.org/10.1353/lan.2015.0005).
- Clackson, James P. T. (2016). Latin as a source for the Romance languages. *The Oxford guide to the Romance languages*. Ed. by Adam Ledgeway and Martin Maiden. Oxford: Oxford University Press, 3–13. DOI: [10.1093/acprof:oso/9780199677108.003.0001](https://doi.org/10.1093/acprof:oso/9780199677108.003.0001).

- Coseriu, Eugenio (1954). *El llamado 'latín vulgar' y las primeras diferenciaciones romances*. Montevideo: Universidad de la República.
- (2008). *Lateinisch - Romanisch. Vorlesungen und Abhandlungen zum sogenannten Vulgärlatein und zur Entstehung der romanischen Sprachen*. Ed. by Hansbert Bertsch. Tübingen: Narr.
- Coseriu, Eugenio and Reinhard Meisterfeld (2003). *Geschichte der romanischen Sprachwissenschaft. Von den Anfängen bis 1492*. Vol. 1. Tübingen: Narr.
- Dardel, Robert de (1985). Le sarde représente-t-il un état précoce du roman commun? *Revue de Linguistique romane* 49, 263–269.
- Desnitskaja, Agnija (1982). Lat. *bucca*. Zur Verbreitungsgeschichte eines gemeinromanischen Wortes. *Festschrift für Johannes Hubschmid zum 65. Geburtstag. Beiträge zur allgemeinen, indogermanischen und romanischen Sprachwissenschaft*. Ed. by Otto Winkelmann and Maria Braisch. Bern: Francke, 237–245.
- Dworkin, Steven N. (2016a). Do Romanists need to reconstruct Proto-Romance? The case of the *Dictionnaire Étymologique Roman* (DÉRom) project. *Zeitschrift für romanische Philologie* 132.1, 1–19. DOI: [10.1515/zrp-2016-0001](https://doi.org/10.1515/zrp-2016-0001).
- (2016b). Lexical stability and shared lexicon. *The Oxford guide to the Romance languages*. Ed. by Adam Ledgeway and Martin Maiden. Oxford: Oxford University Press, 577–587. DOI: [10.1093/acprof:oso/9780199677108.003.0032](https://doi.org/10.1093/acprof:oso/9780199677108.003.0032).
- Ernout, Alfred and Antoine Meillet (1959). *Dictionnaire étymologique de la langue latine. Histoire des mots*. 4th ed. Paris: Klincksieck.
- Eskhult, Josef (2018). Vulgar Latin as an emergent concept in the Italian Renaissance (1435–1601). Its ancient and medieval prehistory and its emergence and development in Renaissance linguistic thought. *Journal of Latin Linguistics* 17.2, 191–230. DOI: [10.1515/joll-2018-0006](https://doi.org/10.1515/joll-2018-0006).
- Fabreti, Luiza Guimarães and Sebastian Höhna (2021). Convergence assessment for Bayesian phylogenetic analysis using MCMC simulation. *Methods in Ecology and Evolution* 13.1 (Jan. 2021), 77–90. DOI: [10.1111/2041-210X.13727](https://doi.org/10.1111/2041-210X.13727).
- Ferguson, Charles A. (1959). Diglossia. *Word* 15, 325–340. DOI: [10.1080/00437956.1959.11659702](https://doi.org/10.1080/00437956.1959.11659702).
- Ferri, Rolando and Philomen Probert (2010). Roman authors on colloquial language. *Colloquial and literary Latin*. Ed. by Eleanor Dickey and Anna Chahoud. Cambridge: Cambridge University Press, 12–41. DOI: [10.1017/cbo9780511763267.004](https://doi.org/10.1017/cbo9780511763267.004).
- Findell, Martin and Paul Heggarty (2023). English. Ed. by Paul Heggarty, Cormac Anderson, and Matthew Scarborough. Zenodo. DOI: [10.5281/zenodo.8089434](https://doi.org/10.5281/zenodo.8089434).

- Finegan, Edward (2009). English. *The world's major languages*. Ed. by Bernard Comrie. London: Routledge, 59–85.
- Garrett, Andrew J. (2006). Convergence in the formation of Indo-European subgroups. Phylogeny and chronology. *Phylogenetic methods and the prehistory of languages*. Ed. by Peter Forster and Colin A. Renfrew. Cambridge: McDonald Institute for Archaeological Research, 139–151.
- (2018). New perspectives on Indo-European phylogeny and chronology. *Proceedings of the American Philosophical Society* 162.1 (Mar. 2018), 25–38.
- Gavryushkina, Alexandra, Tracy A. Heath, Daniel T. Ksepka, Tanja Stadler, David Welch, and Alexei J. Drummond (2016). Bayesian total-evidence dating reveals the recent crown radiation of penguins. *Systematic Biology* 66.1 (Jan. 2016), 57–73. DOI: [10.1093/sysbio/syw060](https://doi.org/10.1093/sysbio/syw060).
- Glare, Peter G. W., ed. (2012). *Oxford Latin dictionary*. 2nd ed. Oxford: Oxford University Press.
- Gray, Russell D. and Quentin D. Atkinson (2003). Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature* 426.6965, 435–439. DOI: [10.1038/nature02029](https://doi.org/10.1038/nature02029).
- Gray, Russell D. and Fiona M. Jordan (2000). Language trees support the express-train sequence of Austronesian expansion. *Nature* 405.6790, 1052–1055. DOI: [10.1038/35016575](https://doi.org/10.1038/35016575).
- Greenberg, Joseph H. (2002). *Indo-European and its closest relatives*. 2 vols. Stanford: Stanford University Press.
- Hall Jr., Robert A. (1950). The reconstruction of Proto-Romance. *Language* 26.1 (Mar. 1950), 6–27. DOI: [10.2307/410406](https://doi.org/10.2307/410406).
- (1974). *External history of the Romance languages*. New York: Elsevier.
- (1976). *Proto-Romance phonology*. New York: Elsevier.
- Hammarström, Harald, Robert Forkel, Martin Haspelmath, and Sebastian Bank (2021). *Glottolog 4.5*. Jena: Max Planck Institute for the Science of Human History. DOI: [10.5281/zenodo.5772642](https://doi.org/10.5281/zenodo.5772642).
- Hartmann, Frederik (2023). *Germanic phylogeny*. Oxford: Oxford University Press. DOI: [10.1093/oso/9780198872733.001.0001](https://doi.org/10.1093/oso/9780198872733.001.0001).
- Heath, Tracy A., John P. Huelsenbeck, and Tanja Stadler (2014). The fossilized birth-death process for coherent calibration of divergence-time estimates. *Proceedings of the National Academy of Sciences of the United States of America* 111.29 (July 2014), E2957–E2966. DOI: [10.1073/pnas.1319091111](https://doi.org/10.1073/pnas.1319091111).

- Heggarty, Paul (2021). Cognacy databases and phylogenetic research on Indo-European. *Annual Review of Linguistics* 7, 371–394. DOI: [10.1146/annurev-linguistics-011619-030507](https://doi.org/10.1146/annurev-linguistics-011619-030507).
- Heggarty, Paul et al. (2023). Language trees with sampled ancestors support a hybrid model for the origin of Indo-European languages. *Science* 381.6656 (July 2023), eabg0818. DOI: [10.1126/science.abg0818](https://doi.org/10.1126/science.abg0818).
- Herman, József (2000). *Vulgar Latin*. Trans. by Roger Wright. University Park: The Pennsylvania State University Press.
- Hillis, David M., Tracy A. Heath, and Katherine St. John (2005). Analysis and visualization of tree space. *Systematic Biology* 54.3 (June 2005), 471–482. DOI: [10.1080/10635150590946961](https://doi.org/10.1080/10635150590946961).
- Höhna, Sebastian, Michael J. Landis, and Tracy A. Heath (2017). Phylogenetic inference using RevBayes. *Current Protocols in Bioinformatics* 57.1, 6.16.1–6.16.34. DOI: [10.1002/cpbi.22](https://doi.org/10.1002/cpbi.22).
- Höhna, Sebastian, Michael J. Landis, Tracy A. Heath, Bastien Boussau, Nicolas Lartillot, Brian R. Moore, John P. Huelsenbeck, and Fredrik Ronquist (2016). RevBayes. Bayesian phylogenetic inference using graphical models and an interactive model-specification language. *Systematic Biology* 65.4 (July 2016), 726–736. DOI: [10.1093/sysbio/syw021](https://doi.org/10.1093/sysbio/syw021).
- Hoinke, Ulrich (2003). Prinzipien der genealogischen Klassifikation der romanischen Sprachen. *Romanische Sprachgeschichte*. Ed. by Gerhard Ernst, Martin-Dietrich Gleßgen, Christian Schmitt, and Wolfgang Schweickard. Berlin: de Gruyter, 124–137. DOI: [10.1515/9783110146943.1.1.124](https://doi.org/10.1515/9783110146943.1.1.124).
- Huang, Wen, Guifang Zhou, Melissa Marchand, Jeremy R. Ash, David Morris, Paul van Dooren, Jeremy M. Brown, Kyle A. Gallivan, and Jim C. Wilgenbusch (2016). TreeScaper. Visualizing and extracting phylogenetic signal from sets of trees. *Molecular Biology and Evolution* 33.12 (Sept. 2016), 3314–3316. DOI: [10.1093/molbev/msw196](https://doi.org/10.1093/molbev/msw196).
- Jäger, Gerhard (2019). Computational historical linguistics. *Theoretical Linguistics* 45.3–4 (Oct. 2019), 151–182. DOI: [10.1515/tl-2019-0011](https://doi.org/10.1515/tl-2019-0011).
- Janda, Richard D. and Brian D. Joseph (2003). On language, change, and language change. Or, of history, linguistics, and historical linguistics. *The handbook of historical linguistics*. Ed. by Brian D. Joseph and Richard D. Janda. Malden, MA: Wiley-Blackwell, 3–180. DOI: [10.1002/9780470756393.ch](https://doi.org/10.1002/9780470756393.ch).
- Janson, Tore (1979). *Mechanisms of language change in Latin*. Stockholm: Almqvist & Wiksell.

- Jud, Jakob (1917). Probleme der altromanischen Wortgeographie. *Zeitschrift für romanische Philologie* 38, 1–98.
- Kassian, Aleksei, George Starostin, Anna V. Dybo, and Vasiliy Chernov (2010). The Swadesh wordlist. An attempt at semantic specification. *Journal of Language Relationship / Вопросы языкового родства* 4, 46–89.
- Kendall, David B. (1948). On the generalized “birth-and-death” process. *The Annals of Mathematical Statistics* 19.1, 1–15.
- Kitchen, Andrew, Christopher Ehret, Shiferaw Assefa, and Connie J. Mulligan (2009). Bayesian phylogenetic analysis of Semitic languages identifies an Early Bronze Age origin of Semitic in the Near East. *Proceedings of the Royal Society B: Biological Sciences* 276 (Apr. 2009), 2703–2710. DOI: [10.1098/rspb.2009.0408](https://doi.org/10.1098/rspb.2009.0408).
- Klinkenberg, Jean-Marie (1999). *Des langues romanes. Introduction aux études de linguistique romane*. 2nd ed. Paris: Duculot.
- Koile, Ezequiel, Simon J. Greenhill, Damián E. Blasi, Remco R. Bouckaert, and Russell D. Gray (2022). Phylogeographic analysis of the Bantu language expansion supports a rainforest route. *Proceedings of the National Academy of Sciences* 119.32 (Aug. 2022), e2112853119. DOI: [10.1073/pnas.2112853119](https://doi.org/10.1073/pnas.2112853119).
- Kühner, Mary K. and Joseph Felsenstein (1994). A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates. *Molecular Biology and Evolution* 11.3 (May 1994), 459–468. DOI: [10.1093/oxfordjournals.molbev.a040126](https://doi.org/10.1093/oxfordjournals.molbev.a040126).
- Leonard Clifford S., Jr. (1980). Comparative grammar. *Trends in Romance linguistics and philology. Romance comparative and historical linguistics*. Ed. by Rebecca R. Posner. The Hague: Mouton, 23–42. DOI: [10.1515/9783110814101.23](https://doi.org/10.1515/9783110814101.23).
- Lewis, Paul O. (2001). A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology* 50.6 (Nov. 2001), 913–925. DOI: [10.1080/106351501753462876](https://doi.org/10.1080/106351501753462876).
- Lloyd, Paul M. (1979). On the definition of “Vulgar Latin”. The eternal return. *Neuphilologische Mitteilungen* 80.2, 110–122. DOI: [10.2307/43343254](https://doi.org/10.2307/43343254).
- Löfstedt, Einar (1959). *Late Latin*. Oslo: Aschehoug.
- Loporcaro, Michele (2005). La sillabazione di muta cum liquida dal latino al romanzo. *Latin et langues romanes. Études de linguistique offertes à József Herman à l’occasion de son 80^{ème} anniversaire*. Ed. by Sándor Kiss, Luca Mondin, and Giampaolo Salvi. Tübingen: Niemeyer, 419–430.
- Mańczak, Witold (1974). La langue romane commune. latin vulgaire ou latin classique? *Revue Romane* 9, 218.
- (1977). *Le latin classique. Langue romane commune*. Wrocław: Ossolineum.

- Mańczak, Witold (1978). Le problème de la langue romane commune. *XIV Congresso Internazionale di Linguistica e Filologia Romanza*. Ed. by Alberto Varvaro. Vol. 2. Amsterdam: John Benjamins, 61–74. DOI: [10.1075/z.9.ii05lep](https://doi.org/10.1075/z.9.ii05lep).
- (1980). Do Romance languages originate from Pompeian Latin? *Lingua Posnaniensis* 23, 145–149.
 - (1987a). Origine des langues romanes. dogme et faits. *Latin vulgaire - latin tardif. actes du 1^{er} Colloque international sur le latin vulgaire et tardif (Pécs, 2–5 Septembre 1985)*. Ed. by József Herman. Vol. 1. Berlin: de Gruyter, 181–188. DOI: [10.1515/9783111652313.181](https://doi.org/10.1515/9783111652313.181).
 - (1987b). Où en est la discussion concernant l’origine des langues romanes? *Kwartalnik Neofilologiczny* 34, 257–263.
 - (1991). Nouvelle classification des langues romanes. *Revue Romane* 26, 14–23.
 - (1994a). La déclinaison romane provient-elle du protoroman ou du latin classique? *Vox Romanica* 53, 17–23.
 - (1994b). Protoroman et origine des langues romanes. *Lingvisticæ Investigationes. International Journal of Linguistics and Language Resources* 18.2, 365–369. DOI: [10.1075/li.18.2.07man](https://doi.org/10.1075/li.18.2.07man).
 - (1994c). Réactions diverses au problème de l’origine des langues romanes. *Revue Romane* 29, 123–129.
 - (1998). Le protoroman est-il une langue soeur du latin classique? *Latin vulgaire, latin tardif. Actes du 4^e Colloque International sur le latin vulgaire et tardif, Caen, 2–5 septembre 1994*. Ed. by Louis Callebat. Vol. 4. Hildesheim: Olms-Weidmann, 29–34.
 - (2006). Latin vulgaire et latin archaïque. *Latin vulgaire, Latin tardif. Actes du VII^{ème} Colloque international sur le latin vulgaire et tardif*. Ed. by Carmen Arias Abellán. Vol. 7. Sevilla: Universidad de Sevilla, 443–448.
 - (2007). Le latin vulgaire est-il une langue soeur du latin classique? *Actes du XXIV Congrès International de Linguistique et de Philologie Romanes*. Ed. by David A. Trotter. Vol. 2. Berlin: de Gruyter, 527–532. DOI: [10.1515/9783110923582.527](https://doi.org/10.1515/9783110923582.527).
 - (2013). Une linguistique romane sans latin vulgaire est-elle possible? *Actas del XXVI Congreso Internacional de Lingüística y de Filología Románicas*. Ed. by Emilie Casanova and Cesáreo Calvo. Vol. 4. Berlin: de Gruyter, 597–602. DOI: [10.1515/9783110299953.597](https://doi.org/10.1515/9783110299953.597).
- May, Michael R., Dori L. Contreras, Michael A. Sundue, Nathalie S. Nagalingum, Cindy V. Looy, and Carl J. Rothfels (2021). Inferring the total-evidence timescale of Marattialean fern evolution in the face of model sensitivity. *Systematic Biology* 70.6 (Mar. 2021), 1232–1255. DOI: [10.1093/sysbio/syab020](https://doi.org/10.1093/sysbio/syab020).

- Meier, Harri (1940). Über das Verhältnis der romanischen Sprachen zum Lateinischen. *Romanische Forschungen* 54.2, 165–201.
- Meyer-Lübke, Wilhelm (1920). *Einführung in das Studium der romanischen Sprachwissenschaft*. 3rd ed. Heidelberg: Carl Winter.
- Muller, Henri François (1921). When did Latin cease to be a spoken language in France? *Romanic Review* 12, 318–334.
- Murray, Robert W. and Naomi Cull (1994). Proto-Romance and the origin of the Romance languages. *Linguisticæ Investigationes. International Journal of Linguistics and Language Resources* 18.2, 371–376. DOI: [10.1075/li.18.2.08mur](https://doi.org/10.1075/li.18.2.08mur).
- Nicholls, Geoff K. and Russell D. Gray (2008). Dated ancestral trees from binary trait data and their application to the diversification of languages. *Journal of The Royal Statistical Society, Series B* 70.3, 545–566. DOI: [10.1111/j.1467-9868.2007.00648.x](https://doi.org/10.1111/j.1467-9868.2007.00648.x).
- Nichols, Johanna (1994). The spread of language around the Pacific Rim. *Evolutionary Anthropology* 3.6, 206–215. DOI: [10.1002/evan.1360030607](https://doi.org/10.1002/evan.1360030607).
- Palmer, Leonard R. (1954). *The Latin language*. London: Faber & Faber.
- Pawley, Andrew (2007). Locating Proto Oceanic. *The lexicon of Proto Oceanic: The culture and environment of ancestral Oceanic society. The physical environment*. Ed. by Malcolm D. Ross, Andrew Pawley, and Meredith Osmond. 2nd ed. Vol. 2. Canberra: Australian National University, 17–34.
- Posner, Rebecca R. (1996). *The Romance languages*. Cambridge: Cambridge University Press.
- Pulgram, Ernst (1958). *The tongues of Italy. Prehistory and history*. Cambridge, MA: Harvard University Press.
- Rama, Taraka (2018). Three tree priors and five datasets. A study of the effect of tree priors in Indo-European phylogenetics. *Language Dynamics and Change* 8.2, 182–218. DOI: [10.1163/22105832-00802005](https://doi.org/10.1163/22105832-00802005).
- Ringe, Donald A. (2022). What we can (and can't) learn from computational cladistics. *The Indo-European language family. A phylogenetic perspective*. Ed. by Thomas Olander. Cambridge: Cambridge University Press, 52–62. DOI: [10.1017/9781108758666.004](https://doi.org/10.1017/9781108758666.004).
- Ritchie, Andrew M. and Simon Y. W. Ho (2019). Influence of the tree prior and sampling scale on Bayesian phylogenetic estimates of the origin times of language families. *Journal of Language Evolution* 4.2 (July 2019), 1–16. DOI: [10.1093/jole/lzz005](https://doi.org/10.1093/jole/lzz005).
- Ronquist, Fredrik, Seraina Klopstein, Lars Vilhelmsen, Susanne Schulmeister, Debra L. Murray, and Alexandr P. Rasnitsyn (2012). A total-evidence approach to

- dating with fossils, applied to the early radiation of the hymenoptera. *Systematic Biology* 61.6 (July 2012), 973–999. DOI: [10.1093/sysbio/sys058](https://doi.org/10.1093/sysbio/sys058).
- Sagart, Laurent, Guillaume Jacques, Yunfan Lai, Robin J. Ryder, Valentin Thouzeau, Simon J. Greenhill, and Johann-Mattis List (2019). Dated language phylogenies shed light on the ancestry of Sino-Tibetan. *Proceedings of the National Academy of Sciences of the United States of America* 116.21 (May 2019), 10317–10322. DOI: [10.1073/pnas.1817972116](https://doi.org/10.1073/pnas.1817972116).
- Savelyev, Alexander and Martine Robbeets (2020). Bayesian phylolinguistics infers the internal structure and the time-depth of the Turkic language family. *Journal of Language Evolution* 5.1 (Jan. 2020), 39–53. DOI: [10.1093/jole/lzz010](https://doi.org/10.1093/jole/lzz010).
- Schuchardt, Hugo (1866). *Der Vokalismus des Vulgärlateins*. Vol. 1. Leipzig: Teubner. DOI: [10.1017/CBO9781139629478](https://doi.org/10.1017/CBO9781139629478).
- Selig, Maria (2008). *La naissance des langues romanes*. Avignon: Université de Avignon. DOI: [10.4000/books.eua.2298](https://doi.org/10.4000/books.eua.2298).
- Stadler, Tanja (2010). Sampling-through-time in birth-death trees. *Journal of Theoretical Biology* 267.3 (Dec. 2010), 396–404. DOI: [10.1016/j.jtbi.2010.09.010](https://doi.org/10.1016/j.jtbi.2010.09.010).
- Stadler, Tanja, Alexandra Gavryushkina, Rachel C. M. Warnock, Alexei J. Drummond, and Tracy A. Heath (2018). The fossilized birth-death model for the analysis of stratigraphic range data under different speciation modes. *Journal of Theoretical Biology* 447, 41–55. DOI: [10.1016/j.jtbi.2018.03.005](https://doi.org/10.1016/j.jtbi.2018.03.005).
- Stefenelli, Arnulf (1962). *Die Volkssprache im Werk des Petron im Hinblick auf die romanischen Sprachen*. Wien: Braumüller.
- (1996). Thesen zur Entstehung und Ausgliederung der romanischen Sprachen / Formation et fragmentation des langues romanes. *Lexikon der romanistischen Linguistik. Latein und Romanisch: Historisch-vergleichende Grammatik der romanischen Sprachen*. Ed. by Günter Holtus, Michael Metzeltin, and Christian Schmitt. Vol. 2. 1. Berlin: de Gruyter, 73–90. DOI: [10.1515/9783110938364.73](https://doi.org/10.1515/9783110938364.73).
 - (2003). Die lateinische Basis der romanischen Sprachen. *Romanische Sprachgeschichte*. Ed. by Gerhard Ernst, Martin-Dietrich Gleßgen, Christian Schmitt, and Wolfgang Schweickard. Berlin: de Gruyter, 530–544. DOI: [10.1515/9783110146943.1.5.530](https://doi.org/10.1515/9783110146943.1.5.530).
- Straka, Georges (1953). Observations sur la chronologie et les dates de quelques modifications phonétiques en roman et en français pré littéraire. *Revue des langues romanes* 71, 247–307.
- (1956). La dislocation linguistique de la Romania et la formation des langues romanes à la lumière de la chronologie relative des changements phonétiques. *Revue de Linguistique romane* 20, 249–267.

- Swiggers, Pierre (2001). De Prague à Strasbourg. Phonétique et phonologie du français chez Georges Gougenheim et Georges Straka. *Modèles linguistiques* 43.3 (Jan. 2001), 21–44. DOI: [10.4000/ml.1459](https://doi.org/10.4000/ml.1459).
- TLL (1900–). *Thesaurus linguae Latinae*. Berlin: de Gruyter.
- Tovar, Antonio (1964). A research report on Vulgar Latin and its local variations. *Kratylos* 9, 113–134.
- Trudgill, Peter (2011). *Sociolinguistic typology. Social determinants of linguistic complexity*. Oxford: Oxford University Press.
- Väänänen, Veikko (1981). *Introduction au latin vulgaire*. 3rd ed. Paris: Klincksieck.
- (1983). Le problème de la diversification du latin. *Aufstieg und Niedergang der römischen Welt. Sprache und Literatur (Sprachen und Schriften)*. Ed. by Wolfgang Haase. Vol. 29. Berlin: de Gruyter, 480–505. DOI: [10.1515/9783110847024-009](https://doi.org/10.1515/9783110847024-009).
- Vallejo, José M. (2012). Del proto-indoeuropeo al proto-romance. *Romance Philology* 66.2, 449–467.
- Varvaro, Alberto (1991). Latin and Romance. Fragmentation or restructuring? *Latin and the Romance languages in the early Middle Ages*. Ed. by Roger Wright. London: Routledge, 44–51.
- (2013). Latin and the making of the Romance languages. *The Cambridge history of the Romance languages. Contexts*. Ed. by Martin Maiden, John Charles Smith, and Adam Ledgeway. Vol. 2. Cambridge: Cambridge University Press, 6–56. DOI: [10.1017/CHO9781139019996.002](https://doi.org/10.1017/CHO9781139019996.002).
- Versteegh, Kees (2022). The ghost of Vulgar Latin. History of a misnomer. *Historiographia Linguistica* 48.2–3 (Dec. 2022), 205–227. DOI: [10.1075/hl.00091.ver](https://doi.org/10.1075/hl.00091.ver).
- Vincent, Nigel (2016). Continuity and change from Latin to Romance. *Early and late Latin*. Ed. by James N. Adams and Nigel Vincent. Cambridge: Cambridge University Press, 1–13. DOI: [10.1017/cbo9781316450826.002](https://doi.org/10.1017/cbo9781316450826.002).
- Wagenmakers, Eric-Jan, Alexandra Sarafoglou, and Balázs Aczel (2022). One statistical analysis must not rule them all. *Nature* 605 (May 2022), 423–425.
- Warnock, Rachel C. M. and April M. Wright (2020). *Understanding the tripartite approach to Bayesian divergence time estimation*. Cambridge: Cambridge University Press. DOI: [10.1017/9781108954365](https://doi.org/10.1017/9781108954365).
- Weinreich, Uriel, William Labov, and Marvin Herzog (1968). Empirical foundations for a theory of language change. *Directions for historical linguistics*. Ed. by Winfred P. Lehmann and Yakov Malkiel. Austin: University of Texas Press, 95–188.
- Weiss, Michael (2020). *Outline of the historical and comparative grammar of Latin*. 2nd ed. Ann Arbor: Beech Stave Press.

- Wittoch, Zdeněk (1984). La naissance des langues romanes, le latin vulgaire et le latin classique. *Philologica Pragensia* 27, 41–47.
- Wright, April M., David W. Bapst, Joëlle Barido-Sottani, and Rachel C. M. Warnock (2022). Integrating fossil observations into phylogenetics using the fossilized birth-death model. *Annual Review of Ecology, Evolution, and Systematics* 53.1 (Aug. 2022), 251–273. DOI: [10.1146/annurev-ecolsys-102220-030855](https://doi.org/10.1146/annurev-ecolsys-102220-030855).
- Wright, Roger (2011). Romance languages as a source for spoken Latin. *A companion to the Latin language*. Ed. by James P. T. Clackson. Malden, MA: Blackwell, 59–79. DOI: [10.1002/9781444343397.ch5](https://doi.org/10.1002/9781444343397.ch5).
- Yang, Ziheng and Bruce Rannala (1997). Bayesian phylogenetic inference using DNA sequences. A Markov Chain Monte Carlo method. *Molecular Biology and Evolution* 14.7 (July 1997), 717–724. DOI: [10.1093/oxfordjournals.molbev.a025811](https://doi.org/10.1093/oxfordjournals.molbev.a025811).