frontiers in Psychology

ORIGINAL RESEARCH
published: 18 December 2020
doi: 10.3389/fpsyg.2020.488871

Formal Syntax and Deep History

*Andrea Ceolin[1], Cristina Guardiano[2], Monica Alexandrina Irimia[2] and Giuseppe Longobardi[3]**

[1] Department of Linguistics, University of Pennsylvania, Philadelphia, PA, United States, [2] Dipartimento di Comunicazione ed Economia, Università di Modena e Reggio Emilia, Reggio Emilia, Italy, [3] Department of Language and Linguistic Science, University of York, York, United Kingdom

We show that, contrary to long-standing assumptions, syntactic traits, modeled here within the generative biolinguistic framework, provide insights into deep-time language history. To support this claim, we have encoded the diversity of nominal structures using 94 universally definable binary parameters, set in 69 languages spanning across up to 13 traditionally irreducible Eurasian families. We found a phylogenetic signal that distinguishes all such families and matches the family-internal tree topologies that are safely established through classical etymological methods and datasets. We have retrieved "near-perfect" phylogenies, which are essentially immune to homoplastic disruption and only moderately influenced by horizontal convergence, two factors that instead severely affect more externalized linguistic features, like sound inventories. This result allows us to draw some preliminary inferences about plausible/implausible cross-family classifications; it also provides a new source of evidence for testing the representation of diversity in syntactic theories.

Keywords: phylogenetics, formal syntax, parameters, language reconstruction, biolinguistics

Edited by:
Antonio Benítez-Burraco, University of Seville, Spain

Reviewed by:
Claire Bowern, Yale University, United States
Johann-Mattis List, Max Planck Institute for the Science of Human History (MPI-SHH), Germany

***Correspondence:**
Giuseppe Longobardi
giuseppe.longobardi@york.ac.uk

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *01 August 2019*
Accepted: *24 August 2020*
Published: *18 December 2020*

Citation:
Ceolin A, Guardiano C, Irimia MA and Longobardi G (2020) Formal Syntax and Deep History. Front. Psychol. 11:488871. doi: 10.3389/fpsyg.2020.488871

INTRODUCTION

The Conceptual Roots of Parametric Comparison

A theory of human language aiming to be part of cognitive science (see Everaert et al., 2015) should try to argue that the structural representations it proposes are: (i) learnable under realistic acquisition conditions; (ii) historically transmitted under the conditions normally expected for the propagation of culturally selected knowledge. The classical theory of generative grammars set itself (i), i.e. the *ontogenetics* of grammars, as its main standard (explanatory adequacy, Chomsky, 1964). We believe that (ii), the *phylogenetics* of grammars, may also provide crucial evidence for the problem of realistic grammatical representations; thus, we test a theory of syntactic diversity inspired by minimalist biolinguistics precisely against the standard in (ii).

Our Goals

We explore the relationship between the historical signal of different levels of linguistic analysis (referred to as *Humboldt's* problem by Longobardi and Guardiano, 2009, and as the problem of the *fabric* of human history by Gray et al., 2010; also see Greenhill et al., 2017). For this purpose, we especially try to assess the historical tree-likeness (the problem of the *shape*, in Gray et al.'s 2010

terms) of syntax. In pursuing these goals, we combine some methods of the quantitative revolution in phylogenetic linguistics[1] with the deductive approach to syntactic diversity that has emerged since Chomsky (1981), and we ask if formal syntactic differences can serve as effective characters for taxonomic purposes, contrary to a long line of skepticism.

Syntax, Cognitive Science, and Historical Taxonomy

Over the past decades, increased attention has been paid to deep-time investigations of human history.[2] A central role in this trend has been played by developments in biology, prompted by the use of genetic evidence for reconstructing the diversification of populations.[3] In the meantime, the rise of cognitive science has produced important breakthroughs in the understanding of human mind as a system of symbolic computations, instantiated e.g., by rules of natural language syntax, most notably in the so-called formal biolinguistic framework.[4] Against this background, a broad methodological question is: can modern cognitive science side with biological anthropology in contributing to a science of long-range history?

As a matter of fact, the study of language pioneered deep historical investigation: linguistic taxonomies and the discovery of remote proto-languages have crucially contributed to pushing back the time limits of human history and prehistory. However, the levels of linguistic analysis that have best substantiated recent cognitive and computational theories have not yet played a part in this enterprise, and the practitioners of formal grammar and phylogenetic linguistics have formed nearly disjoint communities of scholars. In particular, syntax has never been seriously used for reconstructing phylogenies and proto-languages. Morpurgo-Davies (1992/2014) stresses how the earliest researchers[5] already rejected syntax as a tool for language phylogeny on the grounds that it would entail the presence of similar features in languages that can be easily proved to be unrelated, i.e., that it would be subject to pervasive homoplasy.[6] Since the late 18th century, this assumption appears not to have changed, even after Kayne (1975) laid the basis of modern comparative syntax. Consider, for instance the following statement:

(1) "In fact it is quite possible – even likely – that English grammars might be more similar to grammars with which there is less historical connection. From this perspective, looking at the parameters in the current linguistic literature, English grammars may be more similar to Italian than to German, and French grammars may be more similar to German than to Spanish. *There is no reason to believe that structural similarity should be even an approximate function of historical relatedness*..."
(Anderson and Lightfoot, 2002, pp. 8–9: our italic)

The Historical Signal of Syntax

Positions along these lines are widely held in the field (cf. Newmeyer, 2005; Anderson, 2012, a.o.).[7]

Interestingly, at a small scale it is commonly accepted that syntactic variability aggregates across individuals in time and space.[8] For instance, an important facet of the logical problem of language acquisition (Hornstein and Lightfoot, 1981; Lightfoot, 1982, a.o.) makes crucial reference to this kind of similarity among I-languages (how do the children of *a community* converge on the *same* target grammar in certain subtle details, in spite of individual and idiosyncratic primary data?).

It is at a larger scale (e.g., of Romance or Indo-European) that this simple assumption becomes progressively controversial, neglected or altogether rejected, for non-obvious reasons. Normally, culturally transmitted phenomena leave a longer-term historical trace (e.g., some notion of "common Romance vocabulary"). Therefore, that even syntax does so should be the null hypothesis.

It is true that individual syntactic changes may be "catastrophic" and unpredictable: this discovery (Lightfoot, 1979, 1997, 2002, a.o.)[9] has been very instrumental in overcoming the epistemological pitfalls of classical linguistic historicism and reducing inquiry to its appropriate "molecular" units: individual parameters. Yet, if several syntactic parameters are considered at the same time, a historical signal might well emerge. Notice that if such a signal were completely irretrievable, then someone could even argue that generative syntax is inadequate as a model

[1] Ringe et al. (2002); Gray and Atkinson (2003); McMahon and McMahon (2005); McMahon (2010), and the stream of subsequent work.

[2] E.g., Braudel (1958) and subsequent work, Diamond (1997); Smail (2008).

[3] Cavalli-Sforza et al. (1994), as well as subsequent work.

[4] Cf. Hauser et al. (2002); Boeckx and Piattelli-Palmarini (2005); Di Sciullo and Boeckx (2011); Berwick and Chomsky (2015), a.o.; for some specific applications to language diversity see Biberauer (2008); Karimi and Piattelli-Palmarini (2017); Roberts (2019) and much cited literature.

[5] E.g., Kraus (1787; see Kaltz, 1985 for details), Adelung (1806–1817) or Balbi (1826a,b). "Balbi (1826a, xlii f., note) ... stated that grammatical comparison cannot be used to establish kinship and quoted as an example the fact that English and Omagua, a language of Brazil, were clearly not related, though their grammars contrasted in similar ways with the grammars of other languages in their families (ibid., 28)." (Morpurgo-Davies, 1992/2014, p. 51).

[6] Morpurgo-Davies (1992/2014) points out that even Hervás (1778–1787, 1800–1805) or Gyarmathi (1799), though interested in grammatical features, did not go beyond the examination of traits such as declensions, conjugations, degrees of comparison of adjectives, suffixes used to mark comparable functions, pronouns, etc., i.e., the lexically arbitrary coding of form-meaning in functional elements. She notices that later and more established names in comparative reconstruction (Schleicher, for example) equally considered that only phonology and morphology were relevant for historically oriented work.

[7] After the programmatic concepts in Klima (1964, 1965), the question of the potential of grammatical features for historical relatedness was not fully resumed until Nichols (1992); Longobardi (2003); Dunn et al. (2005); Guardiano and Longobardi (2005); Wichmann and Saunders (2007), and a first systematic use of formal syntactic traits was only attempted in Longobardi and Guardiano (2009). An interesting exception regarding syntax as an indicator of relatedness is Chapin (1974), kindly pointed out to us by R. Kayne.

[8] It is normally assumed to be like further features of language and culture, and unlike certain other cognitive faculties (there is a sense to the notion "French syntax," no less than to "French vocabulary," or "French cuisine," though not to "French memory" or "French visual perception").

[9] All this foundational work of Lightfoot's on diachronic syntax, as well as that inspired by Kroch (1989 and subsequent: especially see Pintzuk and Kroch, 1995 on dating) has not been concerned with relatedness, as noted. Nonetheless, this line, along with Kayne's (1975, 2000 and subsequent) insights on comparative syntax, has been essential for conceiving of generative grammars as tools of historical knowledge.

of language transmission (i.e., acquisition across generations), hence as a realistic cognitive model *tout court*.[10]

Syntactic Data and Taxonomic Problems

Two general problems of linguistic taxonomic methods (cf. Guardiano et al., 2020) are especially relevant for our purposes:

(2) a. The *globality* problem
b. The *ultralocality* problem

(2)a refers to the fact that comparative procedures may aspire to long-range or, ideally, global coverage; thus, they should rely on universally definable taxonomic characters, that can apply to any set of languages. (2)b is the converse issue: even if some type of characters does not saturate at the macro-comparative level, it could still fail in resolution when applied to discriminate close dialects, or just fail to correlate altogether with the reduction of their differences in other linguistic aspects.

Even if promising advances in cross-family comparison have recently been made (Jäger, 2015), procedures based on vocabulary data and lexical arbitrariness are generally not appropriate for (2)a, because they mainly rely on family-internal etymologies.[11] Therefore, the development of a non-lexical method is a theoretical *eldorado* in the pursuit of deep language history (Nichols, 1992). Parameters in the theory of generative grammars should lend themselves well to this goal, as they are grounded in a model of the language faculty explicitly designed in universal terms.

Thus, we focused on: (i) a set of syntactic traits modeled along the lines of Longobardi and Guardiano's (2009) Parametric Comparison Method (PCM) and including macro-, meso-, and micro-parameters (Biberauer and Roberts, 2017; Roberts, 2019);[12] (ii) a language sample to test these traits against family-wide taxonomies, but also with respect to cross-family and dialect comparison.

Importantly, we assumed some idealizations about the adopted comparative characters:[13]

(3) a. *Modularity*: they are all purely syntactic traits, drawn from a single module of syntax (the internal structure of nominal phrases);
b. *Deductivity*: they are all coded as abstract primitives of the generative device;
c. *Interdependence*: their known and plausible dependencies are spelt out and built into the parametric structure.

These three properties of our input data are different from those attributed to the structural traits recently used to address similar issues, e.g., in Greenhill et al. (2017). We will explore some consequences of using traits with these three properties for the pursuit of long-range comparison (cf. Section "Input data and phylogenetic results").

MATERIALS AND METHODS

Parameters and Schemata

In classical Principles-and-Parameters models (Chomsky, 1981) it was assumed that variability in human grammars is reducible to a finite list of binary choices, extensionally present in every speaker's mind at the initial state of language acquisition. This "preformistic"[14] view has been criticized recently. In particular, it has been associated with an implausible model of language learnability, as it imposes too heavy a burden on the initial state of the human mind.[15]

Here we 'presuppose' a model of variation which does not necessarily rely on lists of parameters, but rather sketches a universal set of simple possible syntactic relations (i.e., *schemata*: Longobardi, 2005, 2014, 2017; Gianollo et al., 2008); whether, in each language, they apply or not to specific categories and features determines a number of binary choices epigenetically rather than preformistically. This minimalist parametric model (Principles and Schemata in Longobardi's, 2005 terms) has the effect of intensionally defining parameter lists with their familiar properties (including universal definition and ease of value collation for comparative purposes: Roberts, 1998), without attributing such lists extensionally to the common initial state of the language faculty.

Our parameters are formally coded using two symbols, "+" and "−". Specifically, we adopt the system proposed in Crisma et al. (2020): cognitively, just "+" is viewed as an addition to the initial state of the mind. The "−" state of a parameter is not an entity attributed to the speaker's mind, though it is used by the PCM as a symbol to code a difference with "+" at that parameter in another language.

We call "manifestation(s)" the empirical evidence that sets a given parameter. Most parameters have a clustering structure, i.e., are associated with a set of co-varying surface manifestations,[16] with different degrees of saliency. As a consequence of such clustering structure, identifying just one core manifestation (a trigger or *p-expression* in Clark and Roberts', 1993 sense) per parameter will suffice for the learner (and the linguist) to set the parameter to "+." If no relevant

[10] In fact, there have been sporadic, though insightful, suggestions that syntax may be even more conservative than other linguistic levels, at least as a source of primitive diachronic change. This is basically the content of Keenan's (2002, 2009) notion of Inertia, i.e., the hypothesis that linguistic structure tends to stay stable through time "unless acted upon by an outside force or DECAY" (Keenan, 2009, p. 18). "Decay" here refers to phonological erosion and lexical-semantic impoverishment. A slightly more articulated definition of the Inertia hypothesis has been adopted in Longobardi (2001): "...syntactic change should not arise, unless it can be shown to be *caused*—that is, to be a well-motivated consequence of other types of change (phonological changes and semantic changes, including the appearance/disappearance of whole lexical items) or, recursively, of other syntactic changes..." (Longobardi, 2001, p. 278).

[11] For progress in the automatization of lexical comparative methods also see List (2014).

[12] Crucially, we do not use nano-parameters, which involve extensional definitions in terms of lists of lexical items.

[13] See Longobardi and Guardiano (2009) for an extensive justification of these methodological assumptions.

[14] In the terms of early modern biology.

[15] See especially Boeckx and Leivadá (2014); Fodor and Sakas (2017); Lightfoot (2017), and the various problems summed up in Longobardi (2017).

[16] Rizzi (1978, 1982); Taraldsen (1980); Chomsky (1981); Kroch (1989); Kayne (2000), a.o.

manifestation for "+" is present in the data, the grammar's *default* state does not change.

P-expressions are by definition *positive* evidence, i.e., grammatical phrases of a language. In the formulation of the parameters we made sure that the non-default value "+" can be set in all the languages from positive evidence in this sense.

The Syntactic Dataset

In this article, we used the 94 binary syntactic nominal parameters identified in Crisma et al. (2020) by a set of YES/NO questions which define the manifestations of each of them.[17] They are set in 69 contemporary Eurasian languages from up to 13 traditionally irreducible families.[18] Full information about the languages and the parameter states is available in **Supplementary Table 1** and **Supplementary Figure 1**.

The languages were chosen to investigate three different levels of historical depth: the relations of the deepest established families, their internal articulation, and dialect microvariation. To explore the latter, we rely on the sample of Romance[19] and Greek[20] dialects included in the dataset.

Some Numerical Properties of the Syntactic Data

The parameters of our system display an intricate implicational structure (Guardiano and Longobardi, 2017), i.e., many parameter states turn out to be predictable, or completely irrelevant, given the states of other parameters.[21] In the dataset used in this article, 2925 states out of 94 × 69 (= 6486) are null, perhaps the most impressive instantiation of the insight (sometimes attributed to Meillet, but cf. Toman, 1987) that natural languages are "un système où tout se tient." The effect of such null states on the number of possible languages has become measurable since Bortolussi et al. (2011), proving to reduce it by several orders of magnitude (cf. Section "Possible Languages" in **Supplementary Material**).

A related numerical feature of the syntactic dataset is that in a system with two non-null states ("+" and "−") and a null state (coded as "0" and representing no independent information) the only relevant comparisons for a pair of languages are provided by parameters for which neither language displays a "0": namely an identity ("+/+" or "−/−") or a difference ("+/−" or viceversa). The average number of parameters for each language pair that does not display "0" in either language is 39 (in the range of 14 to 66). Thus, the historical signal which can be found in this dataset will be generated by an average of taxonomic characters no higher than 39 (a figure much lower than that of the taxonomic units investigated)[22]: if a significant signal is indeed found, this will suggest that the selected characters have a high degree of resolution.

From a practical viewpoint, it is also important to stress that, thanks to the structure of the parameter system, in order to fill in the states of the 94 parameters for each language it is only necessary to find positive evidence for the "+" values; this is so because "0" is totally deducible information and "−" is a default state. In our dataset the total amount of "+" is 1386, thus, the mean is 20 "+" per language; the median is also 20. Hence, the amount of parameter values which must be set from positive empirical evidence is only about one quarter of the whole parameter list.[23]

Taxonomic and Phylogenetic Methods

We have performed a series of experiments using some standard computational tools, although none of them was conceived for − or specifically adjusted to − syntactic, rather than biological or lexical data. Such tools belong to two major types: distance-based and character-based programs.

Distance-Based Methods

We used three distance-based tools: heatmaps,[24] PCoAs,[25] and UPGMA phylogenetic trees.[26]

Heatmaps can be used to identify clusters in a distance matrix: in the heatmap, each cell (corresponding to a language pair) is assigned a color according to its distance value; then, through a hierarchical clustering algorithm, cells can be arranged on the basis of their color: language pairs which share small distances are arranged along the diagonal of the square matrix.

Principal Coordinate Analyses (PCoAs) represent a distance matrix on a Cartesian plane by plotting the taxa on a bidimensional space, using a linear transformation of the distance matrix.

[17] Several parameters concerning the Determiner category and Genitive Case used in this article are analyzed in syntactic detail in Crisma and Longobardi (in press) and in Crisma et al. (to appear). Notice, however, that, in order to conform to the requirement that the "+" state must be settable on the basis of positive evidence only, the formulation of some parameters here can have reversed the "−" and "+" values (see Crisma et al., 2020).

[18] Considering Turkic, Mongolic, Tungusic, Japanese, and Korean as separate families, since there is no consensus in the field about their genealogical relatedness (see e.g., Ceolin, 2019).

[19] The Italo-Romance dialects of our sample belong to three major groups (Pellegrini, 1977; Loporcaro, 2009): (1) Gallo-Italic: Casalasco (Vezzosi, 2019), Reggio Emilia, Parma. (2) Extreme southern: Reggio Calabria (Southern Calabria dialects are usually clustered with Sicilian dialects), Salentino (traditionally classified as an Extreme southern dialect but geographically separated from the rest of the Extreme group, while it has enjoyed an uninterrupted road connection to Rome and Naples since the Via Appia was built between 312 and 264 BC), two dialects from Sicily (Ragusa and Mussomeli; see Guardiano et al., 2016). (3) Upper southern: Teramano, Campano, Barese, and Northern Calabrese. The latter belongs to a particularly conservative area (Lausberg, 1939) characterized by morpho-phonological features which single it out from the rest of Italian dialects (Rohlfs, 1972; Rensch, 1973; Fanciullo, 1988, 1997; Martino, 1991; Romito et al., 1996, a.o. and also Silvestri, 2013 and Guardiano et al., 2016 about its nominal syntax).

[20] In the Greek group, we selected the following varieties: Standard Modern Greek, Cypriot Greek, and three varieties of Italiot Greek (one from Salento and two from Calabria which display different degrees of conservativity, Guardiano and Stavrou, 2014, 2019, 2020; Guardiano et al., 2016).

[21] Also see Baker (2001); Roberts (2019), a.o.

[22] This figure goes down to 20 if only "+/+" is computed as an identity: cf. footnote 29.

[23] The language that has the highest amount of "+" is Romanian (29), while the language with the smallest amount of "+" is Cantonese (9).

[24] Eisen et al. (1998); Cordoni et al. (2016).

[25] Davis (1986); Podani and Miklos (2002).

[26] Sneath and Sokal (1973).

The distance-based algorithm that is typically used to generate phylogenetic trees from a distance matrix is Neighbor-Joining.[27] Previous work on syntactic data showed that identifying a root and imposing the same branch length between a root and the leaves (i.e., assuming a molecular clock) through an updated version of Neighbor-Joining (the UPGMA algorithm) improves the classification.[28] Hence, for our distance-based phylogenetic experiments, we adopted UPGMA (using the package PHYLIP, Felsenstein, 2005).

Measuring Syntactic Distances

One of the main challenges about our data is dealing with null characters ("0"). Distance-based methods allow us to do so in a simple way: whenever one of the languages of a pair has a "0" for a certain parameter (cf. Section "Some numerical properties of the syntactic data"), we can just ignore the parameter in calculating the distance of the pair. To deal with this problem, we first normalized a standard distance metric (Hamming, 1950) by dividing, for each pair of languages, the number of differences by the sum of their identities and differences.

Our background parameter theory (cf. Section "Parameters and Schemata") assumes that, of the two potential states of a parameter, the value "−" instantiates a default state: thus, identities on two "−" should *a priori* be less marked than identities on two "+." In other words, the former could be less likely than the latter as shared innovations in the phylogenetic history. However, it is difficult to assess the actual weight of the potentially less informative "−/−" correspondences: therefore, we explored the radical idealization of counting as identities only the "+/+" ones. This amounts to using a Jaccard (1901) metric:[29]

$$(4)\quad \Delta\ \text{Jaccard}\,(A, B) = [N_{-+} + N_{+-}]/[N_{-+} + N_{+-} + N_{++}]$$

where N_{XY} indicates the number of positions where the string A has value X and B has Y.

To measure the impact of the idealization, we performed experiments both through a Jaccard distance and a normalized Hamming distance (in which "+/+" and "−/−" are both counted as valid identities) and the results are slightly worse for Hamming[30] (cf. Section "Phylogenetic Analysis – Hamming Distances" in **Supplementary Material**); therefore, we decided to simply proceed with the more restrictive Jaccard formula.

The heatmap, the PCoAs and the phylogenetic tree shown in **Figure 3** were generated from the Jaccard distance matrix inferred from the parametric characters of **Supplementary Figure 1**.

Character-Based Methods

Character-based methods were specifically devised to reconstruct the sequence of changes in the character states of a dataset.[31] Character-based phylogenetic methods have mostly been used to calculate linguistic splits and dates.[32] In particular, Bayesian inference has been recently implemented to evaluate the probability of different evolutionary models: for instance, whether the rate of change is uniform across branches and across characters, or whether it can be modeled according to some mathematical distribution. Evolutionary models are then used to generate phylogenetic trees. We employed the software BEAST 2 (Bouckaert et al., 2019), which is the most up-to-date tool to perform Bayesian phylogenetic analysis.

Finally, we calculated two tree-likeness metrics, Δ-scores and Q-residuals,[33] from a network generated through the algorithm NeighborNet, from SplitsTree.[34] These measures estimate the robustness of the vertical signal, and indicate which taxa are weaker due to the possible presence of horizontal convergence or homoplasy.

Some Problems With Current Methods

Both methods require some idealization about the data structure, and therefore either methodological choice can be expected to misrepresent some aspect of the information contained in the dataset.

When using distance-based algorithms, reducing all pairs of strings (languages) in the dataset to a distance matrix implies that the exact position of identities and differences between them becomes irretrievable. Moreover, the choice of distance metrics has an impact on how differences are weighted against identities.

Character-based algorithms, on the contrary, are the closest automatic analog to the linguists' consolidated procedure of reconstructing all ancestral states (e.g., sounds and etymologies) and changes, and of postulating taxa on this basis (Greenhill et al., 2020); however, a straightforward exploitation of their potential for our data is still partly hampered by at least two features of these algorithms.

First, these methods assume character states and their changes to be independent, an assumption which is not true in our case. Therefore, they do not offer any intuitive solution to deal with implied values ("0"), because they were not devised to incorporate interdependence among characters. Coding the state "0" as a third, independent value, would be an arbitrary manipulation of the data, because "0" represents completely predictable information rather than additional information or points of uncertainty.[35] To mitigate this problem, we coded the

[27] Saitou and Nei (1987).

[28] Rigon (2009); Longobardi et al. (2013).

[29] The average number of parameters that are comparable in our dataset according to the Jaccard metric (i.e., parameters where either language displays a "+" without the other displaying a "0") turned out to be 20, with a range between 7 and 30.

[30] Cf. Franzoi et al. (2020) for an attempt to develop metrics alternative to Hamming and Jaccard in order to capture structural dependencies among characters. Their work interestingly shows that variation in the choice of distance formulae produces limited perturbations of the robustness of the signal when applied to syntactic data.

[31] Cf. Swofford (2001); Schmidt et al. (2002); Ronquist and Huelsenbeck (2003); Yang (2007); Drummond and Rambaut (2007); Tamura et al. (2011); Rambaut et al. (2018), a.o.

[32] E.g., Gray and Atkinson (2003); Bouckaert et al. (2012); Chang et al. (2015).

[33] Gray et al. (2010); Greenhill et al. (2017).

[34] Bryant and Moulton (2004).

[35] So coding "0" would force the method to postulate multiple changes when in fact a single one occurs, and in many cases this would lead the algorithm to reconstruct the wrong node for a certain group, and then spreading the error through the tree.

implied states ("0") as missing characters, to allow the algorithm to ignore redundant characters as a source of information.[36]

The second problem is that character-based algorithms are not *a priori* informed about asymmetries in the likelihood of state transitions. Historical phonology clearly shows several cases of this kind: for instance, Honeybone (2016) shows that a change from the voiceless interdental fricative [θ] to the labial fricative [f] is common, but the reverse is virtually unattested outside of contact areas. Other classic examples are [p] > [f], [p] > [h] or [p] > Ø, all recurrent changes in Indo-European and beyond, and [f] > [p], [h] > [p] or Ø > [p], all extremely rare. With respect to our parameters, we know that there are, for example, several cases of languages acquiring grammaticalized definiteness and no cases of languages dropping this feature,[37] something likely to be reduced to principled explanation, based on the combination of general conditions on change like *Inertia* (Keenan, 2002, 2009) and *Resistance* (Guardiano et al., 2016). An efficient character-reconstructing algorithm will have to be eventually endowed with most such information, but this is not yet the case.

We may expect these problems to affect the topology retrieved by such algorithms. As a consequence, on the other side, any positive taxonomic results retrieved by these methods will attest to the robustness of the signal even *in spite of* the present limitations.

RESULTS

Distance-Based Experiments

Heatmap

The information contained in the syntactic distances was first examined by means of the Heatmap in **Figure 1**. Colors from white to dark blue signal distances lower than the median (spanning from 0 to 0.429), those from yellow to dark red signal distances higher than the median (spanning from 0.430 to 0.857). The overall distribution of colors in **Figure 1** shows that the distances are scattered enough from dark red to dark blue to be potentially informative.

To assess if their distribution has any empirical significance, we considered the maximal aggregations of (white and blue-shaded) cells containing no yellow/red ones which are identified through the clustering option of the program (cf. Section "Distance-based methods"); we compared them to the established genealogical clusters in the sample. In the figure, there are 6 such aggregations which are unambiguous. They correspond to:

(5) a. The Indo-European (henceforth IE) languages.
b. The two Dravidian languages and the two NE-Caucasian ones.
c. Malagasy.
d. The two Basque varieties.
e. The two Sinitic languages.
f. Korean and Japanese.

Two further groups of clusters are also identified along the diagonal. They are more ambiguously interpretable, owing to the fact that they display a partial overlap; in principle, they could single out either the groups in (6) or in (7):

(6) a. Uralic.[38]
b. Turkic,[39] Tungusic,[40] Buryat (i.e., the languages traditionally attributed to the controversial[41] Altaic group) and Yukaghir.

(7) a. Balto-Finnic.
b. The rest of Uralic, Tungusic, Buryat, and Yukaghir.

The clustering algorithm suggests that (6) is the more plausible hypothesis, as highlighted in the tree-like structure on the left and top borders. Hence, the distance distribution in the Heatmap only identifies established taxa (families or isolates: (5)a, c,[42] d, e, (6)a) or supersets of them ((5)b and f; (6)b): thus, no cluster challenges any known historical information, and three of them suggest possible though not yet established supertaxa.

There is also a weaker aggregation of white/pale blue cells next to the sides of the clusters identified along the diagonal. It corresponds to pairs of languages from different families dwelling in the central part of Eurasia (Indo-Iranian, Dravidian, and NE-Caucasian, Altaic, Yukaghir, Uralic except for the three languages now spoken in central and Northern Europe). However, no possible aggregation of white/blue cells displays an average internal distance lower than those of the aggregations identified in (5) and (6) (cf. **Supplementary Material**).

PCoA

The PCoA obtained from the syntactic distances between all the language pairs of the dataset is in **Figure 2**. The first coordinate, which accounts for 59% of the variance, highlights the split between:

(8) a. Non-IE languages (left area).
b. IE languages (right area).

In the left half, the further split corresponding to the second coordinate (accounting for 18% of the remaining variance) separates:

(9) a. Upper-left quadrant: the four languages of the Far East, Malagasy (which has known roots in the same area), and the two Basque varieties, in a rather scattered shape.
b. Bottom-left quadrant: all the other languages of the dataset, i.e., a cloud containing Uralic, Altaic, and Yukaghir and another one with Dravidian and NE-Caucasian.

[36] Note that this does not prevent the algorithm from considering and sometimes selecting reconstructions of ancestral states incompatible with the implicational structure of the dataset.

[37] Roberts and Roussou (2003); Heine and Kuteva (2005).

[38] More specifically Finno-Ugric, represented by two Balto-Finnic languages, three Ugric varieties, two Udmurt (Permic) and two Mari (Volgaic) ones.

[39] Kazakh and Kirghiz (Kipchak sub-branch, Northwestern Turkic, Johanson and Csató, 1998); Turkish (Oghuz sub-branch, Southwestern Turkic: Menges, 1968; Schönig, 1997–1998, a.o); Uzbek (Karluk sub-branch, Southeastern Turkic, Schönig, 1997–1998, a.o.); Yakut (Northeastern Turkic).

[40] Ewenic: Evenki, Even1, and Even2, Khabtagaeva (2018), a.o.

[41] Vovin (2005); Robbeets (2005); but also see Doerfer (1985); Tekin (1994); Soucek (2000); Shimunek (2017), a.o.

[42] There are no other Austronesian languages in our sample.

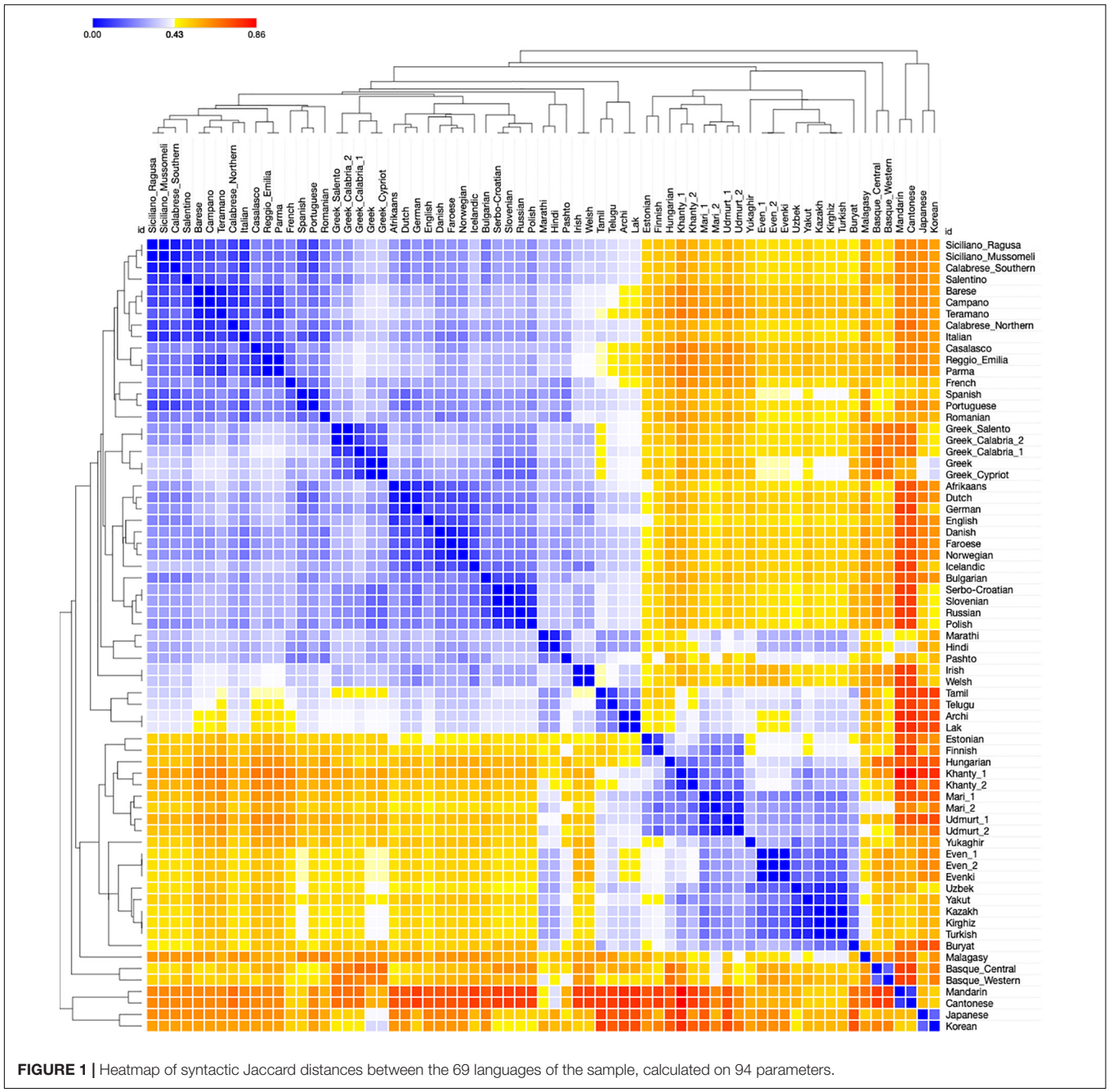

FIGURE 1 | Heatmap of syntactic Jaccard distances between the 69 languages of the sample, calculated on 94 parameters.

In order to obtain a higher resolution, we generated a sequence of further PCoAs from the various subsets of languages progressively identified by the previous ones (cf. Section "PCoAs" in **Supplementary Material**), and they continue to distinguish sets and supersets of independently acknowledged taxa.

Distance-Based Phylogeny

The tests above have preliminarily suggested that a good deal of syntactic diversity is roughly distributed in agreement with genealogical affiliation. Next, we applied phylogenetic algorithms to our data. **Figure 3** displays a (bootstrapped) UPGMA tree. Every cluster identified in the Heatmap also appears in the UPGMA tree.

Character-Based Experiments

Character-Based Phylogeny

The taxonomic results obtained from syntactic distances were finally confirmed by a character-based phylogeny even in spite of the limitations pointed out in Section "Some problems with current methods". The phylogenetic tree calculated with BEAST is in **Figure 4**. The best model was determined by comparing different models using the software Tracer (cf. Section

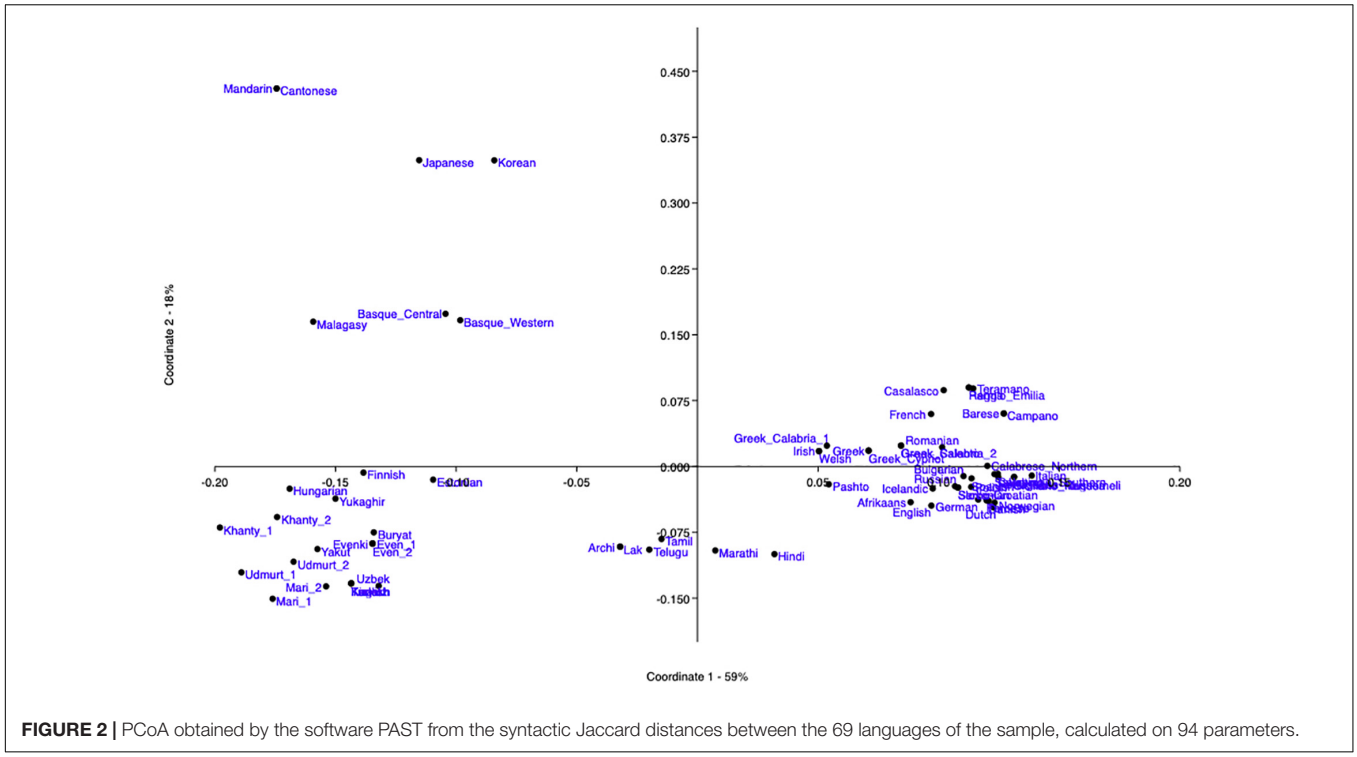

FIGURE 2 | PCoA obtained by the software PAST from the syntactic Jaccard distances between the 69 languages of the sample, calculated on 94 parameters.

"Phylogenetic Analysis – BEAST 2" in **Supplementary Material**). We noticed that most of the nodes were robust across different replications, and the variation was limited to the lower nodes, but a salient exception was the node grouping together Finnish and Estonian, which appeared in different positions of the tree in different replications, and almost always outside of the Uralic node. For this reason, in the tree presented here, we placed a monophyletic constraint on the Uralic languages. An unconstrained tree is available in **Supplementary Figure 8**.

Apart from the Uralic issue, the main differences with UPGMA are:

(10) a. The first two splits, singling out Malagasy along with Sinitic, Japanese, Korean, and Basque[43] from all the rest, recalling the other distance-based visualizations (**Figures 1**, **2**).
b. The clustering of the Archi, Lak, Tamil, and Telugu node with that grouping the so-called Altaic languages and Yukaghir.
c. The reversed position of Buryat and Yukaghir.
d. The intermediate node which combines Celtic with Greek.

Differences in the sub-articulation of Germanic and Romance are discussed below (cf. Section "On the genealogical information in the syntactic trees").

Like in the UPGMA tree, Japanese and Korean fall together, with a posterior probability of 1. Interestingly, both trees are able to assign the languages sharing some similarity in Central Eurasia (cf. **Figure 1**) into their different families (e.g., Indo-Iranian, Dravidian, NE-Caucasian, Uralic, Turkic).

Δ-Scores and Q-Residuals

A graph displaying Δ-scores and Q-residuals (Holland et al., 2002; Gray et al., 2010; Wichmann et al., 2011; Greenhill et al., 2017), along with a SplitsTree network from which they were calculated, can be found in **Supplementary Material**. The median of the Δ-scores is 0.302, and the variance is particularly low (standard deviation: 0.037). The 10 languages associated with the highest values (cf. Section "Network Analysis – NeighborNet" in **Supplementary Material**), i.e., those for which the signal is the least treelike, properly include the languages listed in (9)a, which correspond to the first two outlying branches of the BEAST tree (Mandarin, Cantonese, Korean, Japanese, the two Basque varieties, and Malagasy).

The median of Q-residuals is 0.054, but in this case the variance is quite high, in proportion (standard deviation: 0.021). Again, among the languages with the 10 highest scores, six correspond to the outliers of the BEAST tree (Malagasy has the 11th Q-residual: 0.0805). In particular, while the mean for the Δ-scores is the same as the median, the mean for the Q-residuals is higher (0.058), signaling that the distribution is skewed toward the higher values. In fact, 46 of the 69 languages show a Q-residual lower than the mean, and crucially this subset contains all the 39 Indo-European languages of the sample.

On the Genealogical Information in the Syntactic Trees

With few exceptions, discussed in Section "Sources of deviation", both the UPGMA and BEAST trees capture all the taxa of our

[43] I.e., the languages of the upper left quadrant of **Figure 2** above.

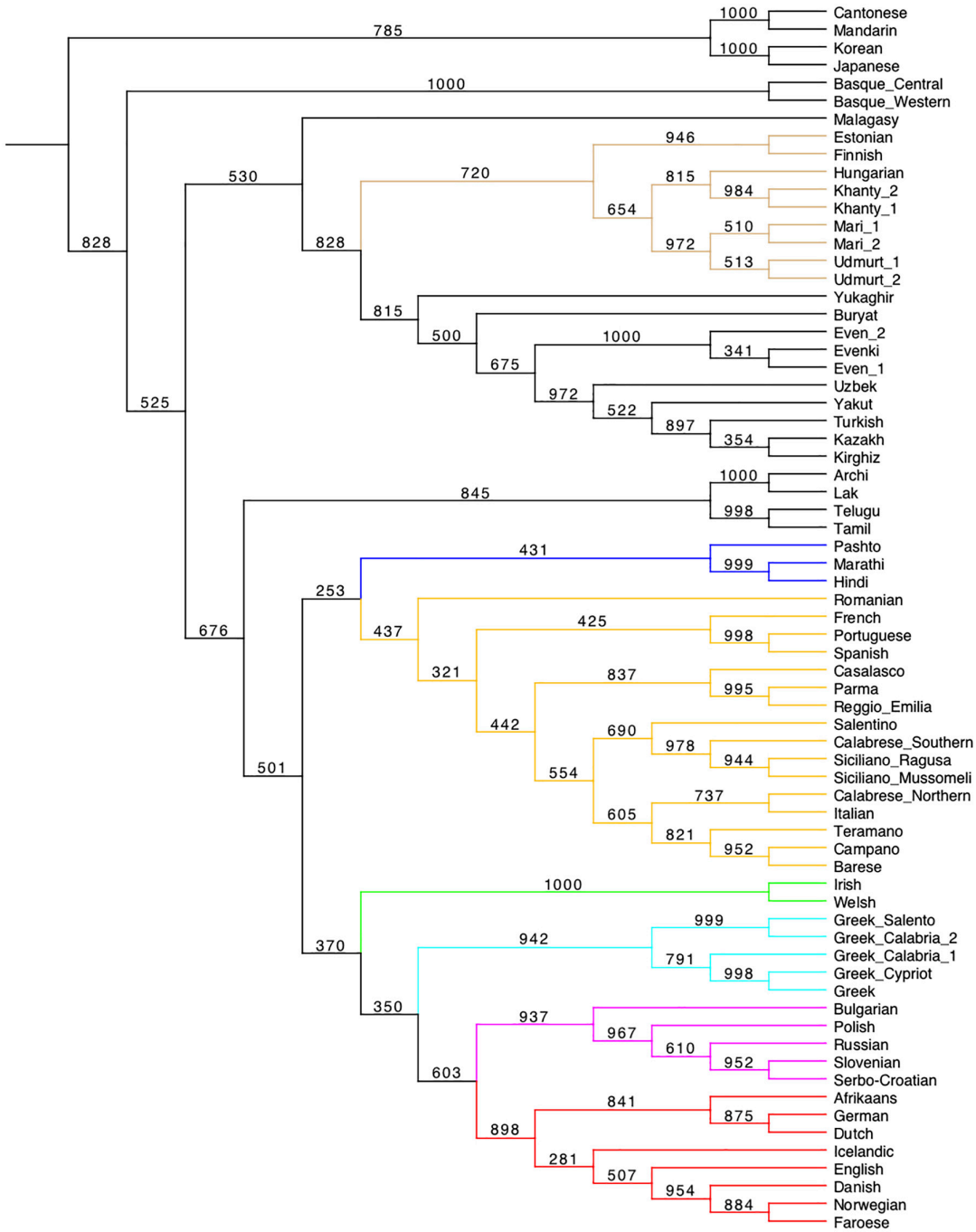

FIGURE 3 | UPGMA tree from syntactic Jaccard distances between the 69 languages of the sample, calculated on 94 parameters. The tree has been produced using Mesquite (Maddison and Maddison, 2007). For information on the bootstrapping procedure adopted, cf. Section "Phylogenetic Analysis – UPGMA" in **Supplementary Material**.

sample that are safely acknowledged by the near-unanimous judgment of historical linguists, based on lexical etymological comparison: this set will be referred to as the "Gold Standard".[44]

Table 1 summarizes the Gold Standard nodes (second column from left), and, in the two last columns, specifies if they are captured by our UPGMA or BEAST trees. UPGMA retrieves

[44]This is the most reliable procedure to evaluate the results of a phylogenetic analysis (cf. Greenhill et al., 2020). From the Gold Standard set we excluded the possible clusters of the micro-variation level, throughout all the families, since their identification in traditional literature is often based on non-vertical evidence and involves geographical and sociolinguistic considerations.

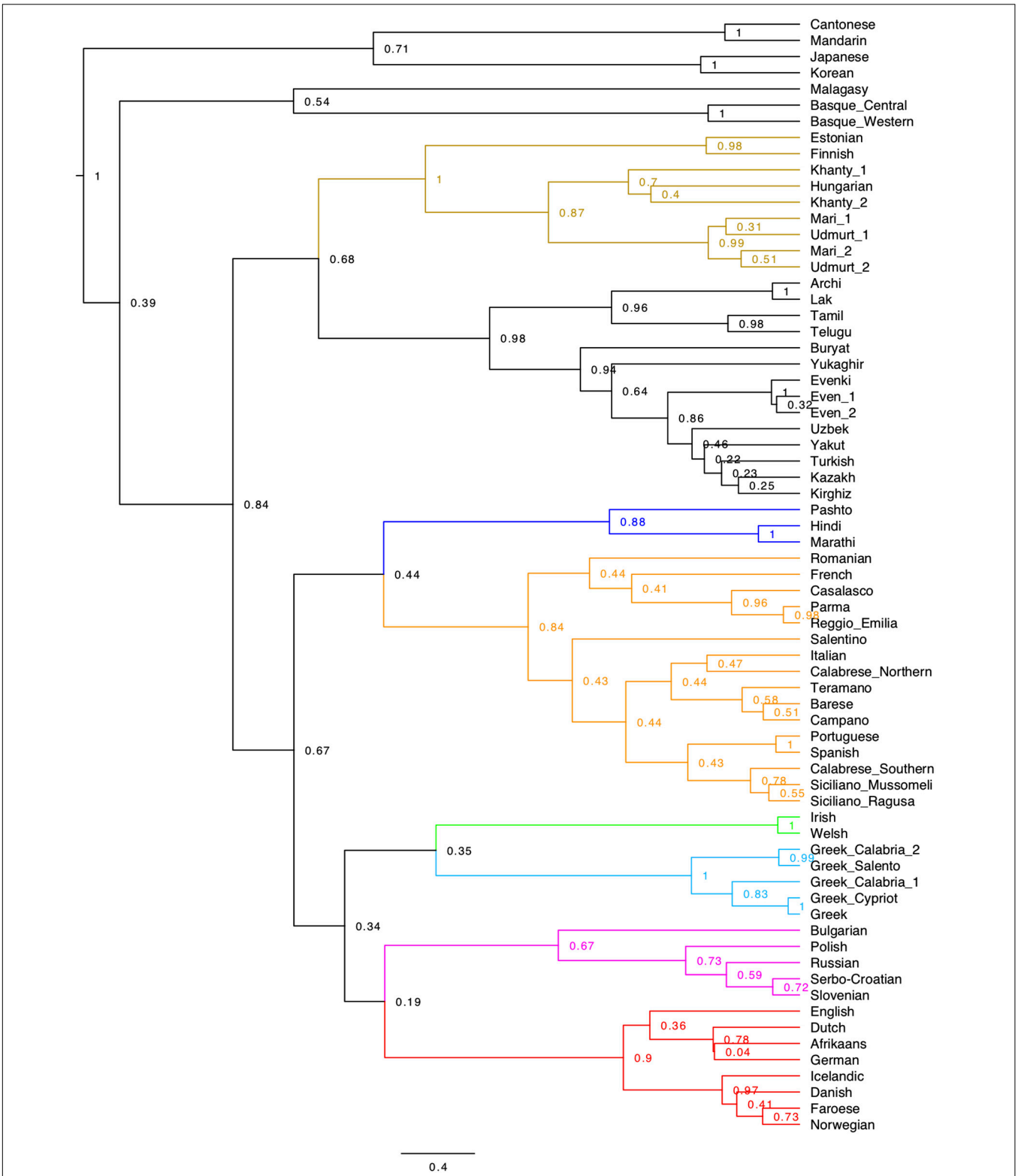

FIGURE 4 | BEAST tree from the 94 syntactic parameters set in the 69 languages of our sample. The best model that we determined is a Gamma Site Model with Substitution Rate = 1, a Mutation Death Model with death $p = 0.1$, a Relaxed Clock (Logarithmic) with clock rate = 1, and a uniform Yule model for the birth rate. The Monte Carlo Markov Chain produced 10,000,000 trees, 25% of which were used for the burn-in and discarded for the purpose of the calculation of the consensus tree. The tree is a consensus tree of 7500 different trees sampled through the 7,500,000 trees (with a sample stored every 1000 generated trees) produced by the Monte Carlo procedure.

TABLE 1 | Our results against the Gold Standard.

	Group	Languages	UPGMA	BEAST
1	Sinitic	Mandarin, Cantonese	YES	YES
2	Dravidian	Tamil, Telugu	YES	YES
3	Basque	Basque_Central, Basque_Western	YES	YES
4	Uralic	Mari_1, Mari_2, Udmurt_1, Udmurt_2, Hungarian, Khanty_1, Khanty_2, Estonian, Finnish	YES	NO[a]
5	Altaic	Kazakh, Kirghiz, Turkish, Yakut, Uzbek, Evenki, Even_1, Even_2, Buryat	YES	NO
6	IE	Irish, Welsh, Marathi, Hindi, Pashto, Greek, Greek_Cypriot, Greek_Calabria_1, Greek_Calabria_2, Greek_Salento, Bulgarian, Serbo-Croatian, Slovenian, Polish, Russian, Faroese, Norwegian, Danish, Icelandic, German, Dutch, English, Afrikaans, French, Casalasco, Reggio_Emilia, Parma, Spanish, Portuguese, Romanian, Siciliano_Ragusa, Siciliano_Mussomeli, Salentino, Calabrese_Southern, Italian, Barese, Campano, Teramano, Calabrese_Northern	YES	YES
7	NE-Caucasian	Archi, Lak	YES	YES
8	Balto-Finnic	Estonian, Finnish	YES	YES
9	Ugric	Hungarian, Khanty_1, Khanty_2	YES	YES
10	Turkic	Kazakh, Kirghiz, Turkish, Yakut, Uzbek	YES	YES
11	Tungusic	Evenki, Even_1, Even_2	YES	YES
12	Kipchak[b]	Kazakh, Kirghiz	YES	YES
13	Celtic	Irish, Welsh	YES	YES
14	Indo-Iranian	Hindi, Marathi, Pashto	YES	YES
15	Greek	Greek, Greek_Cypriot, Greek_Calabria_1, Greek_Calabria_2, Greek_Salento	YES	YES
16	Slavic	Bulgarian, Serbo-Croatian, Slovenian, Polish, Russian	YES	YES
17	Germanic	Faroese, Norwegian, Danish, Icelandic, German, Dutch, English, Afrikaans	YES	YES
18	Romance	French, Spanish, Portuguese, Romanian, Italian, Casalasco, Parma, Reggio_Emilia, Siciliano_Ragusa, Siciliano_Mussomeli, Salentino, Calabrese_Southern, Barese, Campano, Teramano, Calabrese_Northern	YES	YES
19	Indo-Aryan	Hindi, Marathi	YES	YES
20	South-Slavic	Bulgarian, Serbo-Croatian, Slovenian	NO	NO
21	North Germanic	Faroese, Norwegian, Danish, Icelandic	NO	YES
22	West Germanic	German, Dutch, Afrikaans, English	NO	YES
23	Continental West-Germanic	German, Dutch, Afrikaans[c]	YES	YES
24	Ibero-Romance	Spanish, Portuguese	YES	YES

[a]*Recall that the Uralic node in the BEAST tree presented in the text is the product of an explicit constraint placed on this set of languages.* [b]*Northwestern Turkic, Johanson and Csató (1998).* [c]*We included the latter subfamily following Hutterer (1975, p. 195).*

20/23 (87%) major families and subfamilies (21/24: 87.5%, if we include Altaic). BEAST retrieves 21/23 (91.3%) of them (or 21/24: 87.5%). Summing up, the two syntactic trees capture ∼90% of the Gold Standard.

DISCUSSION

The Historical Signal

The results, which are consistent across all the tests performed (Heatmap, PCoA, trees), are largely at odds with statements such as Anderson and Lightfoot's italicized quote in (1), and with the century-long assumptions behind them: syntax has provided, as a whole, a historical signal very close to that of etymological methods. We will now examine the possible roots of the deviations exhibited by syntactic parametric comparison from the expected genealogy.

Sources of Deviation

Deviations from the vertical historical signal can in principle be regarded as due to two factors: secondary convergence (language interference) or homoplasy (parallel independent developments produced by chance). Both are normally *a priori* removed from the input data of automatic lexical phylogenies: one wonders, then, which of these factors is really relevant to produce the deviations above. Let us focus then on the few sources of exceptions to the Gold Standard expectations as they emerge from **Table 1**.

The BEAST tree's failure to capture the Uralic unity (taxon 4) is influenced by few characters in Estonian and Finnish (and their implicational consequences on some other parameters), in which these languages have a value opposite to that of the other Uralic languages and coinciding with that of all IE languages of Europe. For Estonian they are three: p15, CGB, p31, GFP, and p58, NRC, of **Supplementary Figure 1**. For Finnish the relevant ones are p15, CGB, again, and p32, GFN. Parameter CGB defines a macro-areal feature whose value in Balto-Finnic is shared with all IE languages of Europe, while the opposite one is shared by the rest of Uralic, the IE languages of Asia, Altaic, Caucasian, and other Asian languages. Parameter GFP has major implicational consequences on the whole Genitive system, including parameter GFN. Finally, the Estonian value of parameter NRC is the same as in all IE languages, except for some Indo-Iranian ones. These changes have assimilated Finnish and Estonian precisely to their

IE neighbors, with whom very ancient loanwords have also been exchanged.[45]

Also, if an Altaic unit (taxon 5) has ever existed, a part of our experiments (cf. **Figures 1**, **4**) expands it, by placing Yukaghir inside the supposed Altaic family. In fact, the differences of Yukaghir from Eastern Uralic are minimally more numerous than those from the Altaic languages, with which a century-long situation of bilingualism/diglossia as a lingua franca in NE Siberia is well documented.[46]

The outlying position of Bulgarian in both trees (which fail to capture the South Slavic unity, taxon 20) can be traced to relatively recent horizontal parametric convergence; in particular, there are two relevant parametric differences making Bulgarian slightly eccentric with respect to the rest of Slavic:[47] Bulgarian is the only Slavic language (with Macedonian) which selects the value "+" for p17, DGR, like its neighbors Romanian and Greek (it has developed a definite article, and indeed an enclitic one, like Romanian: p24, DCN[48]), and has developed a prepositional Genitive/Dative, like Romanian (cf. p41, GAD).[49] These have long been considered among the areal features of the Balkans.[50] So-called Old Bulgarian (Old Church Slavonic) had the value "−" for DGR. Notice also that DGR starts a long sequence of implications, so that its "−" setting in other Slavic languages *a priori* neutralizes a large number of potential similarities with Bulgarian.

Finally, the UPGMA tree fails to identify West Germanic (taxon 22). As a matter of fact, issues concerning the internal classification of Germanic have been acknowledged in all the quantitative literature.[51] In particular English (along with Afrikaans) has historically experienced most contacts with other Germanic and non-Germanic languages. Furthermore, English has also been recently the focus of a debate between Emonds and Faarlund (2014) and their reviewers and critics[52] about whether, from the Middle English period on, it must be considered a prevailingly Scandinavian rather than West-Germanic offspring (if not the continuation of a creolized version of the two). The unstable position in our experiments confirms that the question is at least a meaningful one. Anyway, it is a fact that English was in close contact with Nordic tribes in both its prehistoric[53] and historic dwelling areas.

In all the cases above, two properties hold: (i) the syntactic detachment of a language from a traditionally expected position in the tree correlates with exhibiting similarity with some neighboring languages; (ii) these deviations from the Gold Standard appear to always be tied to situations of horizontal transmission independently witnessed by other linguistic levels.[54] This confirms Thomason and Kaufman's (1988) conclusion that syntactic borrowing takes place in conditions of "intense" contact, quantitatively measurable by other linguistic variables.

Given the binary nature of our syntactic characters, as opposed to the virtually infinite possibilities provided by lexical arbitrariness, one might think that homoplasy (hence accidental failure of the signal) plays the main role in the deviations from the Gold Standard. On the contrary, the picture suggests that the differences between the syntax trees and the accepted lexical wisdom are always imputable to interference (itself a historical factor), and do not necessarily call for the intervention of homoplasy.

Vertical and Horizontal Transmission

Even horizontal effects have relatively little impact on the general topology of the tree. For instance, under all our experiments, the Italiot Greek varieties cluster with Standard and Cypriot Greek: the protracted contact and documented syntactic interference between Romance and Greek in Southern Italy[55] have not disrupted the overall vertical signal of either family. To measure the conflict between vertical and horizontal information in the signal, we used Δ-scores and Q-residuals. Recall that a lower value of these indices speaks for a sharper vertical signal.

Δ-scores in our experiment, with a median as low as 0.302, yield better results than those obtained in both datasets used in Greenhill et al. (2017), where lexical characters displayed a median of 0.38 and structural characters displayed one of 0.44.

The Q-residuals perform less well: Greenhill et al. (2017) had a median of 0.0062 for lexical characters and 0.0354 for structural characters, against our median of 0.054.[56] Notice, however, that Wichmann et al. (2011) tested the two measures on a group of languages of the Automatic Similarity Judgment Program database,[57] and noticed that Δ-scores distributed uniformly with respect to age and size of the language family; Q-residuals instead correlated with such factors, becoming higher and less informative for chronologically deep and numerous and internally diverse families. Based on these results, they argued precisely in favor of Δ-scores as more accurate measures of non-tree-likeness. This seems to be true in our experiment as well: the highest Q-residuals are associated with languages occurring on the higher branches, whose genetic affiliation is still unclear; but all Indo-European languages display Q-residuals lower than the mean, suggesting that the measure is indeed sensitive to the age and size or diversity of the family (cf. Section "Δ-Scores and Q-Residuals"). This is not true for Δ-scores: while the outliers equally display high Δ-scores, IE languages are more

[45] Kylstra et al. (1991) suggest that the first contacts between Germanic and Balto-Finnic date from around 1000 BC.

[46] Wurm et al. (2011, pp. 970, 978).

[47] Cf. Longobardi et al. (2013).

[48] For this circum-Pontic isogloss see Guardiano et al. (2016).

[49] This, in turn, may have enabled the resetting of p43, GFO, as well, i.e., the disappearing of an inflected Genitive.

[50] See Sandfeld (1930) and now, specifically for syntactic borrowing, Tomić (2006).

[51] Dyen et al. (1992); Ringe et al. (2002); Jäger (2015).

[52] Barnes (2016); Bech and Walkden (2016); Stenbrenden (2016); Crisma and Pintzuk (2019), and the contributions to the 2016 issue, 6.1, of *Language Dynamics and Change*.

[53] Hutterer (1975), a.o.

[54] Even the internal comparison between the UPGMA and the BEAST trees turns out to be informative to confirm cases where the signal is conflicting, i.e., one or more languages can be associated with different phylogenetic histories.

[55] Guardiano and Stavrou (2014, 2019, 2020); Guardiano et al. (2016); Ledgeway (2006); Ledgeway (2013); Ledgeway et al. (2018), a.o.

[56] Greenhill and his collaborators (p.c.) suggest that this difference can be explained as a result of the fact that while Δ-scores might be more sensitive to conflicting signal (i.e., the presence of two alternative histories for a taxon), Q-residuals might be more sensitive to noise in the data.

[57] ASJP, Wichmann et al. (2020).

evenly distributed above and below the mean (23 vs. 16). If Wichmann et al. (2011) are right, then, our result is expected: it is likely that Q-residuals cannot meaningfully apply to long-range classifications across many different families.

Ultralocality: Hints About Microvariation

The internal articulation of the Romance dialects of Italy retrieved by the UPGMA tree is consistent with their traditional classification.[58] The tree clusters them together, then identifies the Gallo-Italic group (Reggio Emilia, Parma, and Casalasco), the Extreme southern group (Siciliano, Southern Calabrese, and Salentino), and one that clusters three Upper southern dialects (Campano, Teramano, and Barese) but not Northern Calabrese: this may reflect the isolation of this dialect as representative of an area known to exhibit several peculiarities with respect to the whole Italian group.[59]

At this level of microvariation, no taxonomy can be really projected onto a genuine phylogeny, because of the uninterrupted contact and diffusion of isoglosses among contiguous dialects (cf. the network and the PCoA in **Supplementary Figures 14, 16**; also cf. Sarno et al., 2014 on strong genetic admixture in Southern Italy). This may have produced the differences between the UPGMA and BEAST trees: the BEAST tree may rather highlight the actual secondary relations which have occurred between Sicilian and Ibero-Romance, some closeness between Gallo-Italic and French, and also plausible interference of Balkan languages with Salentino, which appears as the outlier of all of Romance.

Thus, even minimally different character strings and very short parametric distances have good resolution power. Moreover, the fact that parametric distances become very low at this level of comparison is exactly what we expect if syntax evolves proportionally to other historical variables.

The resolution we obtain in micro-variation is inevitably based on parameters which must have undergone recent changes, i.e., which, virtually by definition, are not as stable as others. Yet, their instability has not produced any conceivable disruption of the correct topology in other areas of the phylogenies. This very consequential observation is discussed in Section "Input data and phylogenetic results".

Globality: Hints About Long-Range Relations

The most salient feature of parametric systems is their potential universality. Accordingly, our phylogenetic analyses provide some preliminary insights about possible or proposed long-range groupings. They will eventually have to be evaluated through more elaborate statistical analyses, but provide a list of heuristic suggestions for further testing.

First, nearly all the experiments single out a set of languages as outlying the rest of the sample: Japanese, Korean, the two Sinitic and two Basque varieties, and, except for the UPGMA tree, Malagasy. The other languages are always identified as a monophyletic structure and Δ-scores and Q-residuals suggest that they have a more reliable vertical articulation.

In addition to recognizing all classical families, our data suggest that Indo-Iranian, Dravidian, NE-Caucasian, Turkic, Tungusic, Buryat, Yukaghir, and part of Uralic partake of some similarity, which is especially highlighted in **Figure 1**; however, such similarity turns out to be weaker than the respective family affiliations (cf. the trees in **Figures 3**, **4**). The methods used cannot decide how much of this similarity is secondary and areal, though the fact that (only) the IE languages of Asia share it, and (only) the Uralic languages that dwell in Central-Western Europe (Hungarian, Finnish, Estonian) do not, suggests that part of it must be.

Next, all experiments point to the unity of part of the controversial Altaic family (Turkic and Tungusic), and a weaker connection of this cluster to Buryat (Mongol), but also to Yukaghir.

Even more robustly, the syntactic analysis argues for a Korean-Japanese relation, although sustained by a relatively low number of non-null comparisons (30 pairs; only 12, according to a Jaccard measure). Statistical support is very high, as is only the case, in our sample, for a few safely established pairs/groups. Notice that some studies have proposed that even sound correspondences support the relatedness of Japanese and Korean.[60]

Notice, instead, that the clustering of Korean and Japanese with Mandarin and Cantonese in both trees should not deceive us, because it is likely to be a bias of the tree algorithms (clustering together data points which are both outliers with respect to the main group of taxa is a common error, usually described as Long-Branch Attraction: Bergsten, 2005). This becomes clear from the distance distribution: in **Figure 1**, the two groups are clearly set apart; moreover, if we draw a PCoA specifically focused on the languages of the upper left quadrant of **Figure 2**, Japanese-Korean and the two Chinese varieties clearly fall into distinct quadrants (cf. **Supplementary Figure 3**).

Finally, none of our experiments hints at a Macro-Altaic grouping.[61] However, the syntactic data cannot exclude some genealogical relation between Korean-Japanese and central Asian languages, with secondary influences from the East Asian area.[62]

A worth exploring relation is that between Uralic and Altaic. Uralic languages are scattered in terms of distance but, with the exception of Balto-Finnic in the BEAST tree, they are recognized as a unit. In spite of the noted similarities with IE languages, the syntactic data provide sufficient evidence that Balto-Finnic is indeed a Finno-Ugric family influenced by IE rather than the opposite, and that, if anything, the whole Uralic is closer to Altaic than to Indo-European. First, when we

[58] Pellegrini (1977).

[59] Lausberg (1939).

[60] For instance, Whitman (2012); also see the discussion in Robbeets (2008a), a.o.

[61] Altaic-Korean-Japanese: see the discussion in Port et al. (2019) and the Trans-Eurasian hypothesis of Robbeets (2008b).

[62] The consequence of such influences is reasonably the degrammaticalization of Person and Number features (p5 FGP and p7 FGN), which are rich in neutralizing implicational effects on further parameters. Indeed, after close consideration of the parameter values, the 0s induced by the lack of value "+" for FGP is the main source of peculiar similarity between Mandarin-Cantonese and Korean-Japanese. Beyond this, the parameters in which the four languages share a value in contrast to all the other languages are only two: p27, FGE, about the necessity of a classifier between a numeral and a head noun (itself a property very frequent in languages without a positive value at FGN: see Cathcart et al., 2020), and p61, LKP, about the presence of a special morpheme linking the noun with essentially any of its arguments.

place a monophyletic constraint on the set of Uralic languages in the BEAST phylogeny, the stable result is that Uralic is clustered with the Altaic-Yukaghir node. Second, the other Uralic languages are never separated from the Altaic group in any experiments (cf. Section "Phylogenetic Analysis – BEAST 2" in **Supplementary Material**). Third, the Genitive systems of Estonian and Finnish (and the pronominal possessive system of Estonian), which oppose them to all the other Uralic (but also Turkic and Tungusic) languages (cf. Section "Sources of deviation"), must be regarded as an innovation with respect to the others: it has involved the loss of agreement between the features of a Genitive and those expressed through a dedicated morpheme on the head noun, a common Uralic feature.[63] The weakening or loss of such morphemes is a well-known diachronic phenomenon, attested, e.g., for verbs and adjectives in the history of Romance and Germanic (possibly an instance of what Keenan, 2009 considers phonological "DECAY"); its creation anew is not easily observed. All this is consistent with the possibility of some Uralo-Altaic unity, blurred by the Indo-Europeanization of the Balto-Finnic languages, while it makes any original Indo-Uralic unity excluding Altaic and Yukaghir highly unlikely.[64]

All experiments also point to significant closeness of NE-Caucasian and Dravidian (average distance 0.23). This similarity, which needs to be investigated, connects to another stable outcome of our experiments: the fact that Basque lies outside the group constituted by the other Eurasian languages except for those of the Far East, and, in particular, does not show any trace of the sometimes proposed relation to the NE-Caucasian languages (average distance 0.51).[65]

The Homology Conjecture

We conclude that (A) syntactic phylogenies are very similar to the lexical-etymological ones, and (B) the small proportion of deviation can be imputed to secondary convergence only (which so far has been *a priori* removed from lexical, though not syntactic, data). These two claims are merged into:

(11) The **Homology Conjecture:** Syntactic and lexical histories provide the same evolutionary topologies once interference is equally taken into account

This hypothesis is in agreement with the expectations of syntactic *Inertia* (cf. Section "Syntax, Cognitive Science, and Historical Taxonomy") and is parallel to the Neogrammarian Regularity hypothesis, in attributing any disruption of an ideal diachronic evolution (in that case, regularity of non-analogical sound change) to dialect admixture.

A Comparison With Phonemic Inventories

We checked then what kind of signal can be retrieved from our language sample through non-lexical (and potentially cross-family) traits that are not characterized by the three formal properties we used to select our syntactic characters (cf. (3)), and that are more remote from the core generative mechanisms of grammar.

For instance, inventories of autonomous phonemes have been used for comparison across different families, e.g., in Creanza et al. (2015). This work employs two large phonemic databases, PHOIBLE[66] and Ruhlen,[67] in an attempt to align phonemes into corresponding classes based on phonetic similarity.[68] To check whether phonemic characters generate informative phylogenies at our scale/density of sampling, we generated a BEAST tree (**Figure 5**) from the entries in Ruhlen's data corresponding to the languages of our study. The only taxa of the Gold Standard above identified by this tree are the 5 (21.7%) listed in (12):[69]

(12) a. Dravidian
b. Indo-Aryan
c. Tungusic (Even, Evenki)
d. Balto-Finnic
e. NE-Caucasian

These pairs are also geographically close and might in part reflect reciprocal secondary influence, as the cluster Spanish/Basque apparently does. Most other clusters do not reflect historical information at all (e.g., Sicilian-Faroese, English-Pashto, Irish-Buryat, Mari-Cypriot Greek etc.).

Our experiment supports Creanza et al.'s (2015, p. 1269) claim that "phoneme inventories are affected by recent population processes and thus carry little information about the distant past":[70] phonemic data exhibit a much shallower historical signal than syntactic data, and are actually prone to detect secondary convergence (see also Wichmann and Holman, 2009). This result shows the relevance of comparing different input data and prompts some considerations about their historical signal.

Input Data and Phylogenetic Results

Some previous phylogenetic experiments found less historical signal when looking at structural traits. For instance, Greenhill et al. (2017) compared the evolutionary rate and signal of lexical etymologies with that of some structural properties in 81 Austronesian languages. They found that, on average, structural properties display higher rates of change than lexical

[63] Collinder (1960).

[64] Also see Marcantonio (2002) and the debate ensued.

[65] Starostin (1996) and Bengtson (2017), a.o.

[66] Moran and McCloy (2019).

[67] http://starling.rinet.ru/typology.pdf

[68] Of course, it is plausible that an interesting historical signal can be retrieved from analyses of more abstract phonological processes and constraints rather than just of the physical resemblance of autonomous phonemes. Promising results on this line, which parallel the ones of our approach, are provided in Macklin-Cordes et al. (2020).

[69] Few other clusters with more indirect genealogical content are those formed by two continental West Germanic languages (German and Dutch), two Northern Germanic languages (Danish, Icelandic), four Slavic languages (Bulgarian, Russian, Slovenian, Serbo-Croatian), and two Romance languages (Portuguese and French).

[70] Creanza et al. (2015) complement this claim with pointing out the limited and historically recent correlations found between phonemic distances and genetic distances. Using syntactic parameters, instead, Longobardi et al. (2015) found that genetic differences correlate with linguistic distances more than with geographic distances in Europe.

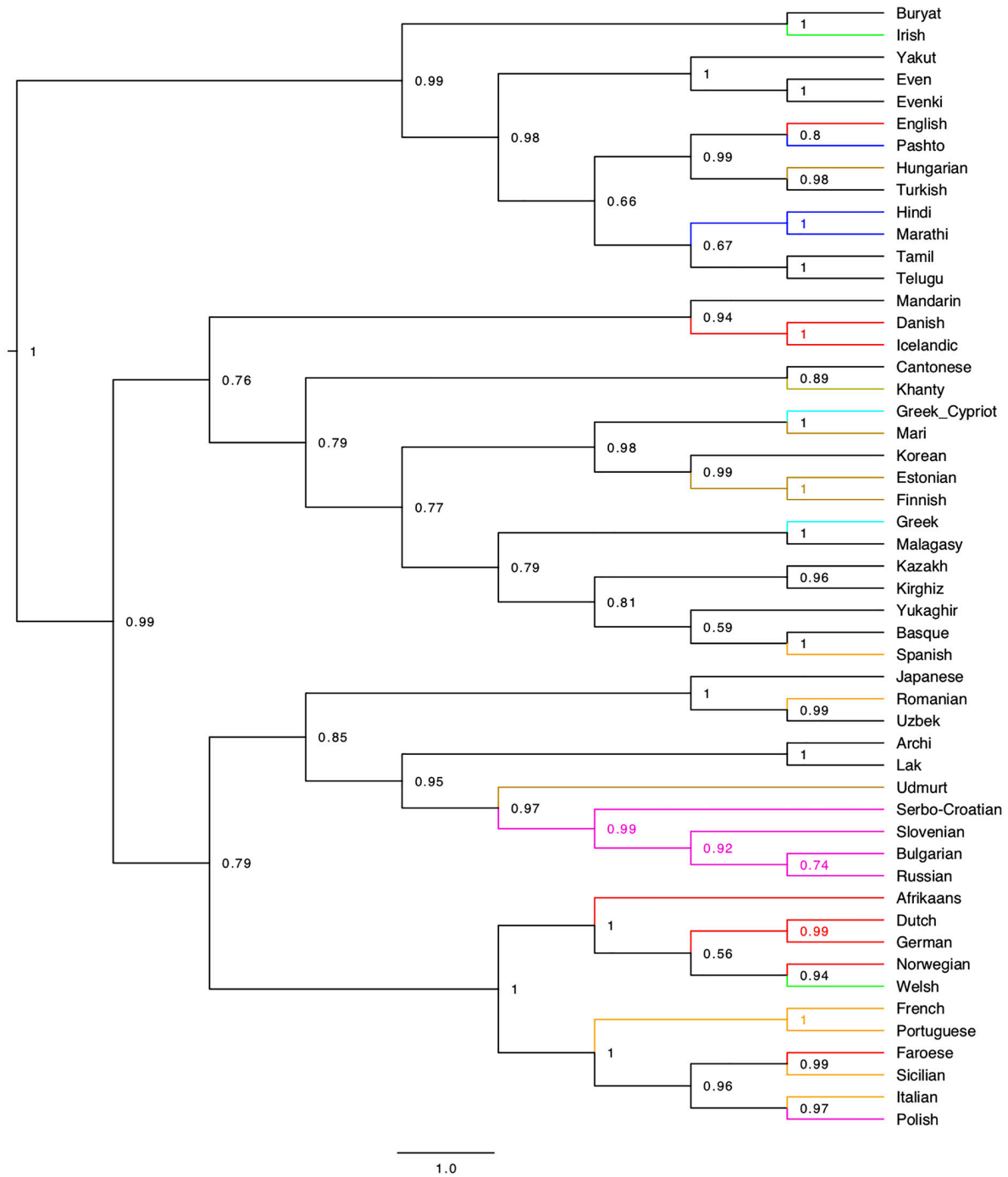

FIGURE 5 | BEAST tree from Ruhlen's phonemic dataset. The tree contains a subset of the languages used in Creanza et al. (2015), consisting of the 52 languages overlapping with those used in this article. The color coding is the same as for the previous phylogenies, visually highlighting the differences in the clustering of the families. The best model that we determined is a Gamma Site Model with Substitution Rate = 1, a Mutation Death Model with death $p = 0.1$, a Relaxed Clock (Logarithmic) with clock rate = 1, and a uniform Yule model for the birth rate. The Monte Carlo Markov Chain produced 10,000,000 trees, 25% of which were used for the burn-in and discarded for the purpose of the calculation of the consensus tree. The tree is a consensus tree of 7500 different trees sampled through the 7,500,000 trees (with a sample stored every 1000 generated trees) produced by the Monte Carlo procedure.

sets, and that there are subsets of properties (both lexical and structural) that change much slower or much faster than the average. For instance, number marking on the noun phrase and the presence of tones showed up as conservative, while article properties and vowel length as features that tend to change over time.

Thus, in certain respects, the historical signal retrieved through the syntactic dataset of the present article is more robust and promising than that obtained with their structural traits: the results are not necessarily in contrast, though, because of the different properties of the input data and of the different idealizations made on them (cf. (3a–c)) in Section "Syntactic data and taxonomic problems".

First, one difference is that the structural traits used in Greenhill et al. (2017), like those employed in a preliminary work by Dunn et al. (2005), include not just syntactic characters but also other non-lexical features, some of which (presence of phonetically defined autonomous phonemes) are shown here to contain a shallow and genealogically very disruptive signal. So, this is a potential cause of the different outcome.

Second, parameters are coded as representations of the generative devices in mental grammars, rather than as generated patterns. It is conceivable that this provides them with a high degree of cognitive realism and deductive information, which in turn provide historical resolution. Recall that only an average of 20 parameters (39 if we consider identities on the "−" values) are fully comparable across the language pairs of our sample, due to the redundancies created by the pervasive implicational structure of parameters (cf. Section "Parameters and schemata"). The correctness of the topologies retrieved by so few characters suggests indeed that parameters do have high-resolution.

Finally, a most interesting property brought to light by our experiments is that all the divergences of syntax from the established or expected topologies can in principle be explained in terms of secondary convergence: neither of the syntactic topologies presents clear cases where an incorrect cluster is exclusively determined by homoplasy. Notice that *a priori* we might expect homoplasy to seriously affect syntactic topologies, given that our characters are binary and that we deal with many independent families. However, this is not the case. This may in part be due to the general robustness of the complementary vertical signal; but a relevant role must be played here by the third property of parametric data, their pervasive interdependence: the redundancy provided by parametric implications neutralizes the effects of the most obvious source of homoplasy. The resolution we obtain in the articulation of families and subfamilies, up to recently and minimally differentiated dialects, comes at the cost of considering at least some traits with a high-rate-of-change, which discriminate between close varieties; thus, by definition, they are less stable than parameters that have remained unchanged for millennia all over large families. In principle, their instability might have produced a great amount of homoplasy elsewhere in the trees, disrupting the correct phylogenies across other families. Yet, this has not happened with our dataset. Many parameters in **Supplementary Figure 1** which make finer distinctions within Romance dialects (and other close varieties) are neutralized in most non-Romance (or non-IE) languages, due to their dependence on hierarchically higher parameters. This has reduced accidental similarities between distant families. It is plausible that any attempt to attain globality with grammatical characters, in order not to crash against homoplastic effects, must indeed take into account the pervasive interdependence of such traits.

CONCLUSION

Five major inferences can be drawn from the results of this article.

The Historical Signal of Syntax

The syntactic structures of I-languages (Chomsky, 1986: the abstract rule systems of computational theories of mind; also see Everaert et al., 2015) are an effective tool of historical knowledge (*pace* contrary positions in comparative philology and in modern formal syntax, as well as some skepticism expressed in quantitative phylogenetics: cf. Dunn et al., 2011): they retrieve most of the phylogenetic information contained in trees produced by lexical etymologies. Strikingly, the trees obtained from syntax are essentially unaffected by the inevitable amount of homoplasy which must be produced by the binary nature of the characters used. Also, the verticality of the syntactic signal and its chronological depth are far stronger than those of more externalized traits, like phonetic similarity in phonemic inventories (in agreement with Creanza et al.'s, 2015 conclusion that such phonemic characters are not informative about deep-time relations). The phylogenies retrieved through syntax must be relatively deep in time, if they are able to sharply separate, e.g., Basque from IE and other Eurasian families: given the limitations of (non-speculative) methods for investigating deeper language evolution, stressed in Hauser et al. (2014), this empirical, bottom-up approach is a promising perspective for studying the past of human syntax.

Historical Support for Generative Grammars

The search for a historical signal represents an unprecedented type of evidence to test the format of representation of mental grammars used in syntactic theories, especially in minimalist approaches to parameters. As in the formal grammatical tradition, we have tried to model the dataset used not simply as a set of experiential facts, but mostly as a deductive structure in which surface data (e.g., E-languages) are largely the product of the combination of simpler and less numerous principles (I-languages). The success in retrieving a historical signal corroborates this general approach on a domain different from the usual ones (synchrony, typology, acquisition) used to support formal linguistic theory.

Generative Grammars and Phylogenetic Evidence

Conversely, this robust historical signal suggests a reconsideration of the practice of formal syntax itself: for example, when a clear deviation of a parameter value occurs in a language from the state of its established family, it will call for an explanation. If the synchronic analysis is correct, then for linguistic theory the question should arise of how, and possibly why, the disconnection from the family pattern has taken place.

Phylogenetics and Language Distances

Beyond some minor complementarity between character- and distance-based models of syntactic history, the topologies

retrieved by the two methods are quite similar. This is in line with Greenberg's (1987) controversial claim that a first approximation to language taxonomy is possible even ahead of step-by-step reconstruction of all ancestral characters.

Tools and Perspectives

We have used a tool for language description (a list of YES/NO existential questions: cf. Crisma et al., 2020) universally applicable and requiring very limited information (in principle no more than one YES answer per parameter set to "+"): this was mainly possible owing to the redundancy and default settings which characterize a minimalist approach to parameters. Beyond phylogenetics, a system with these properties has obvious consequences for the study of grammatical diversity and language learnability (cf. Sakas et al., 2017).

In sum, we regard these results as a breakthrough with respect to a long tradition in linguistics: they indicate that there exists a signal in syntax which might be used for aiming at progressively more comprehensive phylogenies of human languages. We suggest the possibility of adding less visible taxonomic traits, such as syntactic parameters, to the toolkit of phylogenetic linguistics as the basis for a *qualitative* revolution, which may complement the scope and success of the *quantitative* one.

DATA AVAILABILITY STATEMENT

The code used to generate the experiments and the figures can be found at https://github.com/AndreaCeolin/FormalSyntax doi: 10.5281/zenodo.4323165.

AUTHOR CONTRIBUTIONS

GL and CG devised the comparative methodology and the specific parametric structure. GL, CG, MAI, and AC collected the data. AC performed the computational experiments. GL, MAI, and AC wrote the Introduction. GL, CG, and AC wrote the Materials and Methods, the Results, and the Discussion. GL wrote the Conclusion.

FUNDING

This work was funded by: ERC Grant ERC-2011-AdvG_295733 (Langelin, PI: GL); MIUR PRIN 2017K3NHHY "Models of language variation and change: new evidence from language contact" (CG); Fondo di Ateneo per la Ricerca 2014 (Università di Modena e Reggio Emilia), "La microvariazione in Italia Meridionale: convergenze (e divergenze) di geni e lingue nel tempo e nello spazio" (CG); and Fondo di Ateneo per la Ricerca 2020 (Università di Modena e Reggio Emilia) Impulso (MAI).

ACKNOWLEDGMENTS

We thank some of the collaborators on the ERC Adv. Gr. project, LanGeLin, for various contributions to the data collection, and especially G. Cordoni for suggestions on the computational analyses used in this article, as well as T. Biberauer, L. Bortolussi, P. Crisma, F. Fanciullo, L. Franzoi, R. Gray, S. Greenhill, A. Holmberg, H. Koopman, G. Jäger, R. Kayne, D. Kazakov, P. Kiparski, D. Kühnert, the late R. Lazzeroni, M. Mancini, P. Munro, I. Roberts, A. Sgarro, and G. Silvestri for conversations and helpful suggestions on both the data and the argumentation of this article. We are of course indebted to the journal's referees and all the colleagues and language consultants who provided native data about the languages over the last years.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2020.488871/full#supplementary-material

REFERENCES

Adelung, J. C. (1806–1817). *Mithridates oder allgemeine Sprachenkunde mit dem Vater Unser als Sprachprobe in bei nahe fünfhundert Sprachen und Mundarten (fortgesetzt und bearbeitet von Dr Johann Severin Vater)*, 4 vols. Berlin: Vossische Buchhandlung.

Anderson, S. R. (2012). *Languages: a Very Short Introduction*. Oxford: Oxford University Press.

Anderson, S. R., and Lightfoot, D. (2002). *The Language Organ, Linguistics as Cognitive Physiology*. Cambridge: Cambridge University Press.

Arens, H. (1969). *Sprachwissenschaft*. Freiburg/München: Verlag Karl Alber.

Baker, M. (2001). *The Atoms of Language*. Oxford: Oxford University Press.

Balbi, A. (1826a). *Introduction à l'atlas ethnographique du globe, contenant un discours sur l'utilité et l'importance de l'étude des langues appliquées à plusieurs branches des connaissances humaines; un aperçu sur les moyens graphiques employés par les différens peuples de la terre: des observations sur la classification des idiomes décrits dans l' atlas; un coup-d' oeil sur l' histoire de la langue slave, et sur la marche progressive de la civilisation et de la littérature en Russie*. Tome premier, Paris: Rey et Gravier.

Balbi, A. (1826b). *Atlas ethnographique du globe ou classification des peuples anciens et modernes d' après leur langues, précédé d' un discours sur l' utilité... Russie (cf. Balbi 1826a), Et suivi du tableau physique, moral et politique des cinq parties du monde*. Paris: Rey et Gravier.

Barnes, M. (2016). Review of Joseph E. Emonds and Jan Terje Faarlund: *English: The Language of the Vikings. Maal og Minne* 108, 173–179.

Bech, K., and Walkden, G. (2016). English is (still) a West Germanic language. *Nord. J. Linguist.* 39, 65–100. doi: 10.1017/s03325865150 00219

Bengtson, J. D. (2017). *Basque and its Closest Relatives: a New Paradigm*. Cambridge, MA: Mother Tongue Press.

Bergsten, J. (2005). A review of long-branch attraction. *Cladistics* 21, 163–193. doi: 10.1111/j.1096-0031.2005.00059.x

Berwick, R. C., and Chomsky, N. (2015). *Why Only Us? Language and Evolution*. Cambridge, MA: MIT Press.

Biberauer, T. (ed.) (2008). *The Limits of Syntactic Variation*. Amsterdam: John Benjamins.

Biberauer, T., and Roberts, I. (2017). "Parameter setting," in *The Cambridge Handbook of Historical Syntax*, eds A. Ledgeway and I. Roberts (Cambridge: Cambridge University Press), 134–162.

Boeckx, C., and Leivadá, E. (2014). On the particulars of Universal Grammar: implications for acquisition. *Lang. Sci.* 46, 189–198. doi: 10.1016/j.langsci.2014.03.004

Boeckx, C., and Piattelli-Palmarini, M. (2005). Language as a natural object - Linguistics as a natural science. *Linguist. Rev.* 22, 447–466. doi: 10.1515/tlir.2005.22.2-4.447

Bortolussi, L., Longobardi, G., Guardiano, C., and Sgarro, A. (2011). "How many possible languages are there?," in *Biology, Computation and Linguistics*, eds G. Bel-Enguix, V. Dahl, and M. D. Jiménez-López (Amsterdam: IOS Press), 168–179.

Bouckaert, R., Lemey, P., Dunn, M., Greenhill, S. I., Alekseyenko, A. V., Drummond, A. I., et al. (2012). Mapping the origins and expansion of the Indo-European language family. *Science* 337, 957–960. doi: 10.1126/science.1219669

Bouckaert, R., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., et al. (2019). BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* 15:e1006650. doi: 10.1371/journal.pcbi.1006650

Braudel, F. (1958). Histoire et sciences sociales: la longue durée. *Annales* 13-4, 725–753. doi: 10.3406/ahess.1958.2781

Bryant, D., and Moulton, V. (2004). NeighborNet: an agglomerative algorithm for the construction of planar phylogenetic networks. *Mol. Biol. Evol.* 21, 255–265. doi: 10.1093/molbev/msh018

Cathcart, C. A., Hölzl., A., Jäger, G., Widmer, P., and Bickel, B. (2020). Numeral classifers and number marking in Indo-Iranian. A phylogenetic approach. *Language Dynamics and Change* 1–53. doi: 10.1163/22105832-bja10013

Cavalli-Sforza, L. L., Menozzi, P., and Piazza, A. (1994). *The History and Geography of Human Genes*. Princeton, NJ: Princeton University Press.

Ceolin, A. (2019). Significance testing of the Altaic family. *Diachronica* 36, 299–336. doi: 10.1075/dia.17007.ceo

Chang, W., Cathcart, C., Hall, D., and Garrett, A. (2015). Ancestry-constrained phylogenetic analysis supports the Indo-European steppe hypothesis. *Language* 91, 194–244. doi: 10.1353/lan.2015.0005

Chapin, P. G. (1974). Proto-Polynesian *ai. *J. Polyn. Soc.* 83, 259–307. doi: 10.2307/20705006

Chomsky, N. (1964). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. (1986). *Knowledge of Language*. New York, NY: Praeger.

Clark, R., and Roberts, I. (1993). A computational model of language learnability and language change. *Ling. Inq.* 24, 299–345. doi: 10.2307/4178813

Collinder, B. (1960). *Comparative Grammar of the Uralic Languages*. Uppsala: Almvqvist & Wiksell.

Cordoni, G., Woodward, M. J., Wu, H., Alanazi, M., Wallis, T., and La Ragione, R. M. (2016). Comparative genomics of European avian pathogenic *E. coli* (APEC). *BMC Genom.* 17:960. doi: 10.1186/s12864-016-3289-7

Creanza, N., Ruhlen, M., Pemberton, T. J., Rosenberg, N. A., Feldman, M. W., and Ramachandran, S. (2015). A comparison of worldwide phonemic and genetic variation in human populations. *Proc. Natl. Acad. Sci. U.S.A.* 112/5, 1265–1272. doi: 10.1073/pnas.1424033112

Crisma, P., Guardiano, C., and Longobardi, G. (2020). Syntactic parameters and language learnability. *Studi e Saggi Linguistici* 58, 99–130. doi: 10.4454/ssl.v58i2.265

Crisma, P., Guardiano, C., and Longobardi, G. (to appear). "Toward a unified theory of Case form and Case meaning," in *The Place of Case in Grammar*, eds E. Anagnostopoulou, D. Mertyris, and Ch. Sevdali (Oxford: Oxford University Press).

Crisma, P., and Longobardi, G. (in press). "The parametric space associated with D," in *The Oxford Handbook of Determiners*, eds S. Armoskaite and M. Wiltschko (Oxford: Oxford University Press).

Crisma, P., and Pintzuk, S. (2019). The noun phrase and the Viking hypothesis. *Lang. Var. Chang.* 31, 219–246. doi: 10.1017/s0954394519000127

Davis, J. C. (1986). *Statistics and Data Analysis in Geology*. New York, NY: John Wiley and Sons.

Di Sciullo, A. M., and Boeckx, C. (eds) (2011). *The Biolinguistic Enterprise. New Perspectives on the Evolution and Nature of the Human Language Faculty*. Oxford: Oxford University Press.

Diamond, J. M. (1997). *Guns, Germs and Steel. The Fates of Human Societies*. New York, NY: W. W. Norton and Company.

Doerfer, G. (1985). *Mongolica-Tungusica*. Wiesbaden: Otto Harrassowitz.

Drummond, A. J., and Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7:214. doi: 10.1186/s12864-016-3289-214

Dunn, M., Levinson, S. C., Greenhill, S. J., and Gray, R. D. (2011). Evolved structure of language shows lineage-specific trends in word order universals. *Nature* 473, 79–82. doi: 10.1038/nature09923

Dunn, M., Terrill, A., Reesink, G., Foley, R. A., and Levinson, S. C. (2005). Structural phylogenetics and the reconstruction of ancient language history. *Science* 309, 2072–2075. doi: 10.1126/science.1114615

Dyen, I., Kruskal, J., and Black, P. (1992). An Indoeuropean classification: a lexicostatistical experiment. *Trans. Philol. Soc.* 82, 1–132.

Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. U.S.A.* 95, 14863–14868. doi: 10.1073/pnas.95.25.14863

Emonds, J., and Faarlund, J. T. (2014). *English: the Language of the Vikings*. Olomouc: Palacký University.

Everaert, B. M., Huybregts, M. A. C., Chomsky, N., Berwick, R. C., and Bolhuis, J. J. (2015). Structures, not strings: linguistics as part of the cognitive sciences. *Trends Cogn. Sci.* 19, 729–743. doi: 10.1016/j.tics.2015.09.008

Fanciullo, F. (1988). "Lukanien/Lucania," in *Lexikon der Romanistischen Linguistik 4*, eds G. Holtus, M. Metzeltin, and C. Schmitt (Tübingen: Niemeyer), 669–688.

Fanciullo, F. (1997). "Basilicata," in *The Dialects of Italy*, eds M. Maiden and M. Parry (New York, NY: Routledge), 349–354.

Felsenstein, J. (2005). *PHYLIP (Phylogeny Inference Package) Version 3.6*. Seattle: Department of Genome Sciences, University of Washington.

Fodor, J. D., and Sakas, W. G. (2017). "Learnability," in *The Oxford Handbook of Universal Grammar*, ed. I. Roberts (Oxford: Oxford University Press), 249–269.

Franzoi, L., Sgarro, A., Dinu, A., and Dinu, L. P. (2020). "Random Steinhaus distances for robust syntax-based classification of partially inconsistent linguistic data," in *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, eds M. J. Lesot, S. M. Vieira, M. Z. Reformat, J. P. Carvalho, A. Wilbik, B. Bouchon-Meunier, et al. (Berlin: Springer Verlag), 17–26. doi: 10.1007/978-3-030-50153-2_2

Gianollo, C., Guardiano, C., and Longobardi, G. (2008). "Three fundamental issues in parametric linguistics," in *The Limits of Syntactic Variation*, ed. T. Biberauer (Amsterdam: John Benjamins), 109–142. doi: 10.1075/la.132.05gia

Gray, R. D., and Atkinson, Q. D. (2003). Language tree divergence times support the Anatolian theory of Indo-European origin. *Nature* 426, 435–439. doi: 10.1038/nature02029

Gray, R. D., Bryant, D., and Greenhill, S. J. (2010). On the shape and fabric of human history. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 365, 3923–3933. doi: 10.1098/rstb.2010.0162

Greenberg, J. H. (1987). *Language in the Americas*. Stanford: Stanford University Press.

Greenhill, S. J., Wu, C.-H., Hua, X., Dunn, M., Levinson, S. C., and Gray, R. D. (2017). Evolutionary dynamics of language systems. *Proc. Natl. Acad. Sci. U.S.A.* 114, E8822–E8829. doi: 10.1073/pnas.1700388114

Greenhill, S. J., Heggarty, P., and Gray, R. D. (2020). "Bayesian Phylolinguistics", in *The Handbook of Historical Linguistics, Volume II*, eds R. Janda, B. Joseph and B. Vance (Oxford: Blackwell), 226–253.

Guardiano, C., and Longobardi, G. (2005). "Parametric comparison and language taxonomy," in *Grammaticalization and Parametric variation*, eds. M. Battlori, M.-L. Hernanz, C. Picallo and F. Roca (Oxford: Oxford University Presss), 149–174.

Guardiano, C., and Longobardi, G. (2017). "Parameter theory and parametric comparison," in *The Oxford Handbook of Universal Grammar*, ed. I. Roberts (Oxford: Oxford University Press), 377–398.

Guardiano, C., Longobardi, G., Cordoni, G., and Crisma, P. (2020). "Formal syntax as a phylogenetic method," in *The Handbook of Historical Linguistics, Volume II*, eds R. Janda, B. Joseph, and B. Vance (Oxford: Blackwell), 145–182. doi: 10.1002/9781118732168.ch7

Guardiano, C., Michelioudakis, D., Ceolin, A., Irimia, M. A., Longobardi, G., Radkevich, N., et al. (2016). South by Southeast. A syntactic approach to Greek and Romance microvariation. *L'Italia Dialettale* 77, 95–166.

Guardiano, C., and Stavrou, M. (2014). Greek and Romance in Southern Italy: history and contact in nominal structures. *L' Italia Dialettale* 75, 121–147.

Guardiano, C., and Stavrou, M. (2019). Adjective-noun combinations in Romance and Greek of Southern Italy. Polydefiniteness revisited. *J. Greek Ling.* 19, 3–57. doi: 10.1163/15699846-01901001

Guardiano, C., and Stavrou, M. (2020). Dialectal syntax between persistence and change. The case of Greek demonstratives. *L'Italia Dialettale* 81, 119–155.

Gyarmathi, S. (1799). *Affinitas linguae hungaricae cum linguis fennicae originis grammatice demonstrata nec non vocabularia dialectorum tataricarum et slavicarum cum hungarica comparata*. Dieterich, Göttingen. Reprinted Indiana University, Bloomington, 1968; English translation with an introduction and notes by Victor E. Henzeli (Amsterdam Classics in Linguistics, 15). Amsterdam/Philadelphia: John Benjamins, 1983.

Hamming, R. W. (1950). Error detecting and error correcting codes. *Bell Syst. Tech. J.* 29, 147–160. doi: 10.1002/j.1538-7305.1950.tb00463.x

Hauser, D., Yang, C., Berwick, R. C., Tattersall, I., Ryan, M. J., Watumull, J., et al. (2014). The mystery of language evolution. *Front. Psychol.* 5:401. doi: 10.3389/fpsyg.2014.00401

Hauser, M. D., Chomsky, N., and Fitch, T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science* 298, 1569–1579. doi: 10.1126/science.298.5598.1569

Heine, B., and Kuteva, T. (2005). *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press.

Hervás y Panduro, L. (1778–1787). *Idea dell'Universo, che contiene la Storia della vita dell'uomo, elementi cosmografici, viaggio estatico al mondo planetario, e Storia della Terra*. Cesena: Gregorio Biasini all'Insegna di Pallade.

Hervás y Panduro, L. (1800–1805). *Catálogo de las lenguas de las naciones conocidas y numeración, división y clases de éstas según la diversidad de sus idiomas y dialectos*. Madrid: Impr. Administración del Real Arbitrio de Benificencia.

Holland, B. R., Huber, K. T., Dress, A., and Moulton, V. (2002). δ Plots: a tool for analyzing phylogenetic distance data. *Mol. Biol. Evol.* 19, 2051–2059. doi: 10.1093/oxfordjournals.molbev.a004030

Honeybone, P. (2016). Are there impossible changes? $\theta > f$ but $f \not> \theta$. *Pap. Histor. Phonol.* 1, 316–358.

Hornstein, N., and Lightfoot, D. (eds) (1981). *Explanation in Linguistics. The Logical Problem of Language Acquisition*. London/New York: Longman.

Hutterer, C. J. (1975). *Die Germanischen Sprachen. Ihre Geschichte in Grundzügen*. Budapest: Akadémiai Kiadó.

Jaccard, P. (1901). Étude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bull. Soc. Vaud. Sci. Nat.* 37, 547–579. doi: 10.5169/seals-266450

Jäger, G. (2015). Support for linguistic macrofamilies from weighted sequence alignment. *Proc. Natl. Acad. Sci. U.S.A.* 112/41, 12751–12757. doi: 10.1073/pnas.1500331112

Johanson, L., and Csató, ÉA. (eds) (1998). *The Turkic Languages*. London: Routledge.

Kaltz, B. (1985). "Christian Jacob Kraus," in *Linguarum totius orbis vocabularia comparativa*, ed. P. S. Pallas (St. Petersburg: Introduction, Translation and Notes), *Historiographia Linguistica*, 12, 229–260.

Karimi, S., and Piattelli-Palmarini, M. (eds) (2017). Special issue on parameters. *Ling. Analy.* 41 (3–4).

Kayne, R. (1975). *French Syntax. The Transformational Cycle*. Cambridge, MA: MIT Press.

Kayne, R. (2000). *Parameters and Universals*. Oxford/New York: Oxford University Press.

Keenan, E. (2002). "Explaining the creation of reflexive pronouns in English," in *Studies in the History of the English Language: a Millennial Perspective*, eds D. Minkova and R. Stockwell (Berlin/New York: Mouton de Gruyter), 325–354. doi: 10.1515/9783110197143.3.325

Keenan, E. (2009). "Linguistic theory and the historical creation of English reflexives," in *Historical Syntax and Linguistic Theory*, eds P. Crisma and G. Longobardi (Oxford: Oxford University Press), 17–40. doi: 10.1093/acprof:oso/9780199560547.003.0002

Khabtagaeva, B. (2018). "The role of Ewenki VgV in Mongolic reconstructions," in *Philology of the Grasslands: Essays in Mongolic, Turkic, and Tungusic studies*, eds B. Kempf, Á. Bertalan Apatóczky, and C. P. Atwood (Leiden: Brill), 174–193. doi: 10.1163/9789004351981_011

Klima, E. S. (1964). Relatedness between linguistic systems. *Language* 40, 1–20. doi: 10.2307/411919

Klima, E. S. (1965). *Studies in Diachronic Transformational Syntax*, Ph. D thesis, Harvard, Cambridge, MA.

Kroch, A. S. (1989). Reflexes of grammar in patterns of language change. *Lang. Var. Chang.* 1, 199–244. doi: 10.1017/s0954394500000168

Kraus, K. J. (1787). Review of ST PETERSBURG: Srawniteljnjije Slowari wsjech Jasikow i Narjetschii sobrannjije Desnizeju wsewisotschaisei Osobji, etc., das ist: Vergleichendes Glossarium aller Sprachen und Mundarten, gesammelt auf Veranstaltung der allerhöchsten Person. Erste Abtheilung, die europäischen und asiatischen Sprachen enthaltend. Erster Band, *Allgemeine Literaturzeitung*, 4 (numm. 235, 236, 237a, 237b), cols.1–29. Partly reprinted in Arens (1969, 136–145). English translation in Kaltz (1985).

Kylstra, A. D., Hahmo, S.-L., Hofstra, T., and Nikkilä, O. (1991). *Lexikon der Älteren Germanischen Lehnwörter in den Ostseefinnischen Sprachen, Band I: A-J*. Amsterdam: Rodopi.

Lausberg, H. (1939). *Die Mundarten Südlukaniens*. Halle: Niemeyer.

Ledgeway, A. (2006). "The dual complementiser system in southern Italy: *spirito greco, materia romanza*?," in *Rethinking Languages in Contact: the Case of Italian*, eds A. L. Lepschy and A. Tosi (Oxford: Legenda), 112–126. doi: 10.4324/9781351195515-9

Ledgeway, A. (2013). "Greek disguised as Romance? The case of Southern Italy," in *Proceedings of the 5th International Conference on Greek Dialects and Linguistic Theory*, eds M. Janse, B. Joseph, A. Ralli, and M. Bagriacik (Patras: University of Patras), 184–227.

Ledgeway, A., Schifano, N., and Silvestri, G. (2018). Il contatto tra il greco e le varietà romanze nella Calabria meridionale. *Lingue Antiche e Moderne* 7, 95–133. doi: 10.17863/CAM.30519

Lightfoot, D. W. (1979). *Principles of Diachronic Syntax*. Cambridge/New York: Cambridge University Press.

Lightfoot, D. W. (1982). *The Language Lottery. Toward a Biology of Grammar*. Cambridge, MA: MIT Press.

Lightfoot, D. W. (1997). Catastrophic change and learning theory. *Lingua* 100, 171–192. doi: 10.1016/s0024-3841(93)00030-c

Lightfoot, D. W. (2002). Myths and the prehistory of grammar. *J. Ling.* 38, 113–136. doi: 10.2307/4176715

Lightfoot, D. W. (2017). Discovering new variable properties without parameters. *Ling. Analy.* 41, 409–444.

List, J. M. (2014). *Sequence Comparison in Historical Linguistics*. Düsseldorf: Düsseldorf University Press.

Longobardi, G. (2001). Formal syntax, diachronic minimalism, and etymology: the history of French chez. *Ling. Inq.* 32, 275–302. doi: 10.1162/00243890152001771

Longobardi, G. (2003). Methods in parametric linguistics and cognitive history. *Ling. Variat Yearbk* 3, 101–138. doi: 10.1075/livy.3.06lon

Longobardi, G. (2005). "A minimalist program for parametric linguistics," in *Organizing Grammar*, eds H. Broekhuis, N. Corver, R. Huijbregts, U. Kleinhenz, and J. Koster (Berlin: DeGruyter Mouton), 407–414.

Longobardi, G. (2014). "Theory and experiment in parametric minimalism: the case of Romance negation," in *Language Description Informed by Theory*, eds R. Pensalfini, M. Turpin, and D. Guillemin (Amsterdam: John Benjamins), 217–261. doi: 10.1075/slcs.147.10lon

Longobardi, G. (2017). Principles, parameters and schemata. A radically underspecified UG. *Ling. Analy.* 41, 517–557.

Longobardi, G., Ghirotto, S., Guardiano, C., Tassi, F., Benazzo, A., Ceolin, A., et al. (2015). Across language families: genome diversity mirrors linguistic variation within Europe. *Am. J. Phys. Anthropol.* 157, 630–640. doi: 10.1002/ajpa.22758

Longobardi, G., and Guardiano, C. (2009). Evidence for syntax as a signal of historical relatedness. *Lingua* 119, 1679–1706. doi: 10.1016/j.lingua.2008.09.012

Longobardi, G., Guardiano, C., Silvestri, G., Boattini, A., and Ceolin, A. (2013). Toward a syntactic phylogeny of modern Indo-European languages. *J. Histor. Ling.* 3, 122–152. doi: 10.1075/jhl.3.1.07lon

Loporcaro, M. (2009). *Profilo linguistico dei dialetti italiani*. Bari: Laterza.

Macklin-Cordes, J. L., Bowern, C., and Round, E. R. (2020). Phylogenetic signal in phonotactics. *arXiv* [Preprint]. Available online at: https://arxiv.org/abs/2002.00527#:~{}:text=Phylogenetic%20methods%20have%20broad%20potential%20in%20linguistics%20beyond%20tree%20inference.&text=We%20extract%20phonotactic%20data%20from,the%20data%20reflect%20phylogenetic%20history (accessed July 9, 2020).

Maddison, W., and Maddison, D. (2007). *Mesquite 2. A Modular System for Evolutionary Analysis*. Available online at: http://www.mesquiteproject.org

Marcantonio, A. (2002). *The Uralic Language Family: Facts, Myths and Statistics*. Boston: Blackwell.

Martino, P. (1991). *L' "area Lausberg" fra isolamento e arcaicità*. Roma: Dipartimento di Studi Glottoantropologici, Università di Roma La Sapienza.

McMahon, A. (2010). "Computational models and language contact," in *The Handbook of Language Contact*, ed. R. Hickey (Hoboken, NJ: Wiley Blackwell), 128–147. doi: 10.1002/9781444318159.ch6

McMahon, A., and McMahon, R. (2005). *Language Classification by Numbers*. Oxford: Oxford University Press.

Menges, K. H. (1968). *The Turkic Languages and Peoples: An Introduction to Turkic Studies*. Wiesbaden: Harrassowitz.

Moran, S., and McCloy, D. (eds) (2019). *PHOIBLE 2.0*. Jena: Max Planck Institute for the Science of Human History.

Morpurgo-Davies, A. (1992/2014). *Nineteenth Century Linguistics, History of Linguistics, Vol. IV*. London/New York: Routledge.

Newmeyer, F. (2005). *Possible and Probable Languages: a Generative Perspective on Linguistic Typology*. Oxford: Oxford University Press.

Nichols, J. (1992). *Linguistic Diversity in Space and Time*. Chicago and London: University of Chicago Press.

Pellegrini, G. B. (1977). *Carta dei dialetti d'Italia*. Pisa: Pacini.

Pintzuk, S., and Kroch, A. (1995). The Dating of Beowulf. ms. University of York-University of Pennsylvania.

Podani, J., and Miklos, I. (2002). Resemblance coefficients and the horseshoe effect in principal coordinates analysis. *Ecology* 84, 2347–2363.

Port, A., Karidi, T., and Marcolli, M. (2019). Topological analysis of syntactic structure. *arXiv* [Preprint]. Available online at: https://arxiv.org/abs/1903.05181?#:~{}:text=We%20use%20the%20persistent%20homology,syntactic%20structures%20of%20world%20languages.&text=We%20show%20there%20are%20relations,relations%20that%20are%20family%2Dspecific (accessed July 9, 2020).

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., and Suchard, M. A. (2018). Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* 67, 901–904. doi: 10.1093/sysbio/syy032

Rensch, K. (1973). *Nordkalabrischer Sprachatlas anhand der Parabel vom verlorenen Sohn*. The Hague: Mouton.

Rigon, G. (2009). *A Quantitative Approach to the Study of Syntactic Evolution*. Tesi di Dottorato, Università di Pisa.

Ringe, D., Warnow, T., and Taylor, A. (2002). Indo-European and computational cladistics. *Trans. Philol. Soc.* 100, 59–129. doi: 10.1111/1467-968x.00091

Rizzi, L. (1978). "A restructuring rule in Italian," in *Recent Transformational Studies in European Languages*, ed. S. J. Keyser (Cambridge, MA: MIT Press), 113–158.

Rizzi, L. (1982). *Issues in Italian Syntax*. Dordrecht: Foris.

Robbeets, M. (2005). *Is Japanese Related to Korean, Tungusic, Mongolic and Turkic?*. Wiesbaden: Harrassowitz.

Robbeets, M. (2008a). "If Japanese is Altaic, why can it be so simple?," in *Evidence and Counterevidence. Essays in Honor of Frederik Kortlandt, vol. 2 (General Linguistics), Series Studies in Slavic and General Linguistics 33*, eds A. Lubotsky, J. Schaeken, and J. Wiedenhof (Amsterdam: Rodopi), 337–368. doi: 10.1163/9789401206365_022

Robbeets, M. (2008b). The historical comparison of Japanese, Korean and the Trans-Eurasian languages. *Rivista degli studi orientali* 81, 261–287.

Roberts, I. (1998). Review of: Harris, A., and Campbell, L. Historical syntax in cross-linguistic perspective. *Roman. Philol.* 51, 363–370.

Roberts, I. (2019). *Parameter Hierarchies and Universal Grammar*. Oxford: Oxford University Press.

Roberts, I., and Roussou, A. (2003). *Syntactic Change*. Cambridge: Cambridge University Press.

Rohlfs, G. (1972). *Nuovi scavi linguistici nell'antica Magna Graecia*. Palermo: Istituto Siciliano di Studi Bizantini e Neoellenici.

Romito, L., Turano, T., Loporcaro, M., and Mendicino, A. (1996). "Micro e macrofenomeni di centralizzazione nella variazione diafasica. Rilevanza dei dati fonetico-acustici per il quadro dialettologico calabrese," in *Fonetica and fonologia dell'italiano parlato*, ed. F. Cutugno (Roma: Esagrafica), 157–175.

Ronquist, F., and Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574. doi: 10.1093/bioinformatics/btg180

Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425. doi: 10.1093/oxfordjournals.molbev.a040454

Sakas, W. G., Yang, C., and Berwick, R. (2017). Parameter setting is feasible. *Ling. Analy.* 41, 391–408.

Sandfeld, K. (1930). *Linguistique balkanique: problèmes et résultats*. Paris: Honoré Champion.

Sarno, S., Boattini, A., Carta, M., Ferri, G., Alù, M., Yao, D. Y., et al. (2014). An ancient Mediterranean melting pot: investigating the uniparental genetic structure and population history of Sicily and Southern Italy. *PLoS One* 9:e96074. doi: 10.1371/journal.pone.0096074

Schmidt, H. A., Strimmer, K., Vingron, M., and von Haeseler, A. (2002). TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* 18, 502–504. doi: 10.1093/bioinformatics/18.3.502

Schönig, C. (1997-1998). A new attempt to classify the Turkic languages I-III. *Turkic Lang.* 117, 130–151.

Shimunek, A. (2017). *Languages of Ancient Southern Mongolia and North China*. Wiesbaden: Harrassowitz.

Silvestri, G. (2013). *The Nature of Genitive Case*. Tesi di Dottorato, Università di Pisa.

Smail, D. L. (2008). *On Deep History and the Brain*. Los Angeles: University of California Press.

Sneath, P. H. A., and Sokal, R. R. (1973). *Numerical Taxonomy. The Principles and Practice of Numerical Classification*. San Francisco: W H Freeman and Company.

Soucek, S. (2000). *A History of Inner Asia*. Cambridge: Cambridge University Press.

Starostin, S. A. (1996). Comments on the Basque-Dene-Caucasian comparisons. *Mother Tong.* 2, 101–109.

Stenbrenden, G. F. (2016). Why English is not dead: a rejoinder to Emonds and Faarlund. *Folia Ling. Hist.* 50, 239–279. doi: 10.1515/flih-2016-0008

Swofford, D. L. (2001). *PAUP*: Phylogenetic Analysis Using Parsimony (and Other Methods) 4.0. B5*. Sunderland, MA: Sinauer Associates.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739. doi: 10.1093/molbev/msr121

Taraldsen, K. T. (1980). *On the Nominative Island Constraint, Vacuous Application and the that-Trace Filter*. Bloomington: Indiana University Linguistics Circle.

Tekin, T. (1994). "Altaic languages," in *The Encyclopedia of Language and Linguistics*, Vol. 1, ed. R. E. Asher (New York, NY: Pergamon Press), 82–85.

Thomason, S. G., and Kaufman, T. (1988). *Language Contact, Creolization and Genetic Linguistics*. Berkeley: University of California Press.

Toman, J. (1987). Not from 1903, not from Meillet. A final (?) remark on "Où tout se tient". *Historiographia Linguistica* 14, 403–406.

Tomić, O. M. (2006). *Balkan Sprachbund Morpho-Syntactic Features*. Dordrecht: Springer.

Vezzosi, M. (2019). *Le strutture nominali del dialetto casalasco: analisi di comparazione parametrica*. Tesi di Laurea. Università di Modena e Reggio Emilia.

Vovin, A. (2005). The end of the Altaic controversy. In memory of Gerhard Doerfer. *Central Asiat. J.* 49, 71–132. doi: 10.2307/41928378

Whitman, J. (2012). "The relationship between Japanese and Korean," in *The languages of Japan and Korean*, ed. N. Tranter (London: Routledge), 24–38.

Wichmann, S., and Holman, E. W. (2009). *Temporal Stability of Linguistic Typological Features*. Munich: Lincom Europa.

Wichmann, S., Holman, E. W., and Brown, H. (eds) (2020). *The ASJP Database (version 19)*.

Wichmann, S., Holman, E. W., Walker, R., and Rama, T. (2011). Correlates of reticulation in linguistic phylogenies. *Lang. Dyn. Chang.* 1, 205–240. doi: 10.1163/221058212x648072

Wichmann, S., and Saunders, A. (2007). How to use typological databases in historical linguistic research. *Diachronica* 24, 373–404. doi: 10.1075/dia.24.2.06wic

Wurm, S. A., Mühlhäusler, P., and Tryon, D. T. (eds). (2011). *Atlas of Languages of Intercultural Communication in the Pacific, Asia, and the Americas*, Vol I: Maps. Vol II: Texts. Berlin: Mouton de Gruyter.

Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586–1591. doi: 10.1093/molbev/msm088

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.