DISHARMONIC HEADEDNESS IN HOMERIC GREEK AND TOCHARIAN AND IMPLICATIONS FOR PROTO-INDO-EUROPEAN RECONSTRUCTION*

Ryan Windhearn
Cornell University

Abstract Traditionally in comparative reconstruction, innovations are not useful for informing the reconstruction of proto-languages (and for good reason, though they are of course useful for subgrouping purposes). In this paper, however, I will show that due to the unique hierarchical properties of syntax, innovation can in fact reveal inherited structural relationships that would otherwise remain opaque. Specifically, based on the complementizers innovated across the ancient Indo-European (IE) languages, along with the strikingly parallel word orders seen in the auxiliary constructions innovated across these languages, I propose a disharmonically headed reconstruction of Proto-Indo-European (PIE) clause structure, with PIE being left-headed in the C domain and right-headed in the T domain.

1 INTRODUCTION

Traditionally in comparative reconstruction, innovations are not useful for informing the reconstruction of proto-languages (and for good reason, though they are of course useful for subgrouping purposes). In this paper, however, I will show that due to the unique hierarchical properties of syntax, innovation can in fact reveal inherited structural relationships that would otherwise remain opaque. To demonstrate this, based on the complementizers innovated across the ancient Indo-European (IE) languages, along with the strikingly parallel word orders seen in the auxiliary constructions innovated across these languages (both of which act as phonological cues for shared syntactic featural retentions), I propose a disharmonically headed reconstruction of Proto-Indo-European (PIE) clause structure, with PIE being left-headed in

the C domain and right-headed in the T domain.

The rest of this introduction discusses the theoretical assumptions that allow for rigorous syntactic reconstruction along with the specific structural assumptions that underly my own analyses here. In §2 I provide a case study of complementizer development across the old Indo-European daughter languages to illustrate how the Comparative Method may be applied to syntactic functional heads even when associated phonological forms cannot be securely reconstructed. §3 provides my own corpus work from Homeric Greek and Tocharian to illustrate the structural parallels seen in their innovated auxiliary constructions, and also gives an overview of the relevant literature for four other branches of Indo-European to demonstrate the extent of the syntactic similarities seen across the family. In §4 I use these similarities to reconstruct disharmonic headedness for PIE and conclude the paper.

1.1 Theoretical assumptions

Robust syntactic reconstruction has often been called implausible for a variety of reasons, most notably due to the difficulty of setting up appropriate correspondence sets.[1] Lightfoot (2002) clearly explains the problem: lexical information can be reconstructed since it is stored intact in a mental lexicon that is transferred to new generations of speakers, while syntax is procedurally built for every utterance. As a result, sentences generated by speakers are not stored in the lexicon, are not passed directly to new speakers, and therefore (according to the argument) cannot be compared to each other for reconstructive purposes. If, however, we could relegate some portion of syntax to the lexicon, we could reconstruct syntax in the same way that we reconstruct phonology and morphology.

The Minimalist Program of Chomsky (1995) offers one solution to the correspondence problem, adopting two critical assumptions about syntactic structure that enable us to deal with the correspondence set problem and provide a strong theoretical foundation for rigorous syntactic reconstruction. First, the computational component of the syntax is assumed to be universal. It is a set of operations assumed to be part of human cognitive makeup and thus invariant over time, space, and population. Second is the Borer-Chomsky conjecture (Baker 2008: 353): the *input* to this universal computational process, features associated with lexical items and functional heads, are stored in the mental lexicon and *do* change. As a result, all synchronic and diachronic syntactic variation is due to the featural differences of lexical items and functional heads, and is not due to the structure-building

---

1 Refer to Ferraresi & Goldbach (2008) for an overview of the history of syntactic reconstruction, from Watkins (1976)'s criticism of typological approaches to syntactic reconstruction, to Kiparsky (1995)'s pioneering application of modern syntactic theory to Germanic reconstruction, to current Minimalist approaches to diachronic syntax.

component of the syntax. As Pires & Thomason (2008) put it, "the task of syntactic reconstruction can then be restricted to identifying variation in the feature specification of (functional) lexical items that determine syntactic structure and syntactic variation" (p. 41).

As noted in the literature (Hale 1998, Roberts & Roussou 2003, Pires & Thomason 2008, Walkden 2014, etc.) this formulation of syntactic variation and change solves the correspondence set problem,[2] since the elements that drive syntactic variation are now stored in the lexicon. We can, therefore, reconstruct syntactic proto-forms using the same tools we use for reconstructing phonological and morphological proto-forms. Importantly, note that the ability to reconstruct syntax in this manner is not inherently limited to Minimalism: as discussed in Walkden (2014), any formalism can be used for reconstruction if it gives us lexical items to reconstruct. As such, this paper is not an attempt to use syntactic reconstruction to make arguments in favor of Minimalism; rather, it is an attempt to use Minimalism to make arguments in favor of a specific syntactic reconstruction. The theoretical contributions of this paper explore the implications of these Minimalist assumptions for reconstructive theory, arguing (1) that syntactic reconstruction is possible even when phonological or morphological reconstruction is not and (2) that the innovation of novel constructions can improve our understanding of inherited syntactic structure.

There are other similar approaches to syntactic comparison. The Parametric Comparison Method (PCM) of Longobardi (2003), for example, uses syntactic parametric variation as the sole basis for language phylogeny. The crucial difference between Longobardi's syntactic comparison and the syntactic reconstruction I will undertake here is that while the PCM is concerned mainly with language phylogeny and does not attempt to reconstruct proto-forms, my parametric analysis of languages already known to be related is specifically intended to produce rigorous reconstructions of proto-language syntax.[3]

---

2 Partially, at least. Walkden (2014: 50-57)'s Double Cognacy Condition explains how correspondence sets must be composed of cognate forms, which themselves occur in cognate contexts. For phonological reconstruction, this means that the cognate sounds being reconstructed must occur in the same location in words that are themselves cognate. According to Walkden, the Double Cognacy Condition cannot be met for syntactic reconstruction, since the cognate features in question do not occur in contexts that are themselves cognate, but he explains how this correspondence problem can partially be rectified through examination of the distribution of lexical items across structures and the use of phonological clues.

3 For a discussion of how the PCM compares to the Minimalist reconstruction here, see Pires & Thomason (2008: 29). In brief, my approach here uses parametric variation as input to the Comparative Method, which reconstructs proto-forms to determine genetic relationships between languages, while the PCM is more a comparison of typological similarity, and does not attempt to reconstruct proto-forms. Pires and Thomason note that "[Longobardi] intends his use of the term 'reconstruction of phylogenetic relations' to exclude actual historical comparative reconstruction."

In this paper I will be working within the Minimalist Program of the Principles and Parameters framework first introduced in Chomsky (1982). Accordingly, after Whitman (2008), I also assume that phrasal heads may select their complements on the left or the right; specifiers, however, are assumed to be invariantly on the left. Adopting the Cyclic Linearization model of Fox & Pesetsky (2005), I assume that headedness is determined by an interpretable feature, [Comp:Left] or [Comp:Right], that determines whether complements are linearized to the right or to the left of their heads at PF, where the feature is interpreted. At the end of each phase the relative ordering of words is fixed, and this relative ordering must not be contradicted by later phases. For this paper, CP and vP constitute phases. Finally, I will assume that "disharmonic headedness", where functional projections in a language may have different headedness, is allowed by the syntax. My choice of headedness for a given functional projection is determined by (1) economy of movement considerations, favoring the headedness analysis that accounts for the most data with the fewest motivated movements and (2) disfavoring headedness analyses that require unmotivated or unnatural movements given the data.

I use "CP" to refer to all projections in the clause above TP, including the left periphery or "expanded CP" as described by Rizzi (1997). Similarly, "within TP" refers to TP and everything between TP and the projection where the subject is externally merged (vP or VoiceP), excluding obviously semantically specialized projections like NegP. For heads within CP, after Walkden (2014)'s reconstruction of Germanic, I assume uninterpretable features corresponding to their heads (e.g. [uFoc] for Foc^0) as well as the [Comp] headedness feature. For TP, I assume the following features: [uϕ] to express ϕ-features, [uV] to motivate v-to-T movement, T(ense)-A(spect)-M(ood) features to trigger TAM morphology, and the [Comp] headedness feature.

Note that when reconstructing functional categories, I only reconstruct features that I can be confident of, and leave the others unspecified rather than speculating. This is consistent with reconstruction methodology for phonology and morphology, as seen, for example, in the reconstruction of the PIE "laryngeals". Indo-Europeanists determined that certain vowels had been colored by adjacent segments that were more sonorant than stop consonants, but less sonorant than the vowels themselves. Researchers therefore reconstruct the feature-poor resonants commonly known as "laryngeals", with unknown qualities that resulted in the coloring of adjacent vowels. In the same way, the features I am primarily concerned with reconstructing for C and T here are [Comp:Left] and [Comp:Right] headedness features, but I will also reconstruct other features where the data allows.

1.1.1 Final Concerns

There are two final considerations I would like to address. The first is discussed by Lehmann (2005) and Balles (2008), who assert that research on diachronic syntax and reconstruction should be supplemented with a theory of syntactic change. The second is brought up by Walkden (2014), who cautions that reconstruction should proceed to a greater time depth only when the acceptance of shallower reconstructions are agreed upon by the scholarly community. Both of these concerns should be kept in mind; indeed, the only reasons the reconstruction offered in the current paper is undertaken are (1) the fact that the correspondence sets presented here overwhelmingly agree in the nature of both their generalizations and their exceptions and (2) the fact that the headedness features being reconstructed do not change from the parent language to the any of the daughter languages, so no syntactic change has occurred.

2 RECONSTRUCTING SYNTACTIC FEATURES WITHOUT RECONSTRUCTING PHONOLOGICAL FEATURES

As a result of the Minimalist assumptions discussed above, we should not only be able to reconstruct syntactic features in the same way that we reconstruct phonology and morphology, but we should even be able to reconstruct syntactic features in the absence of any accompanying reconstructible phonology or morphology. This section provides an example of such a syntactic reconstruction.

2.1 The case of complementizers in PIE

In this section I propose an uncontroversial reconstruction of a [Comp:Left] feature of the C head for PIE, based on a combination of argument complement clause data from six of the earliest IE languages and other arguments for left-headedness in CP from the literature. This reconstruction not only straightforwardly illustrates the nature of the "innovation-based reconstruction" I propose, but also serves to bolster the existing literature reconstructing left-headedness for CP in PIE. I focus on complement clause data instead of incorporating relative or other clause data for three reasons. First, most of the existing literature arguing for left-headedness in CP for PIE focuses mainly on relative and adverbial clauses, so this analysis addresses the "third pillar" of embedded clauses. Second, by restricting my analysis to argument complement clauses, there is some simplification of the left periphery situation in the data. Third, demonstrating conclusively that any CP head is on the left effectively demonstrates left-headedness for the entire expanded CP. Cross-linguistically we have evidence of C-type heads on one side of the

derivation and T-type heads on the other, but not cases of one C-type head being on the left, and another on the right; there are no cases of Force being on the left and Topic on the right, for example.[4]

No discussion of argument complementation in PIE would be complete without mentioning the debate surrounding the nature of subordination in reconstructed PIE. The original position taken by researchers, concisely reviewed and summarized by Kiparsky (1995)'s influential analysis of the development of V2 syntax in Germanic, is that PIE originally had no finite subordinate clausal embedding, instead relying solely on adjunction and correlative constructions. Over time, according to the analysis, the PIE daughter languages separately developed CP structure, clausal embedding, and the complementizers that go with them. Kiparsky takes a lack of reconstructible complementizers as evidence of a lack of CP structure in PIE ("there were no complementizers, and therefore no CP, and no embedding", p.153), but as we will see, the assumptions we've made about lexical storage of syntactic features will allow us to reconstruct CP features without needing to reconstruct individual complementizers.

In addition to objections on the basis of the continuity hypothesis, the idea that all universal properties of current grammars also held for any historical human language (e.g. Pires & Thomason 2008: 40), evidence has recently come to light that may shift the *communis opinio*. Probert (2014) argues that clausal embedding should be reconstructed for the earliest stage of PIE, noting that the infrequency of clausal embedding in the older daughter languages is not an *absence* in any branch, and citing evidence that the infrequency seen in the earliest stages of the daughter languages are due more to literary genre than grammar. The main reason for not reconstructing embedded relative clauses for PIE, she claims, has been due to the belief that the earliest-attested Anatolian languages do not appear to contain the structures. In response, she provides an example of an embedded construction in Old Hittite, and also claims that the construction in general is considerably more common in later Anatolian than previously thought. As embedded clauses are present in the earliest corpora of all of the IE daughter languages, she argues that we have no reason not to reconstruct them for the proto-language as well.

For the purposes of this paper, we will follow Probert in reconstructing both clausal embedding and CP for PIE, though the debate is likely far from decided.

---

4 Thanks to John Whitman for this discussion.

2.1.1 Reconstructing complementizers for PIE

As Kiparsky noted, argument complementizers (henceforth 'complementizers') across the early IE languages are not all cognate. Latin *quod*, Hittite *kuit*, and Tocharian $k_u ce/kucne$ are from the PIE interrogative stem *$k^w o$-, Gothic *þatei* comes from the demonstrative pronoun **to-*, and Sanskrit *yád* and Greek *hóti* and *hoːs* are from the PIE relative pronoun stem *$H\hat{i}o$-*.[5]

Further, as shown by Hackstein (2013), for most of these languages complementizer behavior developed within their attested history. Latin *quod* was only extended from use with factive verbs in the Classical period. Sanskrit *yád* develops its own complementizer usage from relative usage during the Classical period. Hittite *kuit* and Tocharian $k_u ce/kucne$ start as the heads of adverbial adjuncts which later develop true argument complementizer usage.

So, a single overt complementizer cannot be reconstructed for PIE due to the fact that 1) the complementizers in early IE languages are not cognate and 2) overt complementizer behavior itself is innovative in these languages. After Hackstein, zero-embedding is likely the only reconstructible method for embedding sentential complements after verbs of utterance and cognition for PIE. If we follow much of modern generative theory in assuming that all languages project a CP, and that complementizers fill a functional head C, then even without any reconstructible complementizer, we know that PIE had a C head – we just need to know whether it was left-headed or right-headed.

2.2 Setting up a correspondence set

Even if PIE used zero-embedding, it still projected a CP and has a functional head C; C in PIE is just filled with a null complementizer, or a complementizer that we can no longer reconstruct. So, instead of trying to reconstruct both the phonological form *and* headedness of C, I reconstruct just the headedness itself, regardless of what phonological form this position takes in the daughter languages. I therefore set up a correspondence set for the underlying syntactic structure, composed of the feature sets of all of the innovated complementizers of the daughter languages, and ignore the specific phonological form of each complementizer.

---

5 One reviewer asks if the common nominal origin of these complentizers could play a role in the parallel complementation structures we see in the daughter languages. This is possible, but it's unlikely that all three originate through the same mechanism. The complementizers originating from relatives and interrogatives both likely arose through Spec-Head reanalysis, but the complementizer from the demonstrative would likely have had a different structural origin, possibly Head-Head reanalysis. A diachronic analysis that could unify all three of these developmental pathways would indeed be interesting, and could strengthen the case for just how similar argument complement clause syntax is across the daughter languages.

When we ignore the phonological form of the complementizers, we immediately see striking similarities in the daughter languages' syntax in clauses embedded after verbs of speaking or cognition:[6]

(1) **Hittite**

IDI [***kuit***=*za* KUR ^URU*Mizri* KUR ^URU*Ḫatti*=*ya* 1^EN KUR^TIM
know-1SG [COMP land Egypt land Hatti=and one land
kišari]
become-3SG]

"[...] I know that the land of Egypt and the land of Hatti are becoming one land"   (KUB XXI 38 Rs. 13f.)

(2) **Tocharian B**

poñ [***ce*** *ñiś te-ñemtsa* *pañäkte saim*
say.IMP [COMP I this-name-PERL Buddha refuge-OBL.SG.M
yamaskemar]
make-PRS.1SG.MP]

"Say that I, named so-and-so, take the Buddha as refuge!"
                                        (IOL Toch. 92,4)

(3) **Sanskrit**

vaktavyaṃca [***yac*** *candras* *tvām* *atra*
tell-GV.NOM.SG.N [COMP moon-NOM.SG you-2SG.ACC here
hrada *āgacchantaṃ* *niṣedhayati*]
lake-LOC.SG go-PART.ACC.SG.M forbid-3SG]

"[...] he ought to be informed (namely) that the moon forbids you to go to this lake"   (*Pañcatantra* 160.24)

(4) **Homeric Greek**

gnoːtòn [...] *èstin* [***hoːs*** *eːdeː* *Troːessin*
known-NOM.SG.N [...] COP.3SG [COMP already Trojan-DAT.PL
oletʰrou *peirat′* *epʰeːptai*]
destruction-GEN.SG end-N/ACC.PL bound-PERF.MID.3SG]

"It is [...] known that already the ends of destruction are bound on the Trojans"   (Iliad 7.402)

---

6 These examples are all from Hackstein (2013).

(5) **Old Latin**

te   *scire*   *audivi*   [...] [*quod cum*

you-ACC.SG know-INF hear-PERF.1SG.ACT [...] [COMP with

peregrini   *cubui*   *uxore*   *militis*]

foreign-GEN.SG.M sleep-PERF.1SG.ACT wife-ABL.SG.F soldier-GEN.SG.M]

"I heard you know [...] that I slept with the foreign soldier's wife."

(Plautus, Bacchides 1007f.)

(6) **Gothic**

gamelid ist   [*þatei ni bi hlaib*   *ainana libaid*

written COP.3SG [COMP not by bread-ACC.SG alone live-3SG

manna]

man-NOM.SG]

"[...] that man shall not live by bread alone"   (Luke 4:4, Katz 2019)

Every innovated argument complementizer in the daughter languages, when it ultimately shows up, appears clause-initially the vast majority of the time, and can only be preceded by a small class of fronted elements, as shown by Hale (1987). Even with Rizzi (1997)'s split CP model that allows for landing sites above argument complementizers in the left periphery, assuming that the complementizers in (1) through (6) above are all base-generated in their respective C-heads and have not moved somehow, there are no movement operations to my knowledge that can derive these word orders from a right-headed CP.

I argue that these languages aren't all independently innovating a left-headed C domain; they're innovating a phonological form to fill the left-headed C that they already share. We see therefore that our correspondence set, composed of the [Comp:Right] feature of the C head in each of the early IE languages, unanimously points toward a null (or at least unreconstructible) clause-initial C for the proto-language.

Note that this is not "structural" syntactic innovation on the part of any of the daughter languages. The [Comp:Right] feature of C hasn't changed – just whether a phonological form is associated with this syntactic position. The parallel innovation of separate phonological forms to fill the same syntactic position in each of the daughter languages cues us in to the shared structural syntactic reality: CP was also left-headed in their ancestor, Proto-Indo-European.

This conclusion is corroborated by work on other IE complementizers and particles. Many scholars, for example, claim that Wackernagel ("second") position clitics in the old IE languages show behavior indicating that they likely head their own projections in the left periphery. Koller (2013) locates Tocharian A *ne* (as well as its Tocharian B cognate *nai*) in the head of FocP

since it immediately follows wh-phrases (which Koller places in spec-FocP) clause-initially. For Sanskrit, Hale (1996) places Wackernagel clitics in the C head. Danckaert (2012) explores the Latin left periphery in depth, coming to the overwhelming conclusion that functional heads within the expanded CP are left-headed. Finally, Scharf (2015) points out that the Sanskrit question particle *api* occurs clause-initially as seen in (7) below, instead of the clause final position we would expect if CP was right-headed (e.g. *ka* in Japanese, etc.).[7]

(7)   ***api*** *ete*   *asmatputrāḥ kalabhāṣiṇaḥ*   *padbhyām gaccheyuḥ*
       Q   these our.sons   softly.speaking by.feet   go
       "Will these baby-talking sons of ours walk?" (VP 4.2.43, Scharf 2015)

2.2.1   *Conclusion: PIE was left-headed in CP*

By comparing the syntactic features of the various innovated complementizers across the early Indo-European languages, we arrived at a reconstruction for Proto-Indo-European that supports the position taken by most of the IE literature: its CP was left-headed. This serves as an effective proof-of-concept, since it demonstrates the extent to which innovation of new lexical items to fill structural syntactic positions can preserve and even make explicit inherited syntactic relationships.

3   Using innovated auxiliary constructions to reconstruct TP-headedness

The remainder of this paper will be devoted to a much more controversial reconstruction: reconstructing a right-headed TP for PIE based on the similarities seen in separately innovated auxiliary constructions across the earliest-attested Indo-European daughter languages. In this section I will briefly discuss my methodology, give an overview of my corpus work in Tocharian and Homeric Greek, and review the relevant literature for four other ancient IE branches.

---

7 I would be remiss to leave out Sanskrit's clause-final quotative particle *iti* in this discussion of universal clause-initial complementizers across the Indo-European languages, but note that Hock (1982) and Saxena (1995) claim that its complementizer-like usage did not fully evolve until the classical period, and that even so it does not show true complementizer behavior, acting instead only as a particle indicating quotations. According to Biberauer, Holmberg & Roberts (2014), who conclude that *iti* is part of a class of acategorial elements existing outside the extended projection, "we take it to be significant that we do not find this kind of [word] order with true subordinating Cs".

3.1 Why not just use reconstructible lexical verbs?

All other reconstructions of PIE clause structure have focused on the position of reconstructible lexical verbs, so why am I focusing exclusively on periphrastic auxiliary constructions? Put simply, due to the freedom of word order that we see in the ancient IE languages, combined with the multiple landing sites that verbs are able to target cross-linguistically, individual lexical verbs are much less useful for determining the location of inflection in the syntax than complementizers were for determining the location of C in the previous section. Mostly as a result of this structural ambiguity/freedom of movement, the argument about Proto-Indo-European's clausal headedness has lasted over a century, dating all the way back to the initial assertion of Delbrück (1900: 82-83) that PIE was SOV because Sanskrit was (mostly) SOV.

Looking only at auxiliary constructions eliminates much of this ambiguity. Most generative syntactic models generate auxiliaries in T either by external merge (e.g. the English modals), or by movement/internal merge (e.g. English BE and HAVE auxiliaries), and base-generate the auxiliaries' accompanying participles within VP much lower in the clause. Each of these elements may then be manipulated separately by syntactic processes, of course, but due to what we know about their initial syntactic relationship and the possible syntactic transformations that exist, the potential word-order relationships between their individual landing sites are more constrained. For example, clause-initial participle+auxiliary order is very different from clause-initial auxiliary+participle order, and this tells us much more about the syntax than just a clause-initial finite lexical verb. These relationships are therefore more transparent to reverse-engineering, allowing us to triangulate the locations of V and T relative to their arguments with a precision that is not possible using lexical verbs alone.

3.2 What constitutes an auxiliary construction?

For my purposes, an auxiliary construction is a periphrastic verbal construction consisting of a transparently monoclausal structure in which a semantically bleached verb (the auxiliary) moves to, or is generated in, T in order to express the ϕ-features of T. This auxiliary is accompanied by a participle or other closely related verbal adjective lower in the same clause.

I am excluding constructions composed of a modal verb and (usually) an infinitive, such as desiderative, volitional, or purpose constructions, etc., as these are often considered to be multi-clausal in nature, especially in these old IE languages.[8] I will briefly refer to embedded clause constructions throughout the paper as additional evidence of mixed headedness, but they

---

8 For one example of how early IE modal syntactic behavior is significantly different from that of auxiliaries, see Danckaert (2017: 224).

will not be the main focus of my analysis, since I will be most concerned with the relative positioning of verbal elements in the same clause. This restriction will ensure that my triangulation of φ-feature location relative to structural cases and the participle are as accurate as possible for determining the structure of the local clause in each language.

In addition, I will be excluding from my analysis and reconstruction any participle and copula/lexical verb collocations that are clearly compositional. That is to say, if in a given collocation the participle and the lexical verb retain their individual compositional semantics instead of clearly forming a single periphrastic construction, then that construction is not an auxiliary construction, and cannot be relied on to accurately describe the relationship between T and the lower verbal domain.

Finally, often in the ancient IE languages inflected auxiliaries (especially 'be' auxiliaries) will be omitted in auxiliary constructions, most often in conjoined clauses. In the majority of these tokens across the ancient languages the participle does occur clause-finally, as in (8) below, allowing for the possibility that the elided auxiliaries would be located clause-finally, but without definite proof of their location in the syntax, I will be ignoring all such constructions in my analysis.

(8) *(maiwe ne)sau   māwk   ñäś srūka(l)l(e)*
    young COP-1sg not-EMPH I   dying-GV.N/O
    "I a(m young), I will not die yet."   (Tocharian B, THT 1.b2)

3.3   Can we reconstruct auxiliary constructions for PIE?

Just as we saw with complementizers in §2, auxiliaries are ubiquitous in the early IE languages, either already present at their earliest attestation or innnovated during their early attested history, but their presence cannot be securely reconstructed for PIE since 1) the earliest constructions don't always use the same lexical verb for the auxiliary, and 2) the constructions themselves were often innovated within the attested history of the languages. We don't see the Sanskrit periphrastic perfect showing up, for example, until after the Rig Vedic period, initially constructed with√/kr̥ '*do*', and only much later with√/as '*be*' and√/bʰū '*become*'. Latin auxiliary constructions, however, initially use 'be' and later develop with the verb *habere* 'have'. The oldest periphrastic constructions in Greek show up in Homer, mostly with 'be', but there are also a couple of ambiguous examples of proto-auxiliary constructions with *ékʰo*: 'have', which become much more productive in later Greek. Hittite uses *ḫark-* 'have' and 'be' from the earliest attestation, but we suspect from their absence in the other Anatolian languages that the 'have' auxiliaries were innovated within Hittite.

Some scholars have reconstructed prehistoric periphrastic constructions from univerbated inherited verb forms (like the Latin imperfect as described by Weiss (2009: 414), or Balles (2008)'s reconstruction of the Old Indic *cvi* construction), but it is not known if these date back to PIE, or were innovated separately in the branches that possess them. Even more reminiscent of my methods here, Costello (1984) reconstructs periphrastic passive constructions for PIE based on the existence of etymologically unrelated periphrastic passive constructions in the IE daughter languages. This is crucially different, however, from the reconstruction proposed here (and in my opinion constitutes a misapplication of the Comparative Method): I do not believe that if daughter languages have a *construction*, their parent must have it as well regardless of etymology, but rather that if daughter languages show the same structural *features*, we must reconstruct those features for the parent as well regardless of etymology. This is in keeping with recent generative syntactic reconstruction literature (Walkden 2014, Roberts & Roussou 2003, etc.).

If we are eventually able to securely reconstruct these specific constructions for PIE, it would strengthen the case for my own reconstruction since every one of these potentially reconstructible constructions shows exactly the order of constituents I predict, with the univerbated verb forms and passive constructions showing Part-Aux word order, and showing the old IE languages' propensity for clause-final verb order.

So, again, we have independent innovation of the same syntactic construction across a family of related languages, which, as we saw with the innovated complementizers, can give us insight into the structural features of the syntactic heads these innovated constructions fill. This is especially the case if all of the earliest attested daughter languages agree in the syntax of their separately innovated auxiliary constructions.

3.4   Summary of the early IE auxiliary data

As we saw before with early IE complementizer data, once we abstract away from the individual lexical items filling syntactic positions, we see striking similarities across the early IE auxiliary constructions, as seen in examples (9) through (14) below.

(9) **Hittite**

[(našma)] ÉSAG   *kuiš*   ZI-*it*   **kīnu-an**
or   granary somebody by.his.will break-PRTC.NOM.SG
ḫar-z[(i)]
have-3SG

"Or somebody has broken open a granary by his own will"
                    (MH/MS (CTH 261.3) KUB 13.1(+) rev. iv 20'-23')

13

(10) **Tocharian B**

mā tot   *ñiś pintwāt*   ***warpalle***   ***nesau***
not so.much I   alms-N/O.SG accepting-GV COP.1SG.PRES

"I will not accept (any) alms"   (THT 107 b10)

(11) **Vedic Sanskrit**

ásūn   *pitŕ̥bʰyo*   ***gamayā́ṃ***   ***cakāra***
breaths father-DAT.PL going-VBL.NOUN do-PERF.3SG.ACT.IND

"He made his breaths go to the fathers"   (Atharvaveda 18.2.27)

(12) **Homeric Greek**

meːdʼ   *éti*   *Teːlemákʰoio*   *patèːr* ***kekleːménos***
and.not still Telemachus-GEN father called-PART.PERF.MED.NOM.SG.M
eíeːn
COP.1SG.PRES.OPT.ACT

"And may I no longer be called the father of Telemachus" (Iliad 2.260)

(13) **Old Latin**

sed quid tu foras   ***egressa***
but why you outside departed-PART.PERF.PASS.NOM.SG.F
es?
COP.2SG.PRES.ACT.IND

"But why have you come outside?"   (Plautus, Amphitryon 1078)

(14) **Gothic**

witandans þatei du sunjonai   *aiwaggeljons*
knowing that for defense-DAT.SG gospel-GEN.SG
gasatiþs   ***im***
set-PART.PAST.NOM.SG.M COP.1SG.PRES

"Knowing that for the defense of the gospels I have been set"
(Philippians 1:16, Katz 2019)

In all of the oldest Indo-European branches we see examples of auxiliary constructions in which the auxiliary occurs clause-finally, immediately preceded by the participle. In order to determine exactly what percentage of auxiliary tokens in each language show this behavior, and also to see what the exceptions to this word order generalization tell us about the clausal syntax of each of these early IE languages, I have collected auxiliary construction corpora for Tocharian B and Homeric Greek, and am in the process of collecting corpora for the other four.[9]

---

9 For the data from the corpora still in-progress, along with more detail about the classes of exceptions seen in Homeric and Tocharian, see my dissertation, Windhearn (2020).

3.5 Auxiliaries in Tocharian B

Tocharian is an extinct Indo-European branch spoken on the northern edge of the Tarim Basin in what is now the Xinjiang province in northwestern China. It is the eastmost ancient IE language, and had two dialects, called A and B. Our documents date from the 6th to 8th centuries CE. Tocharian is of the utmost importance to Indo-European reconstruction, as many Indo-Europeanists (e.g. Weiss 2018) believe that it was the second language to split off from the parent language, after the Anatolian languages. Tocharian can therefore be compared with the reconstructed ancestor of the other eight branches of Indo-European directly to reconstruct the second-oldest layer of PIE.

According to Adams (2015), Tocharian possesses periphrastic perfect, future, necessitive, and potential constructions consisting of a participle/gerund and 'be'. I gathered all tokens of periphrastic constructions from the translated portion of the Comprehensive Edition of Tocharian Manuscripts (CE-ToM). As we can see in Table 1, there are 88 total periphrastic constructions in my corpus. 65 of these (74%) place the auxiliary clause-finally immediately following the participle. Also, note that there are no examples of prose clauses ending any other way. This generalization closely matches the word order we would expect from a clause structure with a right-headed TP.

Period	Type	Part-Aux	Other
Archaic	Verse	2	0
Classical	Verse	30	20
Classical	Prose	21	0
Late	Verse	7	0
Late	Prose	4	0
Other	Both	1	3
Total		65	23

Table 1   Clause-final word order in Tocharian periphrastic constructions

3.5.1 Postposed exceptions to the Part-Aux generalization

There are four categories of exceptions to the word order generalizations described above. The first category shows the expected order of participle+auxiliary clause-finally, but with a single element postposed immediately following 'be'. There are five examples of postposed structural cases in the corpus, and seven examples of postposed oblique cases or adjuncts, comprising over half of the total exceptions and accounting for fourteen percent of the auxiliary constructions in the Tocharian corpus.

(15) *mentsisā*   *kr$_u$i* **wikalle**   **tako$_ī$**   *läkle*
      grief-PERL.SG if disappear-GV.NOM.SG COP.OPT.3SG pain-N/O.SG

      yesäñ   *sem*
      you-GEN.PL this-NOM.SG

      "If this sorrow of yours could be driven away by grief, [...]"

                                                            (THT 295 b8)

(16) *empele*   *rano* **yāmu**   **ṣey-ne**
      terrible-NOM.SG however done-PART.NOM.SG COP.IMF.3SG-OBJ

      yāmor   *su,*   *onmin*   *no postäṃ*
      deed-N/O.SG this-NOM.SG remorse-N/O.SG but afterwards

      yamaṣate   *mrau(skāte)*
      do-PRET.MID.3SG feel.disgust-PRET.MID.3SG

      "Even though such a terrible deed he had done, nevertheless he felt
      remorse afterwards [and] felt revulsion."   (PK AS 7C a6)

For (15) and (16) above, we have either an entire DP or a stranded piece of
a DP appearing clause-finally immediately following the auxiliary. If the
clause-final Part-Aux word order that we see in 75% of the Tocharian corpus
indeed reflects the base-generated word order, then these examples can be
derived however Right Dislocation is derived, either by rightward movement,
or through iterated movement, first of the postposed DP for focus reasons,
followed by remnant topicalization of the remainder of TP around it.[10]

(17) *tetrīku*   *po*   **trikau**   **nesau**
      confused-PART.NOM.SG all-N/O.SG led.astray-PART.NOM.SG COP.1SG

      lakle(sa)
      suffering-PERL.SG

      "I have gone astray, all confused I am through the suffering."

                                                            (THT 17 a2)

For the examples where the stranded element is an adverbial, if you believe
in right-adjunction, then this subclass of exceptions shows straightforward
right-T syntax with a higher right-adjoined adverbial. If you don't believe
in right-adjunction, or believe that this adjunct is too high in the syntax for
scope reasons, then this subclass can be derived through the same Right
Dislocation process that I discussed for (15) and (16).

---

10 For a detailed discussion of focus, topicalization, and the information structure implications
for these and other left periphery movements in one early IE language (Classical Greek), refer
to Goldstein (2015).

3.5.2   Intervening exceptions to the generalization

The second category of exceptions consists of examples with a single element appearing between the clause-final participle and 'be'. There are four examples in the corpus: two where the interveners are nominative, and two where they are genitive.

(18)   *keklyauṣ(wa)*   ***eṅ(ku***   *p)elaiknenta* *ṣaim,*
       heard-PART.NOM.PL seized-PART.NOM.SG law-NOM.PL COP.IMF.1SG

       po   *märsā(wa)*
       all-N/O.SG forget-PRET.1SG
       "The laws heard I had grasped [but now] all I have forgotten."

                                                                     (THT 15 a2)

These examples are all derivable through topicalizing vP instead of the entirety of TP after focusing the DP (or portion of DP, in this case), leaving a single element intervening between the participle and the auxiliary. In this example specifically, we have *eṅ(ku* 'grasped' being introduced as a contrastive topic (Kuno 1976, Frascarelli & Hinterhölzl 2007) to emphasize the contrast between laws once learned but now forgotten.[11]

   In the interesting example (19) below, we see this vP topicalization interacting with wh-movement. As demonstrated by Hale (1987), many old Indo-European languages have at least one landing site available above a moved wh-word, and it appears that Tocharian is no exception. So, in this example, we first see *ket* undergoing wh-movement, followed by the topicalization of vP to a location just above the wh-word in the A'domain, leaving the auxiliary stranded clause-finally.

(19)   *se(ṃ)*   *t(e)-yäknesa yāmor*   ***yāmu***   *ket*
       this-NOM.SG in.this.way deed-N/O.SG done-PART.NOM.SG who-GEN

       tākaṃ
       COP.SUBJ.3SG
       "By whom such a deed is done."   (PK AS 7C b3)

3.5.3   Initial Part-Aux exceptions to the generalization

Third are examples where the participle and 'be' appear clause-initially. There are four examples that fit this pattern in the corpus.

---

11 Thanks to an anonymous reviewer for pointing me toward relevant discussions of contrastive topicalization.

(20) **yaitu** **ṣai** *sū* *(krentauna)ṣṣeṃ*
adorned-PART.NOM.SG COP.IMF.3SG this-NOM.SG virtues-OBL.PL

tsaiññentsa
ornament-PERL.PL

"He had been adorned with the ornaments of the virtues."

(THT 77 a6)

This word order appears in a few shorter clauses, and requires focusing the subject 'he', followed by topicalization of TP, placing the participle and auxiliary clause-initially. The perlative DP is most likely right-adjoined above TP in this example.

3.5.4 Preposed auxiliary exceptions to the generalization

In the remaining examples 'be' appears before the participle at the end of the clause.

(21) *kestaṣṣe* *ce_u* *laklesa* *(prā)kre* **ṣeyeṃ**
of.hunger this suffering-PERL.SG hard-NOM.SG COP.IMF.3PL

memīyo(ṣ)
stricken-PART.NOM.PL

"By the pain of hunger, they have been terribly stricken."

(PK AS 17J a6)

We can derive these tokens either through Right Dislocation, or by focusing the participle, followed by topicalizing the rest of TP.

Thus, the lion's share of the Tocharian data points directly toward right-headedness within TP, and the various categories we see in the exceptions to clause-final Part-Aux word order are all derived either through Right Dislocation, or through focusing different single constituents before topicalizing the rest of either TP or vP, in a pattern that we will see repeated in the other old IE languages below.

3.6 Auxiliaries in Homeric Greek

Auxiliary constructions in Homeric Greek are limited to periphrastic perfects composed of a participle and either 'be' or 'have'. I looked at all examples of participles collocated with 'be' and 'have' in Homer, collected by Bentein (2016).

As shown in Table 3.6, there are 58 total participle+'be'/'have' collocations across the *Iliad*, *Odyssey*, and *Hymns*. 44 of these examples (76%) place the finite verb immediately following the participle clause-finally as seen in (13), reproduced as (22) below.

Work	Part-Aux	Part-Aux-X	Other	Total
Iliad	22	4	3	29
Odyssey	18	3	2	23
Hymns	4	0	2	6
Total	44	7	7	58

Table 2   Clause-final word order in Homeric participle-verb collocations

(22) *meːd' éti Teːlemákʰoio   patèːr **kekleːménos***
      and.not still Telemachus-GEN father called-PART.PERF.MED.NOM.SG.M
      eíeːn
      COP.1SG.PRES.OPT.ACT
      "And may I no longer be called the father of Telemachus" (Iliad 2.260)

Already we see close parallels between the Homeric data and the Tocharian data we just looked at, with Homeric pointing directly toward right-headedness within TP with the vast majority of its tokens as well.

3.6.1   Exceptions to the Part-Aux generalization

As we might expect, the exceptions to the Part-Aux generalization in Homeric closely resemble those seen in Tocharian as well, with the majority consisting of single postposed DPs or parts of DPs.

(23) *oùtis,   hòn   oú põː pʰeːmi*
      Noman-NOM.SG who-ACC.SG not yet say-1SG
      pepʰugménon   eĩnai   óletʰron
      fled-PART.PERF.MED.ACC.SG.M COP.INF.PRES.ACT ruin-ACC.SG
      "Even Noman, who, I tell thee, has not yet escaped destruction"
                                                    (Odyssey 9.455)

Here as well, these examples can be derived from a right-T structure by Right Dislocation, or by iterated focusing of the postposed DP followed by remnant topicalization of the remainder of TP around it.

   Most of the apparent counterexamples in Bentein's participle+'be'/'have' collocations are in fact compositional, with the auxiliary still showing copular behavior, and *ekʰoː* still showing possessive lexical semantics:

(24) *all' **anapeptaménas**   ék^hon*
      but spread-PART.PERF.MED.ACC.PL.F have-3PL.IMF.ACT.IND
      anéres
      man-NOM.PL
      "but men were holding them flung wide open"   (Iliad 12.122)

(24) is the sole example of *ek^ho:* in the corpus that does not show clause-final Part-Aux word order, and it foreshadows the ambiguity that will ultimately lead to the development of periphrastic perfects with *ek^ho:*. The semantics of this sentence are clear from context, however, and must describe the gates being held open, not having been opened, which tells us that this construction has the compositional reading, not the periphrastic perfect reading.

The single non-compositional non-postposed exception in the Homeric data is (25) below, from Homeric Hymn 3 dated significantly later than the epics, featuring an adverb immediately preceding the copula, and an adverb and oblique DP following. This example can also be derived through right-adjunction of the following adverbials and focusing of the vP, as we saw in the similar Tocharian examples.

(25) *kaì sè   **kekle:méne:**   émpe:s*
      and yours-NOM.SG.F called-PART.PERF.MED.NOM.SG.F nevertheless
      ẽ:a   hr' en   at^hanátoisin
      COP.1SG.IMF.ACT.IND then among immortal-DAT.PL
      "I, who was at least called your wife among the undying gods"
                                            (HH3 324–325)

For Homeric as well, then, we see that the vast majority of true auxiliary constructions point directly toward a right-headed TP, with the exceptions to the clause-final Part-Aux word order generalization straightforwardly derivable from this right-T structure.

3.7   A brief overview of the auxiliary literature for the other old IE languages

3.7.1   Auxiliaries in Hittite

Most syntactic analyses in the Anatolian literature either avoid the topic of headedness, or seemingly default to a head-initial analysis (e.g. Garrett 1994, Huggard 2011). Sideltsev (2014), however, specifically argues instead for right-headedness within TP and left-headedness above TP for Hittite. He bases this claim primarily on the "rigidity" of clause-final verbs and the rarity of postverbal subjects and objects, but more importantly he also notes the behavior of the auxiliaries *ḫark-* 'have' and *ēš-* 'be', which he claims always follow the participle clause-finally, as seen in (26) below:

(26)   *[(našma)] ÉSAG   kuiš   ZI-it*
        or   granary somebody.NOM.SG.C by.his.will
        kīnu-an   *ḥar-z[(i)]*
        break-PRTC.NOM.SG.N have-3SG.PRS
        "Or somebody **has** broken open a granary by his own will"
                                    (MH/MS (CTH 261.3) KUB 13.1(+) rev. iv 20′-23′)

Sideltsev concludes that the only reasonable syntactic structure that can account for these auxiliary word order facts is a left-headed CP and a right-headed TP, which nicely mirrors the Tocharian and Greek facts above.

3.7.2   Auxiliaries in Vedic Sanskrit

Schaufele (1991), one of the most complete analyses of Sanskrit word order, follows most of western scholarship in assuming base SOV word order for Sanskrit, and claims that the majority of phrases are head-final. Similarly, Hock (1984) notes that 97% of Vedic prose texts are verb-final, compared to 65% of poetic texts. The earliest auxiliary construction that appears in the Vedic corpus, and the only one that appears in Vedic poetry, does show clause final Part-Aux word order:

(27)   *ásūn   pitŕ̥bʰyo   **gamayā́ṃ**   **cakāra***
        breaths father-DAT.PL going-VBL.NOUN do-PERF.3SG.ACT.IND
        "he made his breaths go to the fathers"   (Atharvaveda 18.2.27)

Of the remainder of the auxiliary constructions in the Vedic corpus, we know from Hock that nearly all are Aux-final, since they all occur in Vedic prose. We do not yet have any analyses in the literature about the location of the participle in Vedic word order, however, and though these analyses are encouraging, they are not yet enough to conclusively prove that Vedic was right-headed in TP. I am currently gathering a corpus of these Vedic prose auxiliary constructions, which should shed more conclusive light on the question of Vedic headedness.

3.7.3   Auxiliaries in Latin

The most thorough works on phrasal headedness in Italic are Ledgeway (2012) and Danckaert (2012, 2017). Ledgeway describes in detail the gradual change from head-final to head-initial exhibited throughout Latin to the modern Romance languages. His conclusion, however, is that both TP and CP *emerged* over the (pre-)history of Latin and Romance. The CP argument originates in the idea that PIE lacked clausal embedding discussed earlier in the paper; again, see Probert (2014) for a compelling recent argument to the

contrary. Ledgeway himself uses the left periphery to account for much of Latin's free word order, which is mirrored by other early IE languages. Also note that we do see complementizers already in the earliest Latin data, and that when they appear, they show up heading a clause-initial CP.

For TP, the claim is that the development of TP corresponds to the rise of the left-headed auxiliary constructions seen in later Romance. But, clause-final auxiliary constructions with Part-Aux word order are already ubiquitous in Old Latin, both with 'be' and later with 'have', as seen in (28) below:

(28) *sed quid tu foras* **egressa**
      but why you outside departed-PART.PERF.PASS.NOM.SG.F
      es?
      COP.2SG.PRES.ACT.IND
      "But why have you come outside?"   (Plautus, Amphitryon 1078)

I would argue therefore that the major innovation from Latin to Romance was not the development of TP, but was more likely the switch of TP-headedness from clause-final to clause-initial.

Danckaert (2012) explores in great diachronic detail the syntax of the Latin left periphery in embedded clauses, and Danckaert (2017) analyzes the development of Latin clause structure in general. For our purposes, Danckaert (2012) notes the great diversity of topicalization, focus, and left-dislocation constructions in Latin, consistent with our Tocharian data, though in some ways even more productive, especially in the later Classical language. Danckaert (2017) notes that Plautus, who provides us with the largest early Latin corpus, shows OVAux word order 83% of the time, and that 60% of the Latin corpus between 200BCE and 200CE shows OVAux word order. This is both closely consistent with the numbers we saw for Tocharian and Homeric above, and potentially indicative of the later shift from right-headedness in TP to left-headedness explored by Ledgeway. Danckaert proposes a competing grammars analysis for the admittedly complex Classical Latin data, proposing that TP and VP are descriptively right-headed in some derivations and left-headed in others, with earlier Latin requiring more and more right-headedness in TP.

Though it may ultimately be the case that a competing grammars analysis is required to account for the complex word orders we see in Classical Latin (which could be indicative of a change in progress), Danckaert's Old Latin data seems to be significantly more amenable to the sort of right-headed TP analysis I propose for Tocharian and Homeric.

3.7.4 Auxiliaries in Germanic

Sapp (2016) presents a detailed argument for base SOV word order and head-final VPs in Old High German. He derives surface V2 word order

in Germanic through raising of the verb to C, following most analyses of Modern German, and mentions that his analysis is compatible with that of Lenerz (1984), who had earlier posited head-final TP structure for Old High German as well. And then, of course, there's Modern German itself, which many would consider the *Paradebeispiel* for left-headed CP/right-headed TP langauges, especially in embedded clauses. For our purposes here, the main syntactic innovation of the Germanic languages from PIE would be the innovation of V2 word order in matrix clauses through the development of obligatory T-to-C movement.

4   Reconstructing disharmonic clausal headedness for PIE

4.1   Setting up a correspondence set

As with Indo-European complementizers and the C domain, in lieu of a reconstructible auxiliary construction shared by the daughter languages, our correspondence set for the T domain must instead be composed of the syntactic features of the functional heads associated with the various auxiliary constructions innovated by each of the daughter languages. So, instead of trying to reconstruct both the phonological form *and* position of T, I reconstruct just the position itself, regardless of what phonological form this position takes in the daughter languages. Our correspondence set must therefore be composed of the feature sets of the innovated auxiliaries in each of the daughter languages, and should ignore the specific phonological forms of each auxiliary construction.

And, as we have seen both from the word order generalizations in each of the daughter languages and from the nature of the exceptions to clause-final V-Aux order in the Tocharian and Homeric Greek corpora, the ancient IE languages explored here point toward a right-headed T domain for the proto-language. In addition, most of these languages also show clausal embedding directly before the clause-final Part-Aux collocation, indicating that these languages are right-headed in their V domains as well. Our correspondence set therefore points toward Proto-Indo-European having been right-headed in both T and V. When combined with our left-headed reconstruction for CP from §2, we finally arrive at what we've been waiting for: a clause structure reconstruction for Proto-Indo-European that is left-C, right-T, right-V.

4.2   Alternatives to reconstructing mixed headedness for PIE

I have argued here that the disharmonic headedness data that we see in the Indo-European daughter languages was likely inherited from their parent, but there are, of course, other possible explanations. In this section I will show that these are unsatisfactory, and that the proposed reconstruction is
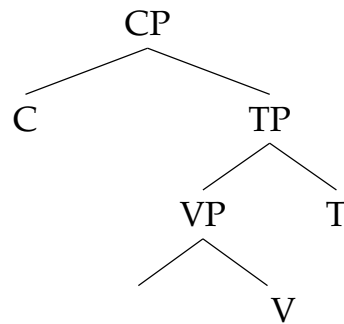
```
                          CP
                        /    \
                       C      TP
                            /    \
                          VP       T
                         /  \
                             V
```

Figure 1   Reconstructed Proto-Indo-European clause structure

the best possible explanation.

The first alternative is that disharmonic headedness was an areal feature that somehow spread throughout the Indo-European languages at a post-PIE date. The problem with this hypothesis is that the IE languages are so widely dispersed that this areal feature spread must have either occurred early enough as to be indistinguishable from PIE (in which case we should likely reconstruct it anyway in absence of data to the contrary), or that this feature spread occurred across an infeasibly broad geographic area.

A closely related alternate hypothesis is that disharmonic headedness was innovated in one (or more) Indo-European language, and later spread to the others through borrowing. Similar to the areal feature hypothesis, though, for this feature to have been borrowed into *all* of the earliest IE languages it would have to either travel unreasonably far, or happen early enough as to be indistinguishable from PIE.

A final alternative is that these auxiliary construction innovations somehow conspired to produce the disharmonic headedness that we see in each of the daughter languages completely independently of each other. In this scenario PIE had harmonic clausal syntax of some sort, most likely left-headedness in C, T, and V, and most/all of the daughter languages separately innovated disharmonic headedness after the breakup of the proto-language. However, the likelihood of all of the daughter languages agreeing this closely by chance is (to put it mildly) prohibitively low, and hypotheticals of this sort, without significant additional evidence, contradict the Comparative Method. For example, if all daughter languages in a given language family show /k/, it would be ludicrous to reconstruct /t/ for the parent language without a very good reason for doing so. Since all of the IE daughter languages show disharmonic headedness, in lieu of evidence to the contrary we must

reconstruct disharmonic headedness for PIE as well.[12]

One criticism of the syntactic reconstruction I undertake here that has been brought to my attention, which is closely related to the conspiracy criticism discussed above, is that similar reconstructions of modern languages that have a documented ancestor show that this sort of syntactic reconstruction gives the wrong results. For example, we know that all of the Romance languages are left-T, and that their shared ancestor Latin was (according to the analysis I propose here) right-T. Wouldn't my reconstruction of Proto-Romance contradict what I've said about Latin, and doesn't this botched reconstruction show that feature-based syntactic reconstruction is unreliable?

In short, no. On the one hand, in comparing the Romance languages we aren't trying to reconstruct Classical Latin; we're trying to reconstruct the latest shared ancestor of these Romance languages, Proto-Romance. On the other hand, Classical Latin fossilized hundreds of years before Latin developed into early Romance, and evidence suggests that late spoken Latin, and certainly early Romance, had already developed left-T clausal syntax. Even if pre-PIE at some point *was* left-headed in T, by the time it split into the separate daughter branches it was right-headed.

4.3   IE and the FOFC

The Final-over-final Constraint of Holmberg (2000) states that a right-headed projection may not dominate a left-headed one, and has been a hot topic in recent diachronic syntactic literature. If my analyses of the daughter languages and the reconstruction for the proto-language are correct, then the FOFC seems to be borne out by the IE data, as predicted by Biberauer et al. (2014). If PIE really was left-headed above TP and right-headed below TP, and the IE daughter languages all inherited the same syntactic structure, then at no time during the reconstructible history of the Indo-European languages (aside from the apparent VOAux word orders innovated in Classical Latin, treated in Danckaert 2017) did a right-headed projection dominate a left-headed one in the clausal syntax. This constraint is borne out both synchronically by the IE daughter language data collected here, and diachronically by our reconstruction of PIE's clausal syntax. For discussion on the theoretical

---

12 One reviewer asks if a cross-linguistically parallel grammaticalization path for the pronominally-derived complementizers could possibly explain a parallel syntactic development path of left-headedness in C. Aside from the likelihood that the interrogative- and relative-derived complementizers had a different origin than the declarative-derived complementizer (as discussed in Fn. 5), we would still have to reconstruct C-left syntax, since good practice dictates that we reconstruct the simplest possible explanation for a set of data. Otherwise, every time we see a cognate /h/ across a family of related languages, we might be tempted to reconstruct */s/ for their shared ancestor, since */s/ > /h/ is a common development, regardless of the fact that /h/ can have multiple possible sources, not least of which is */h/!

details and implications of deriving various O, V, and Aux word orders in LCA-based models or in the PF-interface approach taken in this paper, see Biberauer et al. (2014).

4.4 Conclusions

Better syntactic corpora can only improve our PIE clause structure reconstruction, and to that end I'm finalizing auxiliary corpus data for the other four languages discussed here, Latin, Hittite, Sanskrit, and Gothic, to further solidify the case for PIE disharmonic headedness. Thus far, it appears that clause-final Part-Aux word order for Hittite and Sanskrit is completely exceptionless. Latin shows behavior similar to Greek and Tocharian, with clause-final Part-Aux order the majority of the time, but also making use of the focusing, topicalization, and stranding mechanisms we see in its sisters. Gothic, in the few examples where it diverges from the syntax of the Koine Greek original, appears to show the expected Germanic V2 auxiliary word order in matrix clauses, and clause-final Part-Aux order in embedded clauses.

Overall, this paper builds upon recent syntactic reconstruction literature both by suggesting that the featural composition of syntactic functional heads can be reconstructed without an accompanying reconstructible phonological form, and by showing that certain types of innovation can inform reconstruction, at least in the syntactic domain. Using this innovation-driven reconstruction, I proposed that Proto-Indo-European was disharmonically headed in CP and TP, based on innovated complementizer and auxiliary construction data from six of the earliest-attested Indo-European daughter languages.

References

Adams, Douglas Q. 2015. *Tocharian B: a grammar of syntax and word-formation*. Innsbruck: Inst. für Sprachen und Literaturen der Univ. Innsbruck, Bereich Sprachwiss.

Baker, Mark C. 2008. The macroparameter in a microparametric world. In Theresa Biberauer (ed.), *The limits of syntactic variation*, 351–373. Amsterdam: John Benjamins.

Balles, Irene. 2008. Principles of syntactic reconstruction and "morphology as paleosyntax". In Ferraresi & Goldbach (2008) 161–186.

Bentein, Klaas. 2016. *Verbal periphrasis in Ancient Greek:* have- *and* be- *constructions*. Oxford: Oxford University Press.

Biberauer, Theresa, Anders Holmberg & Ian Roberts. 2014. A syntactic universal and its consequences. *Linguistic Inquiry* 45(2). 169–225.

CET. n.d. A Comprehensive Edition of Tocharian Manuscripts. Retrieved: April 11, 2017. http://www.univie.ac.at/tocharian/.

Chomsky, Noam. 1982. *Some concepts and consequences of the theory of Government and Binding*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Costello, John R. 1984. The periphrastic passive construction in Proto-Indo-European. *Word* 35(2). 125–161.

Danckaert, Lieven. 2012. *Latin embedded clauses: the left periphery*. Amsterdam: John Benjamins.

Danckaert, Lieven. 2017. *The development of Latin clause structure: A study of the extended verb phrase*. Oxford: Oxford University Press.

Delbrück, Berthold. 1900. *Vergleichende Syntax der indogermanischen Sprachen*, vol. 3. Strasbourg: Trübner.

Ferraresi, Gisella & Maria Goldbach (eds.). 2008. *Principles of syntactic reconstruction*. Amsterdam: John Benjamins.

Fox, Danny & David Pesetsky. 2005. Cyclic linearization of syntactic structure. *Theoretical Linguistics* 31. 1–45.

Frascarelli, Mara & Roland Hinterhölzl. 2007. Types of topic in German and Italian. In Kerstin Schwabe & Susanne Winkler (eds.), *On information structure, meaning and form: generalizations across languages*, 87–116. Amsterdam: John Benjamins.

Garrett, Andrew. 1994. Relative clause syntax in Lycian and Hittite. *Die Sprache* 36(1). 29–69.

Goldstein, David M. 2015. *Classical Greek syntax: Wackernagel's law in Herodotus*. Leiden: Brill.

Hackstein, Olav. 2013. The evolution of finite complementation in Tocharian. *Tocharian and Indo-European Studies* 13. 117–148.

Hale, Mark. 1987. *Studies in the comparative syntax of the oldest Indo-Iranian languages*: Harvard University dissertation.

Hale, Mark. 1996. Deriving Wackernagel's law: prosodic and syntactic factors determining clitic placement in the language of the Rigveda. In Aaron L. Halpern & Arnold Zwicky (eds.), *Approaching second: second position clitics and related phenomena*, 165–197. Chicago, IL: University of Chicago Press.

Hale, Mark. 1998. Diachronic syntax. *Syntax* 1. 1–18.

Hock, Hans Henrich. 1982. The Sanskrit quotative: A historical and comparative study in papers on diachronic syntax: Six case studies. *Studies in the Linguistic Sciences* 12. 39–85.

Hock, Hans Henrich. 1984. Rig-Vedic convergence of Indo-Aryan with Dravidian? Another look at the evidence. *Studies in the Linguistic Sciences* 14. 89–107.

Holmberg, Anders. 2000. Deriving OV order in Finnish. In Peter Svenonius (ed.), *The derivation of VO and OV*, 123–152. Amsterdam: John Benjamins.

Huggard, Mattyas. 2011. On Wh-(non)-movement and internal structures of the Hittite preposed relative clause. In *Proceedings of the 22nd annual*

UCLA Indo-European Conference, 104–126. Bremen: Hempen.

Katz, Moses. 2019. The resultative and the perfect in Gothic: resultativity as a model for non-agentive verbs in Early Germanic. Ms., University of Georgia.

Kiparsky, Paul. 1995. Indo-European origins of Germanic syntax. In Adrian Battye & Ian Roberts (eds.), *Clause structure and language change*, 140–169. Oxford: Oxford University Press.

Koller, Bernhard. 2013. On the status of the particle *ne* in Tocharian A clause structure. Paper presented at the 25th Annual West Coast Indo-European Conference, October 25–26.

Kuno, Susumu. 1976. Subject, theme, and the speaker's empathy – a reexamination of relativization phenomena. In Charles N. Li (ed.), *Subject and topic*, 417–444. New York, NY: Academic Press.

Ledgeway, Adam. 2012. *From Latin to Romance: morphosyntactic typology and change*. Oxford: Oxford University Press.

Lehmann, Christian. 2005. On the methodological bases of genetic language comparison. *Language Research* 41. 379–404.

Lenerz, Jürgen. 1984. *Syntaktischer Wandel und Grammatiktheorie: Eine Untersuchung an Beispielen aus der Sprachgeschichte des Deutschen*. Berlin: Walter de Gruyter.

Lightfoot, David W. 2002. Myths and the prehistory of grammars. *Journal of Linguistics* 38. 113–136.

Longobardi, Giuseppe. 2003. Methods in parametric linguistics and cognitive history. *Linguistic Variation Yearbook* 3. 101–138.

Pires, Acrisio & Sarah G. Thomason. 2008. How much syntactic reconstruction is possible? In Ferraresi & Goldbach (2008) 161–186.

Probert, Philomen. 2014. Relative clauses, Indo-Hittite and Standard Average European. In *Proceedings of the 25th annual UCLA Indo-European Conference*, 137–164.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar*, 281–337. Cambridge: Cambridge University Press.

Roberts, Ian & Anna Roussou. 2003. *A Minimalist approach to grammaticalization*. Cambridge: Cambridge University Press.

Sapp, Christopher. 2016. Word order patterns in the Old High German right periphery and their Indo-European origins. *Diachronica* 33. 367–411.

Saxena, Anju. 1995. Unidirectional grammaticalization: diachronic and cross-linguistic evidence. *STUF – Language Typology and Universals* 48. 350–372.

Scharf, Peter M. 2015. Interrogatives and word-order in Sanskrit. In Peter M. Scharf (ed.), *Sanskrit syntax*, 203–218. The Sanskrit Library.

Schaufele, Steven William. 1991. *Free word-order syntax: The challenge from Vedic Sanskrit to contemporary formal syntactic theory*: University of Illinois

at Urbana-Champaign dissertation.

Sideltsev, Andrej. 2014. Wh-in-situ in Hittite. In *Proceedings of the International Conference on the Typology of Morphosyntactic Parameters*, 199–222.

Walkden, George. 2014. *Syntactic reconstruction and Proto-Germanic*. Oxford: Oxford University Press.

Watkins, Calvert. 1976. Towards Proto-Indo-European syntax: problems and pseudo-problems. In Sanford Steever, Carol Walker & Salikoko Mufwene (eds.), *Papers from the parasession on diachronic syntax*, 306–326. Chicago, IL: Chicago Linguistic Society.

Weiss, Michael. 2009. *Outline of the historical and comparative grammar of Latin*. Ann Arbor/New York: Beech Stave Press.

Weiss, Michael. 2018. Tocharian and the West. In Olav Hackstein & Andreas Opfermann (eds.), *Priscis Libentius et Liberius Novis. Indogermanische und sprachwissenschaftliche Studien: Festschrift für Gerhard Meiser zum 65. Geburtstag*, 373–381. Hamburg: Baar.

Whitman, John. 2008. The classification of constituent order generalizations and diachronic explanation. In Jeff Good (ed.), *Language universals and language change*, 233–252. Oxford: Oxford University Press.

Windhearn, Ryan. 2020. *Evidence from innovation: Reconstructing disharmonic headedness for Proto-Indo-European*: Cornell University dissertation.

Ryan Windhearn
Oberbilker Allee 92
40227 Düsseldorf
Germany
rdh238@cornell.edu