

# Innovative reconstruction: disharmonic headedness in early Indo-European

Ryan Hearn

Cornell University

*rdh238@cornell.edu*

June 19, 2018

DiGS 20

# Main goals

1. Define “parallel syntactic innovation” for the purposes of this paper
  - ▶ Establish parallel syntactic innovation as a reliable basis for syntactic reconstruction
2. Examine complementizer development across the early Indo-European (IE) languages
  - ▶ Reconstruct a left-headed CP domain for Proto-Indo-European (PIE)
3. Examine auxiliary construction development across the early IE languages, especially Tocharian and ancient Greek
  - ▶ Reconstruct a right-headed TP domain for PIE

# Syntactic Reconstruction

- ▶ Historically, syntactic reconstruction worked very similarly to phonological/morphological reconstruction
  - ▶ Set up correspondence sets of either exact phrases, or of general word order
- ▶ Now, generative syntax greatly constrains the possible structures we generate
  - ▶ This allows us to go a step further, and reconstruct the structure underlying our word orders
  - ▶ So, not only can we reconstruct the phonological form of function words, but also the features of their functional heads

## But, what if a given function word is not reconstructible for the proto-language?

- ▶ What if the words that fill a functional role aren't cognate
- ▶ Or, what if each daughter language cognate developed this usage separately during each language's attested history?

## A relatively straightforward example: the IE complementizer

- ▶ Complementizers in early IE are not all cognate:
  - ▶ Latin *quod* and Tocharian *k<sub>u</sub>ce/kucne* from *\*k<sup>w</sup>o-* (the PIE interrogative stem)
  - ▶ Gothic *patei* from *\*to-* (a PIE pronominal stem)
  - ▶ Sanskrit *yád* and Greek ὅτι and ὡς, from *\*H<sub>2</sub>iō-* (the PIE relative stem)
- ▶ More importantly, as shown by Hackstein (2013), complementizer behavior mostly developed within the attested history of these languages
  - ▶ 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<sub>u</sub>ce/kucne* start as adverbial adjuncts which later develop complementizer usage.

# An overt complementizer cannot be reconstructed for PIE

- ▶ As shown by Hackstein, **zero-embedding** is likely the only reconstructible method for embedding sentential complements after verbs of utterance and cognition for PIE.
- ▶ So, assuming complementizers fill a functional head (let's call it C), without any reconstructible complementizer, we have no way of locating C in the syntax, right?
  - ▶ The C domain in PIE could be either left-headed or right-headed
  - ▶ But without a reconstructible complementizer in C, how can we know which?

## Setting up a correspondence set

- ▶ Even if PIE used zero-embedding, it still has a functional head C
  - ▶ C in PIE is just filled with a null complementizer
- ▶ So, instead of trying to reconstruct both the phonological form *and* position of C, we reconstruct just the position itself, regardless of what phonological form this position takes in the daughter languages
- ▶ We therefore set up a correspondence set for the underlying syntactic structure, and ignore the specific complementizers used

## Parallel complementizer innovation

- ▶ When we ignore the form of the complementizers, we immediately see striking similarities in the daughter languages' syntax
  - ▶ Every single innovated complementizer, when it ultimately shows up, appears clause-initially
- ▶ These languages aren't all independently innovating a left-headed C domain
  - ▶ They're innovating a phonological form to fill the left-headed C domain that they already share
- ▶ So, we see that our correspondence set unilaterally points toward a null clause-initial C for the proto-language



## Not necessarily structural innovation

- ▶ Note that this is not structural syntactic innovation
  - ▶ The underlying syntax of CP hasn't changed – just which lexical items fill which node
- ▶ The parallel innovation of separate phonological forms to fill the C node, i.e. *complementizers*, cues us in to the shared structural syntactic reality
- ▶ So, **parallel syntactic innovation** refers not necessarily to parallel structural innovations, but in this case to parallel innovation of new lexical items that reveal the underlying similarities in structure.

## Corroborating evidence for left-headed CP in PIE

1. Reanalysis, as defined by Langacker (1977), is “change in the structure of an expression or class of expressions that does not involve any immediate or intrinsic modification of its surface manifestation”
  - ▶ The examples of complementizer development discussed above are spec-head reanalyses
  - ▶ The reanalyzed elements must be surface-adjacent to the null complementizer
    - ▶ The null complementizer must be left-headed to get the correct surface order
  - ▶ This also ties in nicely with the Complementizer Attraction Universal of Bresnan (1972), which states that the landing site of a Comp attraction transformation (i.e. wh-movement) must be adjacent to C
    - ▶ If we reconstruct clause-initial wh-movement for PIE as most do, then we must reconstruct a left-headed CP as well

## Corroborating evidence for left-headed CP in PIE

2. Grammatical particles in Wackernagel positions 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, which may then undergo “prosodic flipping” with adjacent syntactic elements
  - ▶ Further, Scharf (2015) points out that the Sanskrit question particle *api* occurs clause-initially, instead of the clause final position we would expect if CP was right-headed (e.g. *ka* in Japanese)

## Now that we have a hammer, let's find some nails

- ▶ Up next is an age-old problem with significantly more disagreement in the literature
  - ▶ Reconstructing a left-headed CP for PIE is pretty well supported by the other arguments
  - ▶ Now I will show that the method we used to add support to that position is just as useful for reconstruction elsewhere

# Word Order in Indo-European

- ▶ Delbrück (1893) was the first to reconstruct clausal word order for PIE
  - ▶ He concluded that PIE must have been SOV based mainly on Sanskrit word-order evidence
- ▶ Sapp (2016) and Krisch (2017) both reconstruct head-finality within the VP domain
  - ▶ Again, this is mostly due to general SOV word order across the early IE languages.
- ▶ There are plenty of ways to derive SOV word order without needing VP head-finality, however.
- ▶ Let's see if we can some more evidence to better triangulate the exact location of these verbal elements

## Early IE Auxiliaries

- ▶ Just as we saw with complementizers, auxiliaries are ubiquitous in the early IE languages, but their presence cannot be securely reconstructed for PIE
  - ▶ We don't see the Sanskrit periphrastic perfect showing up until the Atharvaveda's *gamayām cakāra* 'he went', constructed with the *do* verb
  - ▶ Latin auxiliary constructions, however, initially use the copula and only later develop with the verb *habere* 'have'
  - ▶ The oldest periphrastic constructions in Greek show up in Homer, mostly with the copula
    - ▶ There are also a couple of ambiguous examples of proto-auxiliary constructions with εἶχον 'have'
  - ▶ Hittite uses *ḫark-* 'have' and the copula

## Parallel innovation in IE auxiliaries?

- ▶ Some have reconstructed prehistoric periphrastic constructions (like the Latin imperfect as described by Weiss 2009), but it is not known if these date back to PIE
- ▶ We may not be able to securely reconstruct a single auxiliary construction for PIE
  - ▶ But, as with complementizers in early IE languages, separate innovation of auxiliary constructions in the daughter IE languages can give us insight into their inherited syntax
  - ▶ This is especially the case if all of the earliest attested daughter languages agree in the syntax of their separately innovated auxiliary constructions
- ▶ So, let's take a look and see just how similar the early IE auxiliary constructions are

## Auxiliary constructions in Tocharian B

- ▶ Tocharian possesses periphrastic perfect, future, necessitive, and potential constructions consisting of a participle/gerund and an inflected copula.
- ▶ I gathered all examples of these periphrastic constructions from the translated portion of the Comprehensive Edition of Tocharian Manuscripts (CEToM).
  - ▶ I also gathered a few additional examples from Adams (2015).



## Auxiliary constructions in Tocharian B

| 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     |       | 1        | 3     |
| Total     |       | 65       | 23    |

- ▶ Tocharian B overwhelmingly (74%) prefers to end auxiliary clauses with a participle followed immediately by the inflected copula
- ▶ In the entire corpus there are no examples of prose sentences ending any other way

## Auxiliary constructions in Tocharian B

- ▶ Per Adams (2015), Tocharian B's "neutral" word order is SOV, comparable to the word order reconstructed for PIE
- ▶ This auxiliary data doesn't fit very well with the head-final VP analysis of SOV word order in PIE, however
  - ▶ Assuming the inflected copula would be sitting in a left-headed T above the right-headed VP
  - ▶ Most of the periphrastic sentences in the corpus would have to move everything into the left periphery
- ▶ I think it more likely that Tocharian B is **right-headed within its TP domain**.
  - ▶ This accounts for the auxiliary order within Tocharian without resorting to the left periphery for the majority of sentences

## Corroborating evidence: Negation in Tocharian

- ▶ According to Adams (2015), *mā* is the most common clausal negator and prohibitive, by itself accounting for 87% of all negated sentences.
  - ▶ *mā* may occur either clause-initially or immediately before the inflected verb much lower in the clause.
- ▶ I was able to find one instance of *mā* collocated with a verbal auxiliary complex:

(1) *teṃ yiknesa weweñu mā tākaṃ*  
 this way spoken not be.3SG.SUBJ

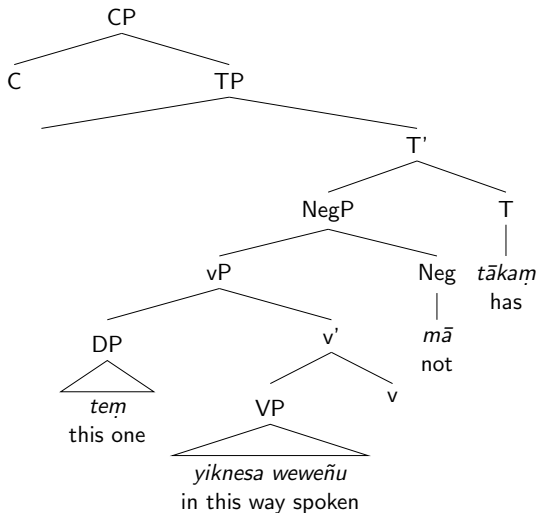
“(If) he has not spoken in this way” (331b3/4<sup>L</sup>, Adams)

## Corroborating evidence: Negation in Tocharian

- ▶ Note how the negation appears precisely between the participle and the copula.
  - ▶ With our posited right-headed TP domain, we would expect a right-headed NegP located between the TP and vP layers.
  - ▶ And, in the one example we have, that's exactly where we find it.

## Corroborating evidence: Negation in Tocharian

(2)



## Auxiliary behavior in early Greek

- ▶ I looked at the examples of auxiliary usage in Homer, collected by Bentein (2016). All were periphrastic perfects.

| <b>Work</b>    | <b>Part-Aux</b> | <b>Part-Aux-NP</b> | <b>Other</b> | <b>Total</b> |
|----------------|-----------------|--------------------|--------------|--------------|
| <i>Iliad</i>   | 22              | 4                  | 3            | 29           |
| <i>Odyssey</i> | 18              | 3                  | 2            | 23           |
| Hymns          | 4               | 0                  | 2            | 6            |
| <b>Total</b>   | <b>44</b>       | <b>7</b>           | <b>7</b>     | <b>58</b>    |

- ▶ 44 of the 58 Homeric examples (76%) place the auxiliary immediately following the participle clause-finally
- ▶ An additional 7 place the auxiliary immediately following the participle clause-finally, except for a single postposed NP or piece of an NP

## Auxiliary behavior in early Greek

- ▶ We see that the Greek auxiliary facts closely resemble the Tocharian B data just discussed
  - ▶ SOV, but with the inflected copula following the participle clause-finally
- ▶ Here too, we see a strong case for right-headedness in TP to best account for this word order, specifically due to the relationship between the participle and inflected copula
- ▶ If I can find a solution (perhaps phonological?) to the postposed NPs and split NPs, the case for right-headedness in TP in the earliest Greek would be even stronger

## Auxiliary behavior in later Greek

- ▶ It is worth noting that, as pointed out by Ceglia (1998), by the time of Herodotus the participle generally follows the copula in the sentence
- ▶ This mirrors the observation of Taylor (1994) that Homeric is primarily OV, with the younger Greek dialects developing more frequent VO word order.
- ▶ I am suspicious that these facts constitute a shift in TP headedness
  - ▶ I think this shift in headedness will be central to any eventual syntactic solution to the “unique degree of word order variation” seen in Classical Greek, as addressed in Goldstein (2015)



## Auxiliaries and word order in the other old IE languages

- ▶ With Greek and Tocharian showing such striking similarities in auxiliary behavior, let's turn to the existing literature on the other old IE languages, and see if we can find some more.

# 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) 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

# Hittite

- ▶ Most importantly, he also notes the behavior of the auxiliaries *ḫark-* ‘have’ and *ēš-* ‘be’, which always follow the participle, as seen below:

(3) [(*našma*)] ÉSAG *kuiš* ZI-*it*  
 or granary somebody.NOM.SG.C by.his.will  
*kīnu-an* *ḫar-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’)

# Anatolian

- ▶ Sideltsev also 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
- ▶ This nicely mirrors what we saw from Tocharian and Greek earlier
- ▶ More work remains to be done on the other Anatolian languages to determine the extent to which their auxiliary facts reflect those of Hittite

# Sanskrit

- ▶ Once periphrastic constructions like the aforementioned *gamayām cakāra* from the Atharvaveda start showing up, they usually occur clause-finally with the auxiliary following the verbal element, mirroring the behavior of the other old IE languages
  - ▶ Schaufele (1991), one of the most complete analyses of Sanskrit word order, follows most of western scholarship in assuming base SOV word order, 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
- ▶ For our purposes, these tendencies are telling but not yet conclusive. I plan to do a more in-depth analysis of periphrasis in Vedic soon

# Latin I

- ▶ The most thorough work on phrasal headedness in Italic is undoubtedly Ledgeway (2012)
  - ▶ He describes in detail the gradual change from head-final to head-initial exhibited throughout Latin to the modern Romance languages
- ▶ The argument seems to be 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; see Probert (2014) for evidence to the contrary.
  - ▶ This argument also seems odd since Ledgeway uses the left periphery to account for much of Latin's free word order, which is mirrored by other IE languages
  - ▶ Also note that we do see complementizers already in the Latin data, and that when they appear, they show up clause-initially

# Latin II

- ▶ For TP, the claim is that the development of TP corresponds to the rise of the left-headed auxiliary constructions in later Romance

- ▶ But, clause-final auxiliary constructions are already ubiquitous in Latin itself, both with the copula and later with *habere*

(4) cum cognitum habeas [...]
   
when known you.have [...]

“When you realize [...]” (Cic. *Fin.* 4.11, Ledgeway (2012))

- ▶ I would argue 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

# 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.
  - ▶ He mentions that his analysis is compatible with that of Lenerz (1984), who had earlier posited head-final TP structure for OHG.
  - ▶ Weiß (2007), on the other hand, argues for head-initial TP, and maintains that surface V2 word order is derived through movement of the finite verb into T itself.



## Germanic II

- ▶ And then, of course, there's Modern German, which many would consider the *Paradebeispiel* for left-headed CP/right-headed TP languages, especially in embedded clauses.
- ▶ For our purposes, the main syntactic innovation of Germanic from PIE would be V2 word order through obligatory T-to-C movement.

## Wrapping up

- ▶ We've seen striking similarities in auxiliary behavior across the early IE languages
  - ▶ Once the daughter languages develop auxiliaries, they overwhelmingly come clause-finally, usually immediately after the participle
- ▶ The daughter languages all point toward the same synchronic structural relationship between inflection (T) and the rest of the verbal domain
- ▶ The syntactic parameter that fits this word order best is right-headedness within TP

## Wrapping up

- ▶ With our structural correspondence set in agreement, we can reconstruct this structural relationship for the parent language as well
  - ▶ These parallel innovations of clause-final auxiliaries show us the relationship between the head-final TP and VP inherited from PIE.
  - ▶ The alternative is that each of these separate auxiliary innovations conspired to produce the same clause-final word order in each of these daughter languages independently

## Wrapping up

- ▶ This idea of parallel syntactic innovation revealing underlying syntactic similarities inherited from the parent language constitutes a new tool available for syntactic reconstruction
  - ▶ It provides a new argument not only in favor of reconstructing SOV word order for PIE, but of reconstructing a specific corresponding underlying structure
  - ▶ Combined with the complementizer data discussed earlier, it provides evidence for reconstructing a left-headed CP and right-headed TP for PIE
  - ▶ This structure was then inherited and made explicit by the earliest IE daughter languages

## Other takeaways: IE and the FOFC

- ▶ If PIE really was left-headed above TP and right-headed below TP, and the IE daughter languages inherited the same syntactic structure
  - ▶ Then at no time during IE's reconstructible history did a right-headed projection dominate a left-headed one
  - ▶ The Final-over-final Constraint (Holmberg, 2000), which states that a right-headed projection may not dominate a left-headed one, seems to be borne out by the IE data, as predicted by Biberauer *et al.* (2014)
  - ▶ Both synchronically by the early IE data, and diachronically by reconstruction

## Where to go from here

- ▶ I plan to gather each auxiliary example from the Vedic poetic and prose texts to see what structural insights the word order variations can give us
- ▶ I'm very interested in the word salad that is post-Homeric ancient Greek.
  - ▶ I will see if treating Greek as a language in the process of switching TP headedness from right to left gives us any insight into its striking word order variations.

Thank you for your attention!

# References I

*A Comprehensive Edition of Tocharian Manuscripts.*

<http://www.univie.ac.at/tocharian/>. Retrieved: April 11, 2017.

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

Bentein, Klaas. 2016. *Verbal Periphrasis in Ancient Greek: Have- and Be-Constructions*. Oxford University Press.

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

Bresnan, Joan W. 1972. *Theory of complementation in English syntax*. Ph.D. thesis, Massachusetts Institute of Technology.

Ceglia, Luca. 1998. L'evoluzione della costruzione perifrastica verbale nel greco del Nuovo Testamento. *Archivio Glottologico Italiano*, **83**, 20–44.

Delbrück, Berthold. 1893. *Vergleichende Syntax der indogermanischen Sprachen*. **3**.



## References II

- Garrett, Andrew. 1994. Relative clause syntax in Lycian and Hittite. *Die Sprache*, **36**(1), 29–69.
- Goldstein, David. 2015. *Classical Greek Syntax: Wackernagel's Law in Herodotus*. Brill.
- Hackstein, Olav. 2013. The evolution of finite complementation in Tocharian. *Tocharian and Indo-European Studies vol. 13*, **13**, 117–148.
- Hale, Mark. 1996. Deriving Wackernagel's Law: Prosodic and syntactic factors determining clitic placement in the language of the Rigveda. *Approaching second: Second position clitics and related phenomena*, 165–197.
- Hock, Hans Henrich. 1984. Rig-Vedic convergence of Indo-Aryan with Dravidian? Another look at the evidence. *Studies in the linguistic sciences*, **14**(1), 89–107.
- Holmberg, Anders. 2000. Deriving OV order in Finnish. *The Derivation of VO and OV*, 123–152.

## References III

- Huggard, Mattyas. 2011. On Wh-(non)-movement and internal structures of the Hittite preposed relative clause. *Pages 104–126 of: Proceedings of the 22nd Annual UCLA Indo-European Conference*. Bremen: Hempen, vol. 83.
- Koller, Bernhard. 2013. On the Status of the Particle ne in Tocharian A Clause Structure. *Pages 25–26 of: West Coast Indo-European Conference, UCLA, October*.
- Krisch, Thomas. 2017. Proto-Indo-European syntax. *Pages 111–152 of: Kapović, Mate, Ramat, Anna Giacalone, & Ramat, Paolo (eds), The Indo-European Languages*. Taylor & Francis.
- Langacker, Ronald W. 1977. Syntactic reanalysis. *Mechanisms of syntactic change*, 58.
- Ledgeway, Adam. 2012. *From Latin to Romance: Morphosyntactic typology and change*. Vol. 1. Oxford University Press.
- Lenerz, Jürgen. 1984. *Syntaktischer Wandel und Grammatiktheorie: Eine Untersuchung an Beispielen aus der Sprachgeschichte des Deutschen*. Vol. 141. Walter de Gruyter.

## References IV

- Probert, Ph. 2014. Relative clauses, Indo-Hittite and Standard Average European. *Pages 137–64 of: Proceedings of the 25th Annual UCLA Indo-European Conference.*
- Sapp, Christopher. 2016. Word order patterns in the Old High German right periphery and their Indo-European origins. *Diachronica*, **33**(3), 367–411.
- Scharf, Peter M. 2015. Interrogatives and Word-order in Sanskrit. *Pages 203–218 of: Scharf, Peter M. (ed), Sanskrit Syntax.* The Sanskrit Library.
- Schaufele, Steven William. 1991. *Free word-order syntax: The challenge from Vedic Sanskrit to contemporary formal syntactic theory.* Ph.D. thesis, University of Illinois at Urbana-Champaign.
- Sideltsev, Andrej. 2014. Wh-in-situ in Hittite. *Pages 199–222 of: Typology of Morphosyntactic Parameters. Proceedings of the International Conference Typology of Morphosyntactic Parameters.*
- Taylor, Ann. 1994. The change from SOV to SVO in Ancient Greek. *Language variation and change*, **6**(1), 1–37.

# References V

Weiß, Helmut. 2007. Manuscript. Die rechte Satzperipherie im Althochdeutschen: Zur Verbstellung in *dass*-Sätzen.

Weiss, Michael. 2009. Outline of the Historical and Comparative Grammar of Latin. *Ann Arbor–New York: Beech Stave Press.*