Syntactic Reconstruction in Indo-European: State of the Art

Thórhallur Eythórsson & Jóhanna Barðdal
University of Iceland & Ghent University

1. Introduction

Although Franz Bopp mainly concentrated on comparing lexis and sounds across the early Indo-European languages, he was still particularly interested in grammar, especially verbal endings and their origins. Bopp also paid close attention to how words combine into sentences in the earliest languages, which involved a comparison of sentence types and syntactic and morphosyntactic constructions. For instance, in his work on the conjugation system (1816), Bopp showed that the aorist of Greek and Sanskrit are cognate formations. According to Wackernagel (1926: 25), this was a pioneering study of syntactic patterns within the paradigm of the Comparative Method (Campbell 1995: 1145). The first steps towards developing the Comparative Method were laid out by Rask in his famous prize essay from 1814 on the origins of Icelandic and the Old Norse language, later published as Rask (1818). It was Bopp, however, who developed and systematized this methodology on which the Comparative Method came to be based.

Not only did Bopp systematically compare lexical and phonological units, but also larger morphosyntactic structures. Thus, it has always been implicit in the standard operating procedure of the Comparative Method that syntactic and morphosyntactic units can be subjected to it, but early grammarians like Bopp, Delbrück and even Wackernagel simply lacked the analytical tools to carry out and formalize such reconstructions. As a result, there was a long period of skepticism regarding the viability of applying the Comparative Method to syntax. In more recent years, however, there is an increasing body of work demonstrating that syntactic reconstruction is not only possible but also of great importance to historical linguistics.

In the remainder of this article, we first give a brief outline of the history of syntactic reconstruction in Section 2, before reviewing current comparative syntactic work on reconstruction on the basis of the Indo-European languages in Section 3. Contemporary syntactic reconstruction of the Indo-European languages can be divided into three different strands, the traditional Indo-European paradigm, the generative paradigm and the Construction Grammar paradigm. As an example, we lay out the details of how fragments of grammar can be reconstructed by means of Construction Grammar, including a reconstruction of grammatical relations for Proto-Germanic. In Section 4 we summarize the main arguments against syntactic reconstruction that have been brought forth in the literature and show why they lack validity. Section 5 contains our conclusions.

---

1 We are indebted to Leonid Kulikov and two anonymous reviewers for their valuable comments. We also thank the editors for inviting us to contribute to this celebration volume in honor of Franz Bopp. This research was supported with a generous research grant to Jóhanna Barðdal from the European Research Council (EVALISA, grant nr. 313461).
2. Earlier Ventures in Syntactic Reconstruction

The work carried out within the neogrammarians paradigm during the latter part of the 19th and the first half of the 20th centuries was focused on phonology, morphology and the lexicon. Syntax was still an underdeveloped field, although a number of important syntactic discoveries were made, like the placement of clitics in second position by Wackernagel’s Law (Wackernagel 1892), the position of the verb in Vedic and other archaic Indo-European languages (Delbrück 1878), the morphosyntactic and functional properties of the imperative form of the verb in Indo-European (Thurneysen 1885), Verb-Second in Old French (Thurneysen 1892), the function of the cases across Indo-European (Delbrück 1907, Havers 1911), and the tentative reconstruction of ergative alignment for (early) Proto-Indo-European (Uhlenbeck 1901, Pedersen 1907, Vaillant 1936). There was also a lively debate on whether subordinate clauses were found in Proto-Indo-European or not (Hermann 1895, Brugmann 1925), and intensive comparative work on the syntax of mood was also conducted (Jolly 1872, Thurneysen 1885, Delbrück 1893–1900).

In the wake of research on typological universals (Greenberg 1966), attempts at developing reconstruction methodology for syntax were made, in particular by Lehmann (1974). Lehmann argued that Proto-Indo-European must have been an SOV language, based on the occurrence of OV word orders in simple sentences in the earliest texts, as well as the word order found in other syntactic contexts, such as with nominal modifying constructions, verbal modifying constructions, mediopassive constructions, and postposed coordinators. The following year, two further publications appeared on Indo-European basic word order: Friedrich (1975) and Miller (1975). On the basis of word order frequencies in the earliest Indo-European texts, especially Homer, Friedrich suggested that the basic word order in Proto-Indo-European was SVO. Friedrich’s reconstruction is thus based on statistics and the pattern of the majority. Miller, on the other hand, argued that the basic word order in Proto-Indo-European must have had all three basic word order types: SOV, SVO, VSO. The arguments for this claim came from the word order patterns found in later stages of the Indo-European languages.

The reconstructions by Lehmann, Friedrich and Miller were met with severe criticism, in particular Watkins (1976), echoed by Jeffers (1976), Lightfoot (1979) and Winter (1984), among others, emphasized the futility of the typological approach to syntactic reconstruction. These scholars also highlighted the fact that seemingly random conclusions can be reached on the patterns of basic word order in Proto-Indo-European, all depending on the the selection criteria of the material on which the reconstruction is based. This criticism had a demoralizing effect on the field of historical Indo-European syntax for decades to come.

Despite his critical tone, Watkins (1976) still had some concrete and constructive proposals on how to conduct syntactic reconstruction, which he argued could be done on the basis of morphological material (Barðdáð & Eythórsson, in press). In his 1995 book, Watkins fleshed this proposal out in detail, giving countless examples from Indo-European poetic language of formulaic expressions found across the early languages. Focusing on set phrases and identical collocations in comparable contexts from literature within the same genre, Watkins was in effect able to reconstruct fragments of Proto-Indo-European morphosyntax. Thus, Watkins’s results and especially his methodology are of great worth not only for Indo-European poetics but also for syntax.

Further significant advances in the methodology of syntactic reconstruction have been made by contemporary scholars like Harris (1985) Harris & Campbell (1995, 2002), Gildea (1992, 1998, 2000), and Kikusawa (2002, 2003). Harris & Campbell (1995) developed a thorough research program on how to carry out syntactic reconstruction, based on the concept of syntactic pattern. Gildea (1992, 1998), in his meticulous work on the grammar of the Cariban languages, convincingly showed how innovations and archaic morphosyntax could be teased apart, serving as the basis for the reconstruction of the Proto-Cariban alignment system. Kikusawa (2002, 2003), who has also reconstructed alignment for Proto-Central-Pacific on the basis of the development of the pronominal systems, introduced the notion of cognate structures into the discussion on syntactic reconstruction.

More recently, the principles and methodology of syntactic reconstruction have been made even more precise, as is evident in a number of papers in Ferraresi & Goldberg
Further arguments for the legitimacy of syntactic reconstruction have been put forward, as well as suggestions on how to formalize syntactic reconstruction in a precise and theoretically-coherent manner. See Hale (1987a, 1987b, 2015), Garrett (1990), Willis (2011) and Walkden (2013, 2014), within the generative framework, and Barðdal and Eythórsson in several contributions within the construction grammar framework (Eythórsson & Barðdal 2011, Barðdal & Eythórsson 2012a, 2012b, in press, Barðdal 2013, 2015, Barðdal & Smitherman 2013, Barðdal et al. 2013). A further explication of the contributions made by these different approaches will be outlined below.

3. Syntactic Reconstruction within the Indo-European Language Family

One can identify three strands of research focusing on syntactic reconstruction: Work carried out within i) the traditional Indo-European paradigm, ii) the generative paradigm, and c) the construction grammar paradigm. We will now discuss each in turn (3.1–3.3).

3.1 The Traditional Indo-European Paradigm

Starting with the first strand, syntactic work done within the traditional Indo-European paradigm, we can mention, as some of the most recent work of this kind, Gamkrelidze & Ivanov (1995), Mendoza (1998) Bauer (2000), Lühr (2008), Balles (2008), Fritz (2010), Hock (2013), Keydana (2013), Kulikov & Lavidas (2013a), Coticelli Kurras & Rizza (2013), Viti (2014), Luraghi (2016), and Luján & López Chala (2016). Several book volumes on comparative Indo-European syntax and/or syntactic reconstruction have also seen the light of day in recent years (Ferraresi & Goldbach 2008, Kulikov & Lavidas 2013b, Viti 2014, Luján, Barðdal & Gildea (in press), speaking for the increased interest in syntactic reconstruction. We would like to comment in particular on Lühr (2008) and Hock (2013).

Lühr (2008) argues, and quite convincingly so in our opinion, that subordinate clauses not only can, but must be reconstructed for Proto-Indo-European. She focuses on the early reflections of modern-day that-clauses, and not only does she demonstrate that such clauses must themselves be reconstructed for the proto-language, but also a complementizer corresponding to that on the basis of data from Late Hittite, Old Avestan, Old Persian, Homeric Greek, and Old Saxon. The particular complementizer form, *kw;i-, is also cognate across at least Hittite, Ancient Greek, and Germanic, while the early Indo-Iranian languages have another etymon, *io-, in the complementizer position. While the neogrammarian debate evolved around whether the proto-language had embedded clauses or not, the modern debate is focused on the nature of the subordinate clauses that must be reconstructed for the proto-stage, namely whether they really are embedded or whether they only are adjoined. Lühr’s work is a solid contribution to the long-standing debate on the status of subordination in Proto-Indo-European.

Hock’s (2013) article focuses on the position of the finite verb in Proto-Indo-European. Contra earlier scholarship, Hock points out that recent typological research does not exclude the occurrence of relative clauses containing relative pronouns and finite verbs in SOV languages. Hence, such structures are in principle reconstructable for a proto-stage with an SOV order, despite earlier arguments against such a reconstruction (Lehmann 1974, Friedrich 1975). In addition, Hock argues that the fact that non-initial finite verbs in main clauses in Vedic lose their accent can only be motivated under the assumption that the verb was final. The reason is that there seems to have been a prosodic conflict between the high-pitch accent of the verb and the falling intonation characteristic of clause-final position. This conflict was resolved by the verb losing its accent utterance-finally (see also Hock 1986 and Klein 1992). A further argument for Proto-Indo-European SOV order stems from i-apocope in Italic, Insular Celtic and Baltic-Slavic, whereby finite verbs lose their word-final -i. Hock (2006, 2013) and Weiss (2009) have argued that this i-apocope is indeed an utterance-final reduction, compatible with the assumption that the finite verb occupies clause-final position.
Hock thus shows that it is absolutely possible to formulate a fruitful hypothesis on word order in Proto-Indo-European, contra claims made by Watkins (1976), Lightfoot (1979, inter alia), Pires & Thomason (2008: 50) and Fortson (2010).

Most of the otherwise excellent publications on comparative and reconstructed syntax listed above suffer from one non-trivial shortcoming, which is that an actual reconstruction is not really carried out. That is, the final step of the Comparative Method is not taken. Surely, correspondence sets are set up, the alternatives are compared, and a conclusion is made on the basis of the comparative facts and how they fit with the data and the grammar in general. It is claimed that a syntactic or a morphosyntactic object must or can be reconstructed. The final output, however, the reconstruction itself, is not given a formal representation. Such a representation is needed in order to explicate the analysis and flesh out the details of the reconstruction. An analysis that does not employ a thorough and precise representational system does not really result in a grammatical model, only vague ideas about how such a model may look like.

In some work within Indo-European comparative syntax (Gamkrelidze & Ivanov 1995, Kulikov & Lavidas 2013, Cotticelli Kurra & Rizzà 2013, Luján & López Chala 2016), a diachronic scenario may be drawn up, where different developmental stages of the analysis are outlined. Implicit in such analyses is a reconstruction of the original stage, although the exact status and the details of the reconstructed final output are not really fleshed out. Thus it turns out that, valuable though they may be, most syntactic reconstructions within the traditional Indo-European paradigm are in fact incomplete.

### 3.2 The Generative Paradigm

Early comparative syntactic work within the generative paradigm is represented by Hale (1987a, 1987b) and Garrett (1990). Hale carried out extensive work on clitics and the Wackernagel position, based on comparative evidence from Vedic and Hittite, which resulted in a reconstruction of a hierarchical clause structure around the Wackernagel position in Proto-Indo-European. Hale convincingly showed that topic and wh-elements must have occupied two different syntactic positions, which were represented in the tree structure he posited for Proto-Indo-European. Two arguments were provided for this analysis: The first argument is based on the position of fronted elements in the clause and the second argument on the distribution of clitics. Hale’s analysis was further developed by Garrett (1990) for clitics in Hittite. There is no doubt that Hale’s work was a sensation at the time and marked the beginning of an era, inspiring confidence in the historical linguistics community that it was possible to formalize comparative Indo-European syntax in a precise manner and thereby to make new linguistic discoveries in this long-neglected field of study.

More recent work within the generative paradigm has been done by Willis (2011) and Walkden (2009, 2013, 2014). Willis (2011) reconstructs aspects of the grammar of Common Brythonic, involving free relative clauses, cognate relative markers, negation and stress patterns, on the basis of data from all the medieval and modern Brythonic languages. Willis reconstructs a chain of changes involving both reanalysis and extension of existing patterns, hence uniting a set of complex data into a coherent story of grammatical change.

Employing the generative notions of I-language and E-language, Willis points out that they restrict the possible reconstructions one can posit. That is, the data that language learners produce (E-language) must be compatible with their input data, which in turn is based on the grammar of the source language (I-language). This means that a reconstruction of a grammar cannot generate data that is radically different from the inherited grammar, on which the reconstruction is based.

An important contribution of Willis’s work relates to his distinction between universal and local directionality. By the term *universal directionality*, Willis refers to morphosyntactic changes known through the large amount of work on grammaticalization, for instance. Drawing on data from a wide variety of languages, the grammaticalization program has documented that there is a compelling tendency towards unidirectionality of change, although this is first and foremost manifested as a robust statistical preference. By the term *local directionality*, however, Willis emphasizes that often the linguistic data under scrutiny...
only allow for one interpretation of the observed historical changes, and hence the reconstruction falls out directly from these very local factors. This is, indeed, a very important observation, one that we have also highlighted ourselves in our own work (Barðdal & Eythórsson 2012a: 267–268, Barðdal 2013: 448); hence Willis's documentation of what local directionality involves is a most welcome contribution to the field of syntactic reconstruction. It follows from this that neither syntactic “laws” nor any universal “directionality” in language change is needed for reconstructing syntax.

Another example involving local directionality has been brought forward by Walkden (2009, 2013, 2014). This is the verbal “middle” ending -sk in Old West Nordic, -s in Old East Nordic, which has developed from a cognate reflexive pronoun, sik ‘self’, first into a clitic and then later into a derivational ending. This development, however, has neither taken place in West nor East Germanic, even though the reflexive pronoun is also found there, cf. Gothic sik and Old High German sih. Hence, a comparison between the Germanic daughter languages reveals that, although not involving identity, the North Germanic situation must be regarded as an innovation. Walkden uses this example, contra Lightfoot (2002a), to highlight the fact that even when there is no identity across the daughters, syntactic reconstruction can still be carried out.

Furthermore, Walkden (2009, 2014), drawing on data from Eythórsson (1995) and Ferraresi (2005), suggests a reconstruction of the word orders where the finite verb occupies second position (V2) or third position (V3) in declarative main clauses for Proto-Northwest-Germanic. Using Rizzi’s (1997) expanded CP analysis to formalize his reconstruction of the position of the finite verb, Walkden develops existing analyses of the conditions underlying V2 and V3 further, and argues that the position of the verb in both V2 and V3 structures is conditioned by information-structural factors. V2 occurs in main clauses in both North and West Germanic under two types of conditions: a) with focused wh-elements in first position, and b) in neutral word order, most likely through a process of generalization. Walkden further claims that there are also arguments for taking V3 to be a variant of “neutral V2”, consisting of V2 together with an additional element in preverbal position. Although there may be disagreement on the details of the analysis, it is clear that V2/V3 can be reconstructed for Proto-Northwest-Germanic in structures where the requirements posited by Walkden are fulfilled.

3.3 The Construction Grammar Paradigm

We have argued extensively in our work on syntactic reconstruction that the emergence of Construction Grammar as a research framework is ideal for reconstructing linguistic objects at a larger level than the level of the word or morpheme, as it emphasizes all aspects of a construction: the phonological, morphological, syntactic, semantic, and discourse-pragmatic ones. Since the basic assumption within Construction Grammar is that constructions are form–meaning (or function) correspondences, consisting of a form part and a meaning part, syntactic units also qualify as constructions and can as such be reconstructed.

In the reconstruction program that Watkins (1995) envisaged, the role of morphology is emphasized. Watkins argued that syntax can indeed be reconstructed provided there are some morphological clues to guide the reconstruction. This is of course visible in both the proposed reconstruction of Willis’s (2011) of the relative marker in Common Brythonic, and Walkden’s (2013, 2014) reconstruction of the “middle” construction in Proto-Northwest Germanic. As we have emphasized elsewhere (Barðdal & Eythórsson 2016), the possible criteria for reconstruction accepted by Watkins constitute by no means an insignificant portion of grammar – the whole of morphosyntax.

Some of our own syntactic reconstructions are indeed based on morphological pointers, as for instance our reconstruction of argument structure constructions, mostly focusing on the Accusative and the Dative Subject Constructions (Barðdal & Eythórsson 2012b, 2016, Barðdal 2013, Barðdal & Smitherman 2013). Thus, we have suggested a reconstruction of a lexically-filled subconstruction of the Dative Subject Construction, DAT-iso-woe, for Proto-Indo-European (Barðdal 2013, Barðdal et al. 2013). The reconstruction is based on lexical items, the internal order between these items, the dative case, and the
discourse-pragmatic function of the construction, being first and foremost used in situations of severe adversity. The pragmatic properties of the DAT-is-woe construction in language use also entail a high degree of speaker involvement, reconstructed as ‘speaker’s dismay’.

In relation to our work on the DAT-is-woe construction, we have also reconstructed several construction types for Proto-Indo-European that are purely schematic: a bare exclamative construction, a dative exclamative construction, a predicative construction, a subject–predicate construction, and finally a fronted-predicate focus construction, of which the last two involve different word order patterns. The subject-predicate construction involves neutral word order with the subject preceding the predicate. The fronted-predicate focus construction, in contrast, shows the opposite word order with the predicate preceding the subject, due to the focal nature of the predicate in such contexts.

More specifically for the DAT-is-woe construction, we have reconstructed one type of language change, namely the lexicalization of an exclamative variant, involving woe-DAT, which exhibits a fixed word order. This construction always occurs with woe in clause-initial position, followed by a dative noun phrase (DAT), and an elliptical ‘be’. This is the form that is most widely attested in the daughter languages, and it is most plausibly analyzed as a lexicalized variant of the fronted-predicate focus construction, due to its exclamative nature. Only on this analysis can we account for the strict word order of woe-DAT, and as its corollary, the absence of DAT-woe word order in the daughter languages.

Another language change that we have reconstructed is the loss of a productive rule of anticausativization involving oblique subjects in the history of North-Germanic (Barðdal 2015). In Old Icelandic, there is a systematic relation between a causative construction with an accusative or a dative object and its anticausative variant where the original object case marking is maintained on the subject of the anticausative. One example is the verb drífa ‘drive’, which can instantiate the causative transitive construction (1a) as well as its anticausative alternant (1b), thus showing a special type of lability also involving case marking:

(1a) Þegar vindurinn dreif þá burt.  
when wind.the.NOM drove them.ACC away  
’When the wind drove them away.’

(1b) Bátinn dreif langt frá ströndinni.  
boat.the.ACC drifted long from shore.the  
’The boat drifted far away from the shore.’

It is clear that this alternation is not productive in Modern Icelandic, although speakers of Icelandic may upon reflection deduce a relation between the two. At some point in history, however, this anticausativization process, involving P-lability and maintained case marking, was indeed productive. Whether it was still fully productive in Old Icelandic, or whether one has to go further back to Proto-Norse, is difficult to know; what matters here is that at some point this anticausativization process ceased to be productive. We have reconstructed two language stages, separated by this change. During the first stage, there is an anticausativization link between the two constructions, whereas at the second stage the non-productivity warrants a reconstruction of the grammar not containing such a link. This results in a lack of systematic synchronic relation between the two constructions in the conception of speakers of Modern Icelandic. In traditional and generative terms, this would correspond to the loss of a productive rule.

As another example of how syntactic reconstruction may be carried out within the framework of Construction Grammar, we would like to epitomize one of our own reconstructions from Proto-Germanic, namely our reconstruction of grammatical relations (Barðdal & Eythórsson 2012b). We have chosen this example for the simple reason that it takes current syntactic reconstruction even further than most prevalent syntactic reconstructions which focus on a morphosyntactic construction (like Walkden’s “middle” in Proto-North-Germanic) or a single schematic construction (like Hale’s Wackernagel position). Reconstructing grammatical relations, in contrast, is a multifarious procedure which requires the following steps:
To appear in *Veleia* 33 (2016)

a. by reconstructing argument structure constructions
b. by identifying potential subject tests for each of the daughter languages
c. by establishing which of the attested subject test constructions are reconstructable
d. by reconstructing the constructions that function as subject tests for the proto-language
e. by defining for each subject test construction how the subject (and the object) argument of the argument structure constructions behaves

When argument structure constructions and the constructions functioning as subject tests have been reconstructed, which includes defining the behavior of subjects (and objects) with regard to the subject tests, grammatical relations simply fall out from these reconstructed fragments of grammar.

For Proto-Germanic, the relevant reconstructable subject tests that we have established on the basis of comparison between Gothic, Old English, Old Saxon, Old High German and Old Icelandic are the following:

- subject–verb word order
- subject–verb inversion
- raising-to-subject
- raising-to-object
- control

We will let it suffice to present our reconstructions of one argument structure construction and its interaction with the Raising-to-Subject construction in Proto-Germanic. Consider the following examples of the verb ‘hunger’ from Gothic, Old Norse-Icelandic, Old High German and Old English:

**Gothic**

(2a) þana gaggandan du mis ni huggreið
this.one.ACC going to me not hungers
‘the one who comes to me will not starve’ (St. John 6:35)

**Old Norse-Icelandic**

(2b) mann hungrar þá til líkamligra krása
man.ACC hungers then to bodily delicacies
‘a man starts having cravings for fleshly delicacies’ (*Leif*. 4814.18.20)

**Old High German**

(2c) Mih hungrita, inti ir gabut mir ezzan
me.ACC hungered and you gave me eat.INF
‘I was hungry and you fed me’ (Tatian 152:3)

**Old English**

(2d) seðe cymes to me ne hynegreð hine
he.who comes to me not hungers him.ACC
‘the one who comes to me will not starve’ (Lindisfarne Gospels 1, St. John 6:35)

The verb ‘hunger’ occurs systematically with an accusative subject across the earliest Germanic languages, as evident from (2) above. The lexical item ‘hunger’ has already been reconstructed for Proto-Germanic by earlier scholars as *hungrian-, hence we only need to reconstruct it together with its argument structure for Proto-Germanic. The correspondence set is shown in Table 1, showing identity across the daughters.
Table 1: Correspondence set for the ACC-hungers argument-structure construction

<table>
<thead>
<tr>
<th>Language</th>
<th>Alt 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gothic</td>
<td>ACC-hungers</td>
</tr>
<tr>
<td>Old Norse-Icelandic</td>
<td>ACC-hungers</td>
</tr>
<tr>
<td>Old English</td>
<td>ACC-hungers</td>
</tr>
<tr>
<td>Old High German</td>
<td>ACC-hungers</td>
</tr>
</tbody>
</table>

The verb ‘hunger’ can thus be reconstructed together with its argument structure for Proto-Germanic as in Figure 1.

Figure 1: A reconstruction of the accusative-subject verb ‘hunger’ and its argument structure in Proto-Germanic

The FORM field in Table 1 is filled with the phonological material *hungrian-, while the SYN field contains the number of arguments and their case marking, only one argument in the accusative in this case. The SEM field defines the semantics of the construction and is here specified in terms of semantic frames. The relevant semantic frame is the need-for-intake-of-nourishment frame, occurring with only one argument in our examples, namely the NEEDER, which is here indexed with an i and co-indexed with the NP-ACCi in the SYN field. This is all it takes to reconstruct the verb ‘hunger’ with its accusative subject in Proto-Germanic.

Let us now turn to the Raising-to-Subject construction in Proto-Germanic and its interaction with argument structure constructions. Consider the following examples from Old Norse-Icelandic, Old Saxon, Old High German and Old English, all involving Raising-to-Subject:

Old Norse-Icelandic
(3a) að mér tekur nú að þykja minna gaman að gulli en var that me.DAT begins now to think.INF less entertaining at gold than was ‘that now I’ve started to take less pleasure in gold than before’ (Hreiðars þáttur, ch. 5)

Old Saxon
(3b) ððo beginnad imu than is uuerk tregan, an is hugi hreuen easy began him.DAT then his work regret at his mind rue ‘Easily, he began to regret what he did, to rue it in his mind’ (Heliand 3233)
Old High German

(3c) So imo daranah nôten gestat
so him.DAT thereafter get.into.trouble.INF began
‘and then he began to get into trouble’ (Notker, Boeth)

Old English c. 971

(3d) þa ongan hine eft langian on his cyþþ
then began he.ACC again long.INF for his kin
‘then he started to long for his family again’ (Blickling Homilies 113,15)

In these examples, the oblique subject verbs are þykja ‘think’, tregan ‘regret’, nôten ‘get into trouble’ and langian ‘long’, all occurring in the infinitive. The first three verbs select for dative subjects and the last one for an accusative subject. In all four cases, these oblique subjects behave syntactically as the grammatical subjects of the relevant finite aspectual verbs, which are taka ‘begin’, biginnan ‘begin’, stantan ‘begin’ and onginnan ‘begin’, while at the same time they maintain the case marking of the subjects of the lower verbs.

In all four examples in (3), there is a sentence adverb, nú ‘now’, than ‘then’, daranah ‘thereafter, and eft ‘again’, respectively, demarcating the boundaries of the infinitive clause. Since the oblique subject occurs before the sentence adverb, it is clear that the oblique is not located within the infinitive clause. In (3a) and (3c) the dative subject immediately precedes the finite aspectual verb, while in (3b) and (3d) there is subject–verb inversion with the finite verb occurring immediately before the dative vs. the accusative. There is thus no doubt that the oblique subject-like argument, assigned by the lower verb, behaves syntactically as the subject of the finite verb, which is indeed what Raising-to-Subject involves.

On the basis of the comparative evidence, it is clear that Raising-to-Subject was a subject test in the earliest Germanic daughter languages, and thus it can be reconstructed for Proto-Germanic. We have suggested the reconstruction shown in Figure 2 (Barðdal & Eythórsson 2012b: 385):

![Figure 2: A reconstruction of Raising-to-Subject for Proto-Germanic](image)

The formalism in Figure 2 consists of an outer box which demarcates the whole construction and smaller boxes within, which represent different parts of the construction. The ellipsis preceding the whole construction indicates that we are not reconstructing subject-first word order in Proto-Germanic in this construction. This must be done independently, as we have shown elsewhere (Barðdal & Eythórsson 2012b: 376–380). Each of the inner box representations contain a FORM field and a SYN field, respectively. Since the Raising-to-Subject construction we are reconstructing denotes the inceptive aspect, i.e. identifying the beginning stage of the event, there is also a SEM field in the middle box where the finite verb is located.

The asterisk outside the larger box signals that this is a reconstruction. The FORM field is systematically left empty throughout since the reconstructed object is a schematic construction. The SYN field specifies the syntactic category of each element. The SYN field in the rightmost inner box is defined as $V_{\text{INF}}$, representing the nonfinite verb. In the middle
box it is defined as $V_{\text{FIN}}$, representing the finite verb, while in the leftmost box it is defined as $\text{NP}_{i}$, representing the “raised” subject. The SEM field in the middle box is defined in terms of the begin-fr.

The schematic construction in Figure 2 also interacts with a lexical-class construction in Proto-Germanic, in which all inceptive verbs meaning ‘begin’ are listed. This lexical-class construction can be reconstructed as in (4), where a lexical reconstruction of the form ‘-*ginnan*’ is also included on the basis of the evidence from East and West Germanic where inceptive verbs with ‘*ginnan*’ are found.

(4) raising-to-subject-\(lxm\) => ‘-*ginnan*’

The schematic reconstruction in (4) can be laid out with the box formalism of Construction Grammar as in Figure 3, where it is shown very clearly by lack of any listed arguments in the SYN field that ‘begin’ verbs do not select for an argument of their own, let alone a subject. As a consequence, when an argument structure construction, Raising-to-Subject construction, and the verb-class-specific construction ‘begin’ interact, indeed no argument comes from the aspectual ‘begin’ itself, meaning that there is only one argument that can take on the subject behavior of the Raising-to-Subject construction, and hence of ‘begin’ in this construction. This argument is the \(i\)-indexed argument of the argument structure construction, a nominative in case of nominative subject verbs and an accusative in case of accusative subject verbs.

![Figure 3](image)

**Figure 3**: A reconstruction of the verb-class-specific construction of ‘begin’ in Proto-Germanic

We have here shown how fragments of the grammar of Proto-Germanic can be reconstructed. We have demonstrated how one can reconstruct argument structure constructions with an example of a reconstruction of an established subject test, namely the Raising-to-Subject construction. The behavioral subject properties of oblique subjects are found in the interaction between these two, the argument structure construction and the construction functioning as subject test. The indexing of the accusative subject of ‘hunger’ with an \(i\) indeed defines the behavior of the accusative subject in the Raising-to-Subject construction where the argument preceding the finite verb is also indexed with a subscript \(i\). Hence, when the argument structure construction merges with the Raising-to-Subject construction, it is already coded in the construction that the accusative subject will take on the subject behavior of the inceptive finite verb, through the indexing of the preverbal position as \(i\).

As mentioned in passing above, ordinary argument structure construction with a nominative subject also have the nominative indexed with a subscript \(i\). This means in effect that the subject of the lower verb will always take on the subject behavior of the higher verb when the argument structure and the Raising-to-Subject construction merge. This is irrespective of whether the lower verb is a nominative, accusative or a dative subject verb. This is how subject behavior falls out from our reconstructions of argument structure
constructions, our reconstruction of the constructions that function as subject tests, and the interaction between the two.

The next section is devoted to the main problems that have been pointed out as pertaining to syntactic reconstruction; we will lay out our argumentation against the validity of these alleged problems.

4. Arguments against syntactic reconstruction and why they do not hold

Several arguments have been used against the feasibility of syntactic reconstruction during the last century (Watkins 1976, Jeffers 1976, Winter 1984, Lightfoot 1979, Pires & Thomason 2008, Mengden 2008, inter alia). These include:

a) lack of cognates and hence lack of correspondences
b) lack of syntactic laws and hence lack of directionality in syntactic change
c) lack of arbitrariness
d) lack of form–meaning correspondences
e) lack of continuous transmission of syntactic structures during acquisition

We will now discuss each of these arguments in turn (see also Eythórsson & Barðdal 2011, Barðdal & Eythórsson 2012a, 2012b, 2016, Barðdal 2013). Starting with the assumption that there are no cognates in syntax, this was first claimed by Jeffers (1976) and has since been echoed by Lightfoot and others in numerous publications. Harris & Campbell, in their (1995) research program, argued for the usefulness of the notion of syntactic pattern (see Section 2 above), which serves as input for their correspondence sets. Kikusawa, in her work on the Polynesian languages, identifies cognate structures through the pronominal system. We have argued for the validity of the notion of cognate argument structure constructions in our earlier work, and devote an entire chapter to this topic in a forthcoming book on syntactic reconstruction (see Barðdal & Eythórsson 2016). The research community thus seems to have no problems at all with identifying cognates in syntax or with setting up correspondence sets.

The second problem with carrying out syntactic reconstruction is the alleged lack of syntactic laws, and hence of directionality in syntactic change. Postulating that linguistic change is “chaotic” due to its presumed “abductive” nature (Andersen 1973), Lightfoot (e.g. 2002a: 135) argues that except in cases of identity, reconstructing syntax is totally out of the question. This alleged impossibility is due to syntactic change being, by its own very nature, unconstrained. However, as we have argued in several papers (Barðdal & Eythórsson 2012a, Barðdal 2013, Barðdal et al. 2013), and as has been independently shown by Willis (2011), universal regularity of syntactic change is simply not needed in order to carry out syntactic reconstruction. In many cases, the order of the changes can be deduced directly from the data.

The third suggested problem for syntactic reconstruction is the lack of arbitrariness in syntax. This, however, is not a real argument because the arbitrariness requirement is first and foremost needed for demonstrating genealogical relatedness (cf. Harrison 2003). Syntactic reconstruction, however, is usually carried out only after genealogical relatedness has already been established. In addition, there is more arbitrariness in syntax than often assumed; collocational restrictions, case marking, and argument structure are some of the syntactic structures where a considerable degree of arbitrariness is found. Lack of arbitrariness is therefore no real argument against syntactic reconstruction.

The fourth argument against the viability of syntactic reconstruction is the alleged lack of form–meaning correspondences in syntax. There is no doubt that the 19th and early 20th century Indo-Europeanists who developed the Comparative Method emphasized the role of the linguistic sign for reconstruction. In order to be able to reconstruct, they argued, it is essential that both the meaning part and the form part show a correspondence across the dataset being compared. The reason for this strict criterion was to enable scholars to identify inheritance and to distinguish between forms that are inherited and those exhibiting accidental formal similarities without being inherited. This criterion has, indeed, been a fundamental methodological principle in establishing genealogical relatedness.
The most influential theoretical framework in modern linguistics, the generative program, does not assume that syntactic structures have any meaning or function. As a consequence, a reconstruction based on the main criterion of the Comparative Method, namely form–meaning correspondences, is excluded by definition. However, in Construction Grammar, where the basic linguistic building blocks are form–meaning correspondences, syntactic reconstruction is perfectly viable and falls directly out from the basic assumptions of the model. Not only does Construction Grammar take the assumption of form–meaning correspondences as the core of its program, it also formalizes constructions in terms of both form and meaning (as illustrated above). This makes Construction Grammar the ideal framework for reconstructing syntax and superior to frameworks denying the fundamental status of form–meaning correspondences, as reported on above.

The transmission problem, constituting the fifth argument against syntactic reconstruction, involves the alleged lack of continuous transmission of syntactic structures during acquisition. This means that language learners do not “inherit” the grammar of the previous generation; rather, they must construct their own grammar on the basis of the previous generation’s output. This stands in stark contrast to lexical items which are assumed to be passed on from one generation to the next. We have argued against this view and pointed out that lexical items are also abstractions. That is, lexical items are signs which consist of abstract phonological features and meaning, and this pairing of form with meaning that language learners must go through in order to acquire lexical items is a cognitive process in all and every respect comparable to the process of extracting the grammar from the input found in the linguistic environment. In keeping with the tenets of Construction Grammar, we thus reject the dichotomy between lexical items and grammar, and maintain that this is not an obstacle to the reconstruction of syntax.

Moreover, the goal of syntactic reconstruction is not in and of itself to reconstruct a proto-language as such, but rather to model the grammar of a proto-stage. Such a model is intended to reflect the state-of-the-art of our knowledge at each time. A reconstruction that is confined to phonology, morphology and the lexicon will not yield a complete grammar for the relevant proto-stage. This makes syntactic reconstruction a legitimate domain in historical syntax in particular and in historical linguistics in general.

One major goal with syntactic reconstruction is to identify the mechanisms of syntactic change, as has been emphasized by Goldbach & Ferraresi (2008), among others. In addition, we have argued that syntactic reconstruction can be seen as an important part of identifying the development of certain syntactic structures in order to reveal their origin (Barðdal & Eythórsson 2012a). This may result in reconstruction on the basis of identical elements in the correspondence set, the kind of reconstruction which has been dismissed by Lightfoot (2002a) as being trivial. Nevertheless, when investigating certain aspects of, say, Modern English, it is natural that one reverts back to Middle English for an historical explanation. And, likewise, when investigating the same aspects in Middle English, it appears as self-evident to explore the development from Old English. Reconstructing a proto-language enables one to take the investigation one or even more steps further back in time to a period predating existing written sources. In that sense, syntactic reconstruction of a proto-stage is no more trivial than investigating Old English when searching for answers on linguistic developments in Middle and Modern English.

At this juncture, let us address the question of whether there may be even further reasons for denying the validity of syntactic reconstruction. One possibility would be that scholars are blinded by their own theoretical paradigm, such as the generative framework, where semantics only has a very limited role, if any. Instead, syntax is an autonomous component in the model, where interaction with the phonological and semantic components is confined to the “interfaces” (e.g. Chomsky 1993, and especially Chomsky 1986: 363 for the claim that “syntax is semantics”). On such an approach, it becomes almost impossible to imagine how to reconstruct syntax, since the guidelines otherwise provided by semantics, morphology and phonology are unavailable. Such a theoretical framework may therefore become a straightjacket hampering one’s ability to confront the fact that syntax can be reconstructed.

Consider again Walkden’s reconstruction of the “middle” for Proto-Northwest Germanic. Walkden–used the phonology and morphology of the reflexive pronoun and its
distributional properties in the clause, in addition to the semantic content of the reflexive itself, in his comparison of North Germanic with its West and East Germanic sister languages. This Walkden did in order to reconstruct the reflexive middle for Proto-Northwest Germanic. Now, even if another element, a non-reflexive *sík*, had existed in Northwest Germanic, it would not have been used as a comparandum, because it would have been another morpheme with a different function from the ones in East and West Germanic. This is self-evident, of course, but only because Walkden’s analysis is indeed based on both form and function. However, even so, there is no designated place for either the phonology or the semantics in Walkden’s formalism, in reality denying the guiding role that these properties of language have, when determining the comparandum of the correspondence set in North and West Germanic.

The same criticism can be directed at Willis’s (2011) reconstruction of free relative clauses for Common Brythonic. His analysis is based on the relative marker and the function of the clause as a free relative clause, in addition to other factors such as negation and stress. In essence, this means that both form and function are used when establishing the comparanda for a particular correspondence set. Once again we see the problem of there being no designated place in the formalism for function or intonation, and as a consequence there is a failure to acknowledge the guiding role such factors have in determining the comparandum of the correspondence set for syntactic reconstruction. In this way, generative reconstructions in syntax are also incomplete.

5. Conclusions

We began this article with an overview of current work on syntactic reconstruction, summarizing recent work within the traditional Indo-Europeanist paradigm, the generative paradigm, and the Construction Grammar paradigm. The syntactic reconstruction carried out by traditional Indo-European scholars is generally solid and based on rigorous philological analysis of the early texts, applying the tried and tested techniques of comparative linguistics. The goal seems to be to decide on whether specific syntactic structures are reconstructable for a proto-stage or not. These reconstructions, however, are incomplete in one important sense: the traditional Indo-European scholars do not continue with their analysis beyond the point of deciding on the issue of reconstructability. That is, no actual reconstruction is carried out (although some draw up a diachronic scenario). This is presumably a consequence of the traditional Indo-European scholars’ lack of explicit theoretical framework and thereby of the necessarily representational formalism. Among historical syntacticians working within the generative framework, syntactic reconstruction has also been on the agenda in recent times. This work is based on sound historical syntactic research and modeled within the classical tree structure of generative grammar. The disadvantage of employing this representational model, however, is that there is no place in the formalism for semantics or function, discourse-pragmatic properties, or intonation. So similarly to the traditionalists, the generative attempts at syntactic reconstruction are also incomplete. Certainly not because of a lack of theoretical framework but rather since the formalism is a decisive factor in obtaining a fully accurate and a theoretically adequate reconstruction.

Syntactic reconstruction has also been carried out within the Construction Grammar framework, in which it is assumed that the basic building blocks of language are constructions – form–meaning pairings. Since the Comparative Method presupposes form–meaning correspondences, the leap from a synchronic analysis to syntactic reconstruction is minimal. The Construction Grammar formalism allows for the reconstruction of form, meaning, discourse-pragmatic properties, and intonational patterns, in addition to the reconstruction of higher-level schematic constructions. As such, Construction Grammar is the ideal framework for carrying out syntactic reconstruction.

To conclude, syntactic reconstruction, the long-dormant Sleeping Beauty of historical linguistics, has now been awoken from her sleep that had lasted more than half a century and has been given her deserved seat of honor within historical linguistics. We look forward
to witnessing further breakthroughs and successes in this important research area in the years to come.

References:


Rask, Rasmus Kristian. 1818. Undersøgelse om det gamle Nordiske eller Islandske Sprogs


