Headedness in Tocharian, and its implications for PIE reconstruction*

Ryan Hearn, Cornell University
November 3, 2017

1 Introduction

- The study of constituent/phrase structure in the Indo-European languages is both a very old science and a very young one.
  - Age-old syntactic phenomena have only relatively recently seen rigorous theoretical analysis.
  - This has given us more detailed knowledge of the synchronic syntax of these ancient languages, as well as improved our syntactic reconstructions of the proto-language.
- The current paper is an attempt to continue this trend of underpinning robust word-order generalizations in the IE languages with syntactic theory.
  - Today, I will examine the evidence in favor of left-headedness within CP and right-headedness within TP for the Tocharian languages.
  - Further, I will briefly summarize the evidence for similar analyses of the other old IE languages.
  - Finally, I will discuss the implications of these similarities for the reconstruction of PIE.

2 Mixed Headedness in Tocharian

2.1 Left-headedness in CP

- Left-headedness within CP across the Indo-European languages is effectively the default stance taken in the literature.
  - However, let's review support for this stance, in order to make this claim overtly.
- Grammatical particles showing up in Wackernagel positions show behavior indicating that they likely head their own projections in the left periphery.
  - For example, 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.
- In addition, Hackstein (2013) compares the development of complementizer behavior from relative constructions across the ancient IE languages.
  - The most common means of marking complement clauses in Tocharian is with a null complementizer, which unfortunately does little to show us its syntactic location.
  - However, Tocharian $k_u ce$, *ce* (B) and *kucne* (A), though only showing true complementizer behavior in a minority of complement-taking predicates, do occur clause-initially when used in this fashion.

*Thanks to my committee, Michael Weiss, John Whitman, and Miloje Despic, for their invaluable advice.

- Further, according to the Complementizer Attraction Universal of Bresnan (1972), the landing site of a Comp attraction transformation (i.e. wh-movement) must be adjacent to C.
  - So, since wh-movement in Tocharian (and across the Indo-European languages) is to a position near the left edge of the clause, CP must be left-headed in these languages.
- When taken together, these data clearly point toward left-headedness within CP in Tocharian.

2.2 Right-headedness in TP

- Adams (2015) claims that "neutral" word order in Tocharian B is SOV, based on the majority of clauses in the language showing this surface ordering of constituents.
- This is a good start when looking for right-headedness in general, but is far from conclusive, especially given the left-headed nature of CP in Tocharian. Let's see what else we can find.

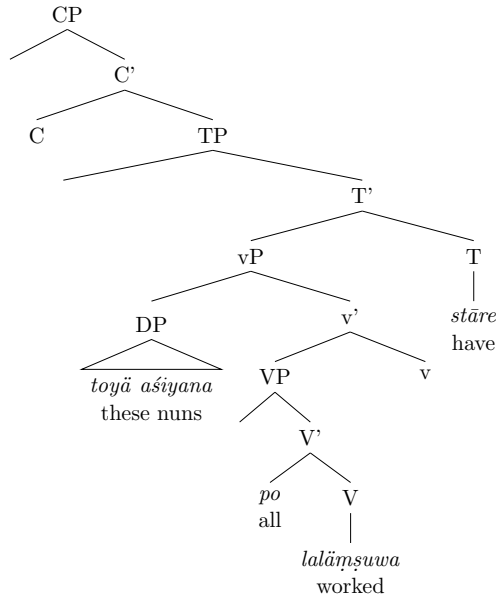
2.2.1 Auxiliary Evidence

- Tocharian possesses periphrastic perfect, future, necessitive, and potential constructions consisting of a participle/gerund and an inflected copula.
  - Notably, the overwhelming majority of these constructions cited by Adams place the auxiliary clause-finally, *after* the main verb.

(1) *toyä aśiyana po laläṃṣuwa stāre*
these nuns all worked be.3PL.PRET
"These nuns have worked everything" (MSL.19.160) Adams (2015)

- Assuming right-headedness within the TP, here's what the structure of this clause might look like after v to T movement (discussed below).

(2)



- The few exceptions to this generalization are straightforwardly derivable through topicalization of TP.

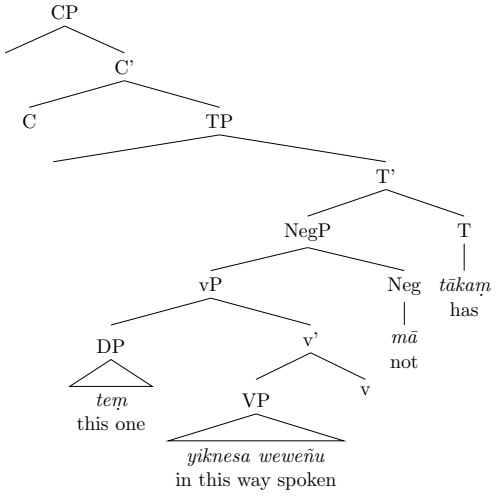
2.2.2 Negation Evidence

- According to Adams (2015), our old friend *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 in Adams (2015):

(3) *teṃ yiknesa weweñu mā tākaṃ*
this way spoken not be.3SG.SUBJ
"(If) he has not spoken in this way" (331b3/4^{L}) Adams (2015)

- Note how the negation appears precisely between the participle and the copula.
  - With our posited right-headed TP domain, we would expect our right-headed NegP to be located between the TP and vP layers.
  - At first glance, it looks like that's exactly where we find it.
    - This would also count as evidence of v to T movement in Tocharian, as the main verb appears to the right of (above) negation.

(4)



- However, if Neg is a head in Tocharian, it should block verbal head movement to T.
- But, we see evidence elsewhere that negation and the inflected verb act as a single constituent:

(5) *ka*[sic] *mā weścer krent (reki)*
why not say.you good word
"Why do you not say the good (word)?" (20b6^{C}) Adams (2015)

- I argue that inflected verbs first move from little v to the right-headed Neg head and merge with the negator (if present), and this newly-created complex itself then moves up to T.
  - Adopting the expanded left periphery of Rizzi (1997), TP may then itself be topicalized, ending up in TopP below the wh-question word in the highest specifier of CP in 5 above.
- Thus, we see that a right-headedness proposal for Tocharian within the TP domain not only accounts for the auxiliary constructions we see in the language, but also gives us a straightforward explanation of the behavior of preverbal negation.

2.3 Headedness in other phrases

- Slightly orthogonal to the otherwise clausal aim of the paper.
- Adams (2015) notes that adjectives and genitives usually precede nouns.
  - There are a few systematic exceptions to this, namely regarding adjectives and genitives referring to days and months.
- According to Adams (2015), there are 23 postposition constructions in Tocharian, compared to only six prepositions.
  - It's also worth noting the secondary case endings themselves as having developed from postpositional elements.
  - We see further synchronic head-final DP behavior evidenced by Gruppenflexion.

3 A brief look at mixed-headedness arguments across Indo-European

3.1 Anatolian

- 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.
  - He also notes the behavior of the auxiliaries *ḫark-* 'have' and *ēš-* 'be', which always follow the main verb, as seen below:

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

3.2 Sanskrit

- One of the most complete theoretical treatments of word order and local headedness is Schaufele (1991).
  - He follows most of western scholarship in assuming base SOV word order, and claims that the majority of phrases are head-final.

  - PPs seem to show mixed headedness, with the majority being head-final, behavior consistent with other ancient IE languages.
  - Hock (1984) notes that 97% of Vedic prose texts are verb-final, compared to 65% of poetic texts.
- Another potential piece of evidence for right-headedness within TP are the periphrastic constructions we see appearing in later Vedic, mirroring those seen in Tocharian and Hittite above.
  - The earliest attestation of the periphrastic perfect, for example, seems to be *gamayắm cakāra* 'he went' from the Atharvaveda.
  - The development of these later constructions could support the argument for underlying head-finality within TP in earlier Vedic.
- As far as evidence for left-headedness in the CP domain goes, 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 question particle *api* occurs clause-initially, instead of the clause final position we would expect if CP was right-headed in Sanskrit (e.g. *ka* in Japanese).

(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)

- I would be remiss to leave out clause-final quotative *iti* in this discussion, but note that Hock (1982) and Saxena (1995) claim that its complementizer-like behavior did not fully evolve until the classical period.
  - Further, this period is also exactly when clause-initial *yád* fully developed its own complementizer behavior according to Viti (2007).

3.3 Italic

- The most thorough work on phrasal headedness in Italic is undoubtedly Ledgeway (2012), who describes in detail the gradual change from head-final to head-initial exhibited throughout Latin to the modern Romance languages.
- Interestingly, however, the clausal 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.
  - For our purposes, we see that complementizers seen in the Latin data appear clause-initially.
- Ledgeway claims that the development of TP corresponds to the rise of the left-headed auxiliary constructions in later Romance.
  - But, clause-final auxiliary constructions already existed in Latin itself, both with the copula and *habere*.

(8) cum cognitum habeas quod sit summi rectoris [...] numen
when known you.have what is supreme.GEN ruler.GEN divine.will.ACC
"When you realize the will of the supreme lord" (Cic. *Fin.* 4.11, Ledgeway (2012))

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

3.4 Greek

- Like in the other old IE languages, the Greek complementizers ὅτι and ὡς occur initially in their clauses.
- As Goldstein (2015) mentions, "Ancient Greek is unique in its degree of word-order variation" (p. 18), which leads him to posit a flat VP structure for Ancient Greek as of Herodotus.
- However, according to Taylor (1994), Homeric Greek is primarily OV, with the younger Greek dialects developing more frequent VO word order later.
  - Further, per Bentein (2012), the oldest periphrastic constructions composed of the copula + the perfect participle show up as early as Homer, and we find them primarily appearing clause-finally as late as Herodotus (Rosén (1957)).
  - So, even if, as Goldstein states, there is not enough evidence to make an explicit decision about the behavior of TP in Herodotus, the older Greek evidence does favor a right-headed TP analysis.

3.5 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.
- And then, of course, there's Modern German, which many would consider the *Paradebeispiel* for left-headed CP/right-headed TP langauges.

4 Mixed Headedness in PIE?

- Delbrück (1893) was the first to attempt an overall clause structure reconstruction for the Proto-language, concluding that PIE must have been SOV based mainly on Sanskrit word-order evidence.
  - Here, I would like to go a step further, and tie this word order generalization to a specific underlying syntactic structure.
- Per Hackstein (2013), overt complementizer behavior seems to develop relatively late in the prehistory (or even history) of many of the IE daughter languages.
  - Zero-embedding is the most securely reconstructible method of sentential embedding across the earliest attestations of the daughter languages.
- However, even though we cannot reconstruct any of the individual complementizers for the proto-language, I argue that we *can* reconstruct their shared underlying syntactic structure, especially given their overwhelming similarities when they do emerge in the daughter languages.
- Also, if, as Hale and others have suggested, Wackernagel clitics head their own projections in the left periphery, then we have further support that PIE CPs were left-headed.

- Finally, we have Bresnan's Complementizer Attraction Universal, mentioned above.
  - With wh-movement (especially in question constructions) being so ubiquitous across the Indo-European languages, it makes sense to reconstruct this behavior for the proto-language as well.
- As far as right-headedness lower in the clause goes, I'm not the first to posit something like this.
  - Sapp (2016) reconstructs head-final VPs for PIE due to the Germanic evidence when considered alongside the SOV word order of the other IE languages.
  - He does not, however, go so far as to recostruct right-headedness for TP in PIE.
  - Similarly, Krisch (2017) reconstructs head-final VPs due to SOV evidence from across the Indo-European languages, but does not reconstruct TP at all for the proto-language.
- I argue that even though we cannot reconstruct any of the individual periphrastic constructions present in the daughter languages for PIE, we *can* reconstruct their shared underlying syntactic structure, especially given their similarities when they do emerge in the daughter languages.
  - Taken with the fact that we can and do reconstruct SOV word order and head-final VP behavior for PIE, the evidence in favor of going a theoretical step further and reconstructing a right-headed TP for the proto-language is worth serious consideration.

5 Future research directions

- First and foremost, I'd like to undertake more detailed projects in the vein of this Tocharian analysis for each of the other old IE languages.
  - The end goal of this project being, of course, securely reconstructing mixed headedness for the proto-language in the manner argued here for Tocharian specifically.
- To that end, I'd like to take a closer look at the behavior of negation specifically with regard to phrase structure for the rest of the IE languages.
- Also, for many IE languages, exceptions to verb-final behavior (other than topicalization of the verb) largely consist of single elements extraposed to the right, commonly called "right-detachment" in the literature.
  - I plan to account for, and motivate, these apparent examples of rightward movement in a leftward manner.
  - Ideally, such a study would also give us a more secure reconstruction of these constructions for PIE.

References

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. 2012. Verbal Periphrasis in Ancient Greek A state of the art. *Revue belge de philologie et d'histoire*, **90**(1), 5–56.

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

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

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 Wackernagels 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 Heinrich. 1982. The Sanskrit Quotative: A Historical and Comparative Study in Papers on Diachronic Syntax: Six Case Studies. *Studies in the Linguistic Sciences*, **12**(2), 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**(1), 89–107.

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.

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.

Probert, Ph. 2014. Relative clauses, Indo-Hittite and Standard Average European. *Pages 137–64 of: Proceedings of the 25th Annual UCLA Indo-European Conference*.

Rizzi, Luigi. 1997. The fine structure of the left periphery. *Pages 281–337 of: Elements of grammar*. Springer.

Rosén, Haiim B. 1957. Die <<zweiten>>Tempora des Griechischen: Zum Prädikatsausdruck beim griechischen Verbum. *Museum Helveticum*, **14**(3), 133–154.

Sapp, Christopher. 2016. Word order patterns in the Old High German right periphery and their Indo-European origins. *Diachronica*, **33**(3), 367–411.

Saxena, Anju. 1995. Unidirectional grammaticalization: diachronic and cross-linguistic evidence. *STUF-Language Typology and Universals*, **48**(4), 350–372.

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.

Viti, Carlotta. 2007. *Strategies of subordination in Vedic*. Vol. 57. FrancoAngeli.

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