

Indo-European Particles and Head-Movement-driven Word Order Change

Luke Smith

November 9, 2017

Abstract

Here I cover the grammaticalization of particles in early Indo-European languages into PP phrasal heads and verb prefixes. Into the fronting tendencies of such particles and then building off pre-generative sources, I argue that the characteristic word orders of Celtic (VSO) and Germanic (V2) are the results of earlier particle movement concomitant with changes in head movement parameters. I move to generalize that diachronic change from an SOV to SVO language can be easily come about gradually from the interface of Wackernagalian and other kinds of particle fronting, along with changes in head movement parameters.

Note: It's not typical to gloss examples in Indo-European linguistics, while I've glossed all relevant morphemes here, those sentences in languages I know less well (Hittite, Sanskrit) may have some errors as I'm often drawing them from sources which have not glossed them. I am 90% sure everything is not wrong.

In case you don't know: Some Hittite sentences might have material in brackets, as you can probably guess, this is partial or unclear material. There will also be words capitalized. This is to designate a Akkadian cuneiform logogram whose pronunciation in Hittite is unknown, instead, the capital letters are the Old Akkadian pronunciation.

1 Empirical Overview

Indo-European languages are notable in that there is a historical overlap between the class of adpositions and a class of preverbal derivational morphemes. For example, Latin verbs, and by extension English borrowings, may be prefixed with modifiers such as *in/de/ante/præ/pro*, etc., all of which also function in the language as individual prepositions.

Proto-Indo-European (PIE) **proti* yields Greek *pros*, Sanskrit *prati*, English *forth*, etc., all of which, in their historical forms can serve either as adpositions or verbal prefixes. Their typical use in earliest forms is illustrated below with examples from the Rig Veda.

- (1) **ápa** svásur uśáso Nág jihīte
away sister.ABL her.ABL Dawn Night goes
“Away from her sister Dawn Night departs.” (Rig Veda 7.71.1ab)
- (2) divyā āpo **abhí** yád enam áyan
gods.GEN water over CONJ him.DAT come
“when the water of the gods comes upon/overcomes him” (Rig Veda 7.103.2a)

Particles did enjoy relatively free placement due to the relatively unrestricted word order of the earliest Indo-European languages, but in every case, they would precede the verb (thus are often referred to “preverbs” in some literature).

Native speakers of these languages did have the percept that these particles were not fully independent ones, but part of the verb itself. The traditional grammatical term for these constructions, *tnesis* lit. “splitting” reflects the fact that early Greek grammarians viewed the constructions as severed verbs with a head and a “prepositional” preverb, hence the typical term “preverb” in historical linguistics.

Still, the actual chronology of the constructions are, as we will discuss, the reverse. The earliest attestations of the earliest languages show very active

tnesis, and only as time progresses do the particle and verb join in one way or another.

At that, particles also cliticize to nouns yielding the adposition now characteristic of all descendent languages. Reconstructing the “headedness” of PIE “adpositional phrases” is particularly difficult because it is not totally clear if any such entity existed. In early Indo-European languages, most of the semantic space governed by PP in English was taken care of by a robust case system, presumably buttressed these particles, not as adpositional heads, but general adverbials, syntactically neither strict adpositions or verbal morphemes.

1.1 Reconstruction of particles

The first literary attestations of Indo-European writing and folklore witness to this fairly directly. The Homeric Epics, the Rig Veda and most early Hittite writings show languages where particles are just that, with no canonical location in the sentence nor adpositional quality with their putative objects. We can see that the location of the particle with reference to its logical object in Hittite is not consistent, (3a) is apparently postpositional, while (3b) is apparently prepositional. It should be noted that (4), similar to the earlier Sanskrit examples appears “prepositional,” despite Sanskrits ongoing emergence into a militantly postpositional language.

- (3) a. ERIN-ti-ma-ssan **ser** GIR ZABAR kitta
troops-LOC-CON-PART on bronze dagger lies
“On the troops a bronze dagger lies.”
- b. huiswatar-ma-pa **anda** hingani haminkan
life-CON-PART in death.DAT.LOC be bound
“Life is bound into death.”
- (4) **prá** vātā vānti
for winds blow
“The winds blow (forth).” (Rig Veda 5.83.4ab)

It could be argued that this is not necessarily proof of a *lack* of quasi-adpositional relationship between the nominal and the particle, as early IE language were subject to excessive hyperbaton, scrambling constituents in unintuitive ways. It might be that these examples are hyperbatized ones. Still, in those languages where adpositions had come to be fully grammaticalized, adpositions seem consistently *unavailable* to this kind of scrambling, usually remaining proclitic on the head noun.

Horace, writing in some of the earliest available Latin, gives us stark examples of hyperbaton, but adpositions are never seen as deviating from their head. The following lines, from his Odes (I, 5) show substantial non-configurationality, but the two prepositional phrases are characteristically intact.

- (5) Quis multā gracilis tē puer in rōsā
 who.NOM much.ABL gracile.NOM you.ACC boy.NOM in rose.ABL
 perfūsus liquidīs urget odoribus grātō Pyrrha sub
 infused liquids.ABL urges aromas.ABL pleasant.ABL P. under
 antrō?
 cave.ABL
 “What slender boy infused with perfumes courts you with so many roses
 in a pleasant cave, Pyrrha?”

It should be noted that both prepositional objects (*rōsā* and *antrō*) are modified by adjectives outside of their PPs (*multā* and *grātō* respectively). But this a principled “transformation” in Latin (compare familiar *summā cum laude* constructions), where a syntactic constituent in a PP may move to a specifier position of PP, and may later be accessible by later operations. What is important is that the adpositional particle *is* in fact adpositional (or perhaps proclitic), and as such, is not subject to hyperbaton. Thus we might indirectly infer that those early Indo-European particles with freer order should not be considered “adpositions” in the strictest sense as even in those languages with liberal word order, they remain fairly well-behaved.

More than that, the earliest texts also yield some examples of pseudo-adpositional

particles without either nominal objects *or* accompanying verbs. These particles appear in the Iliad and Odyssey yielding an almost adverbial meaning:

- (6) a. theoi d=**epi** martyroi estōn!
 gods then=upon witnesses be.SUBJ
 “Let the gods be witnesses **thereupon!**”
- b. **en** d=etitheī melitos... amphiphoreas.
 in then=placed honey.GEN amphoras
 “And **thereon** he sat jars of honey.”

(6a) is particularly interesting in that not only is there no logical “object” to the preposition, but the metaphorical motion behind the particle is clearly *not* an extension of the semantically vacuous copula. This illustrates the primarily “adverbial” nature of the early particles, before they were interpreted as being syntactic units with either verbs or nouns.

However there is strong circumstantial evidence that the relationship between the particle and verb at least in some situations are not merely compositional even at the earliest examples of attested languages. Certain particle/verb combinations, for example, exhibit their own properties independent of both of their parts. Most well known is Latin *in-spicere* “examine, look closely.” Normally Latin verbal objects will take accusative case, while objects of the preposition *in* will take either ablative (with a locative meaning) or accusative (with an allative meaning). *Inspicere*, however takes an object in the *dative*, unlike both its components. Boley (2004) notes the same of the Hittite particle *anda*:

- (7) [u]t[n]e **anda**-le-aut[ti le] arsanesi
 land in-NEG-look NEG envy
 “Do not look into the land; do not envy [it.]”

This level of co-grammaticalization is indeed stark, but is limited to only several combinations of lexical items. The general pattern remains of fairly productive particles that add a coherent compositional meaning to a clause.

1.2 Patterns in particle placement

As I've alluded to here, particle placement is fairly free, save the requirement that the particle linearly precede the verb head. Still, as Boley (2004) notes, gradual patterns emerged in particle placement. These placements will be important for our analysis.

1.2.1 Sentence initial particles

- (8) **amphi** de teikhos helasse polei
around thus wall.ACC built city.DAT
“built a wall around the city”
- (9) **en** de moi alphita kheuon eurrapheessi doroisin
in then me.DAT goats poured wellstitched skins
“pour me goats into wellstitched wineskins”

1.3 Second position particles

Something fishy would be going on if we ever failed to see Indo-European languages employing the second-position at every opportunity.

Some examples of second position particles have already been presented (see (6) and (3)), but Homeric Greek is especially replete with examples. Notice in those examples below, there is true Wackernagelian movement to second position, where we might otherwise expect the “adposition” to be clause initial, before its logical object.

- (10) a. **seo** d=ek tade panta pelontai!
you.GEN then=from these all come
“Because of you all this is happening!” (Iliad 13.632)
- b. **tu** m=ek phasi genesthai.
him.GEN 1S=from they say be sired
“They say that I was sired from him.” (Odyssey 1.220)

- c. **tō** d=ama kēryka proiei menos Alkinooio... Arētē
 him.DAT then=with herald sent self.NOM Alcinous Arete
 d=ara **oi** dmōas am=epempe gynaikas.
 then=thus him.DAT slaves.ACC with=sent women.ACC
 “A herald sent Alcincous with him... Arete sent some maidservants
 with him.” (Odyssey 13.64,66)

Note also that this movement is that movement is “prosodic” in the sense of Agbayani and Golston (2010) in that it violates the contiguity of syntactic constituents freely.

- (11) **bioton** d=**apo** pampan olessei
 substance then=away all will destroy
 “... will destroy all of my substance.” (Odyssey 2.49)

I should note that, although it’s just a statement from lack of evidence, that while second position particles are clearly attested in Greek, Hittite, Avestan and other early languages, they are suspiciously hidden in Sanskrit. This might in fact, be a relevant data point for the generalizations later.

1.4 Multiple exponence

While on the issue of placement, it is important to note that as particles became grammaticalized, there are many examples of multiple exponence of these particles. Below, for example, the Hittite particle *kattan* “under” appears as a preverb and perhaps in a more particle-like or perhaps proto-adposition as well.

- (12) nu eshani **kattan** NINDA **kattan** appanzi.
 then blood under bread under they place
 “They place the flat-bread under the blood.”

This is similarly common in Greek, which gives at least several examples of a triple reflex of a particle, preposition and preverbal all in the same clause, as in (13b).

- (13) a. **en** d=oinon echeuen **en** depai golden
 in then=wine.ACC poured in cup.DAT golden.DAT
 “He poured wine in a golden cup”
- b. **en** d=hyperas te kalus te podas te **en**-edēsen **en** autē
 in then=braces and reefs and sheets and in-fixed in it
 “He fixed the braces, reefs and sheets on it.” (Odyssey 20.260-1)

2 Towards a Theory of Diachronic Word Order

2.1 “Univerbation” as Head Movement

The question of the “typical” word order of PIE, if any, is still an object of some debate (Friedrich 1975), but an emerging consensus circulates around an unmarked SOV word-order, with a common, albeit perhaps marked VSO alternative (Forston 2004). The later, as a generative syntactician may intuit, is mostly due to the same semantic and discourse features that often drive T-to-C movement in their descendant languages: polar questions and often imperatives.

Still, the particular problem of reconstructing a canonical word order in PIE stems from the fact that nearly all Indo-European language families show systematically different word orders. While some of the oldest tend toward SOV (Sanskrit, Hittite, Latin), outside of the Indo-Iranian family, VO word orders are the norm, be it conventional SVO, VSO as in Celtic, or historically verb-second of Germanic.

Now despite this variety, it should be said that in each case of VO order, there is a pattern leading back in time toward a SOV original. In the traditional terms of historical linguistics, all IE languages in one way or another underwent a process of *univerbation* in which the adposition-like particle and the V head became merged in one prosodic word in one way or another.

Indo-Iranian languages did this in such a way that the particle became conventionally realized on the verb in final position, but in each other case, as we will

see, the V head is raised to be merged with the particle in whatever position it commonly occupies in a language.

As addressed in the previous section, above all, particles either ended up in second position or utterance initially, we'll see that different languages show this reflex in different ways.

2.1.1 Celtic

Watkins (1963, 1964) argues that Proto-Irish, and by extension earlier Celtic Languages, began as SOV languages which tended to place verbal particles sentence initially. This created a default structure of PSO_V, with the somewhat awkward situation of the two logical elements of the verb maximally far apart.

Watkins reasons that in Old Irish, VSO becomes increasingly common as an attempt to merge the particle and verb in one phonological constituent, partially evidenced by the lack of mutation when particles are followed by their verbs, and its presence elsewhere.

In formal terms, this is simply a change in the head movement parameter of V (thus moving to T or further, depending on the analysis), and is a parametric change precipitated by a kind of phonological optimization. While in the earlier language, particles may have been fronted as Wackernagelian movement, as particles become more lexically and logically associated with a verbal host, movement of the entire verb becomes preferable.

2.1.2 Germanic

Hock (1991) shows a nearly equivalent course of diachronic change, not with particle verbs *per se*, but with the fronting of the earliest auxiliaries in Germanic. He reasons that Germanic languages gradually developed V2, and then the SVO in some cases in a multistage process. At the earliest level, Proto-Germanic can be said to be SOV, with auxiliaries still to the right, the structure we should

expect from an obedient head-final language.

- (14) flagda faikinaz ist
demons menaced is
“He is menaced by evil spirits” (Runic inscription)
- (15) thrawijan haitinaz was
throes desitined was
“He was destined for the throes.” (Runic inscription)

While this seems to be the state of Germanic at earliest attestation, centuries later, particularly the key texts of the period, we see a stark evolution of AUX-second constructions, still with linearly final Vs. Phonologically minor verbs, particularly copulas are often fronted to second position as well.

- (16) Bēowulfe wearð gūðhrēd gyfeþe.
Beowulf.DAT was battle glory given
“To Beowulf was glory in battle given.”
- (17) þæt wæs wræc micel wine Scyndinga.
that was sorrow great friend.DAT Scyndings.GEN
“That was great sorrow for the friend of the Scyndings.”

This should bring Germanic into the familiar territory of the common analysis of modern German V2, where the highest head in the clausal spine moves into C. The general correlary of this: lack of V2 in subordinated clauses with a filled C, is also attested (Kroch 2009), see below, but this is *not* a *universal* tendency in Old English (cfr. Pintzuk (1993) for more).

- (18) ... ðeah hit ær upahæfen wære
... although it before up-raised was
“...although it was raised up before.”

Still, for a period in the early middle ages, most all Germanic languages converge upon V2 syntax of various militancy. But as Hock notes, V2, in a way similar

to verbal tmesis in other Indo-European languages, is “unstable” in a way. Particularly, it violates (Otto) Behaghel’s First Law, stating...

- (19) “Elements that belong close together intellectually will also be placed close together.”

Put in more contemporary terms, there is a constraint against phonologically distant logical units, an intuition in line with Boley (2004)’s similar statements on Old Irish and other Celtic languages.

2.1.3 Romance

It can also be said in brief that Romance languages, in the shift from Latin’s typical SOV structure to the now near universal SVO underwent a series of changes equivalent to Germanic languages.

Although it has been subject to ongoing debate (MacKenzie 2010), medieval Romance languages seem to have gone through a period of verb second syntax (be it German-like or Yiddish-like) on their way to the contemporary structure (Vance 1997; Wolfe 20015).

2.1.4 Lol greek

It hit me literally at the last minute that maybe i should talk about Greek word order change, since it’s actually the most consistently attested language all the way back. that would probably be the best validation/refutation, but I actually don’t know anything about Greek since the the common era, so that’s for next time I guess!

3 Closing Thoughts

Here we've overviewed the conundrum of Indo-European particle verbs. These particles, which in PIE may have been general sentential modifiers become further grammaticalized as both prepositions and verb prefixes in the earliest languages. These "prefixes," however are often split from the verbs they accompany, found most commonly in first (possibly spec CP or C itself) or second position (possibly C).

Yet the ubiquitous tendency in all daughter languages is to phonologically merge the verb and its particle, and depending on how this is done, it results in different canonical word orders for each language.

Indo-Iranian languages keep particles low, maintaining the traditional SOV order of the earliest known Indo-European. Celtic languages, where particles resided in the clause-initial position, underwent a change in head movement parameters bringing the V up in the clausal spine, resulting in VSO. We can compare this to similar syntactic changes in Germanic and Romance, both of which move toward SVO orders by particle and auxiliary raising to second position which is later generalized to all verbs (in some languages).

This I think moves to a pretty tangle nexus of how word order change can happen systematically. Indo-European languages fronted particles, perhaps out of the general Wackernagelian motive (or perhaps due to constraints analogous to Behaghel's Fourth Law), but as they became logically associated with verbs, the distance between them and the verb head accrued a more significant "constraint violation" so to speak. This caused a parameterized shift in particle positioning or verb movement to fix the new leak.

Hegelian. All too Hegelian.

References

- Agbayani, Brian and Chris Golston (2010). “Phonological movement in Classical Greek”. In: *Language* 86.1, pp. 133–167.
- Boley, Jacqueline (2004). *Tmesis and Proto-Indo-European Syntax*. Innsbrucker.
- Forston, Benjamin W. (2004). *Indo-European Language and Culture*. Blackwell.
- Friedrich, Paul (1975). *Proto-Indo-European Syntax: The Order of Meaningful Elements*. Montana College of Mineral Science and Technology.
- Hock, Hans Henrich (1991). *Principles of Historical Linguistics*. Mouton de Gruyter.
- Kroch, Anthony (2009). *The V2 syntax of Old English*.
- MacKenzie, Ian (2010). “Refining the V2 Hypothesis for Old Spanish”. In: *Bulletin of Hispanic Studies* 87, pp. 379–396.
- Pintzuk, Susan (1993). “Verb seconding in Old English: verb movement to Infl”. In: *The Linguistic Review* 10, pp. 5–35.
- Vance, Barbara S. (1997). *Syntactic Change in Medieval French*. Springer.
- Watkins, Calbert (1963). “Preliminaries to a historical and comparative analysis of the Old Irish verb”. In: *Celtica* VI, pp. 1–49.
- (1964). “Preliminaries to the reconstruction of the reconstruction of the Indo-European sentence structure”. In: *Proceedings of the Ninth International Congress of Linguists*. Ed. by Horace G. Lunt. Mouton de Gruyter.
- Wolfe, Sam (20015). “The nature of Old Spanish verb second reconsidered”. In: *Lingua* 164, pp. 132–155.