Reconstructing Syntax

EDITED BY

Jóhanna Barðdal, Spike Gildea,
and Eugenio R. Luján

BRILL

Reconstructing Syntax

Brill's Studies in Historical Linguistics

Reconstructing Syntax

Edited by

Jóhanna Barðdal
Spike Gildea
Eugenio R. Luján

Contents

Figures and Tables

Figures

Tables

Notes on Contributors

Jóhanna Barðdal
Ghent University

Flávia de Castro Alves
Universidade de Brasília

Thórhallur Eythórsson
University of Iceland

Spike Gildea
University of Oregon, Collegium de Lyon

Kikusawa Ritsuko
National Museum of Ethnology, Japan, The Graduate University for Advanced
Studies, Japan

Leonid Kulikov
Ghent University

Nikolaos Lavidas
National and Kapodistrian University of Athens

Ángel López Chala
Universidad Complutense de Madrid (Spain)

Eugenio R. Luján
Universidad Complutense de Madrid (Spain)

Silvia Luraghi
University of Pavia

Sara Pacchiarotti
Ghent University

Na'ama Pat-El
University of Texas, Austin

The Curious Case of Reconstruction in Syntax

Spike Gildea, Eugenio R. Luján and Jóhanna Barðdal

Abstract

The general consensus in the historical linguistics community for the last half a century or so has been that syntactic reconstruction is a bootless and unsuccessful venture. However, this view has slowly but steadily been changing among historical linguists, typologists, and anthropological linguists alike. More and more syntactic reconstructions are being published by respectable and virtuous publication venues. The debate on the viability of syntactic reconstruction, however, continues, and issues like i) lack of cognates, ii) lack of arbitrariness in syntax, iii) lack of directionality in syntactic change, iv) lack of continuous transmission from one generation to the next, and v) lack of form–meaning correspondences have, drop by drop, been argued not to be problematic for syntactic reconstruction. The present volume contributes to two of these issues in detail; first the issue of reliably identifying cognates in syntax and second, the issue of directionality in syntactic change. A systematic program is suggested for identifying cognates in syntax, which by definition is a different enterprise from identifying cognates in phonology or morphology. Examples are given from several different language families: Indo-European, Semitic, Austronesian, Jê, Cariban, and Chibchan. Regarding the issue of directionality for syntactic reconstruction, most of the studies in this volume also demonstrate how local directionality may be identified with the aid of different types of morphosyntactic flags, particularly showcased with examples from Chibchan, Semitic, and various Indo-European languages.

1 Why a Volume on Reconstructing Syntax

In the field of historical-comparative linguistics, the enterprise of reconstruction has first and foremost been focused on lexical, morphological and phonological comparisons, while syntactic reconstruction has been regarded as bootless, frowned upon and even lambasted (cf. Watkins 1964; Jeffers 1976; Lightfoot 1979, 2006; Harrison 2003; Pires & Thomason 2008; von Mengden 2008, *inter alia*). The rationale behind this position stems from the view that syntactic units are radically different from, for instance, morphological and lexical units. There is, of course, no doubt that syntactic structures are

larger and more complex than most morphological structures and phonological units, and at first sight they may certainly appear as ontologically different from word units. In other words, while morphological and lexical units are considered as having their own meaning, the traditional view of syntax is that it is not inherently meaningful. Rather, the meaning of sentences is more often than not taken to be the sum of the meaning of the lexical parts that instantiate them (cf. Klein 2010). Due to this semantic compositionality, reconstruction based on arbitrary form–meaning correspondences has been systematically ruled out. Further arguments against syntactic reconstruction are that syntactic structures are not inherited in the same manner as the vocabulary is (Lightfoot 1979, *inter alia*), that cognates cannot be identified in syntax (Jeffers 1976; at least with the same confidence as in the lexicon, Walkden 2009, 2013), that syntactic change is not regular, as opposed to the regularity of phonological change (Lightfoot 2002a; Pires & Thomason 2008), and that there is no arbitrariness found in syntax (Harrison 2003). The consensus based on these objections is, and has been, that syntactic reconstruction is either ontologically impossible or at the least substantially less certain than phonological and morphological reconstruction.

However, alongside these attacks, research exists within historical-comparative syntax suggesting that these critiques are overly pessimistic. To begin with, the arbitrariness requirement is simply not needed in syntax, as its primary function within the Comparative Method is to determine genetic relatedness; thus, if one limits syntactic reconstruction to languages known to be genetically related, this critique becomes orthogonal (Harrison 2003; Barðdal & Eythórsson 2012a; Barðdal 2013, 2014). Turning to the issue of cognacy, a growing number of studies within multiple theoretical perspectives identify cognates also in syntax, including Harris (1985, 2008), Campbell (1998), Bowern (2008), Barðdal & Eythórsson (2012a–b), Willis (2011), Dunn et al. (2017), Pooth et al. (2019), and many more studies cited in Section 3.4.

With regard to regularity of syntactic change, there are multiple considerations to take into account. First, we question the accuracy of the absolutist rhetoric surrounding the alleged regularity of sound change: in fact, there are many less regular sound changes that have simply been excluded from the definition of sound laws (Hoenigswald 1978); for example, no relevance has been accorded to irregular phonetic reduction or erosion associated with grammaticalization processes (Heine & Kuteva 2004: 2–3, 9). Also, the primary function of sound laws in historical linguistics is that they stand in for a similarity metric when deciding upon cognate status (Harrison 2003). However, historical linguists have always relied heavily on context, rather than on absolute sound laws, in determining which specific sound changes are more and less plausible

for each specific reconstruction (a point made nicely by Seržant 2015: 121). The same should be true of syntax, with directionality of change not being required to follow from absolutist principles, but rather being extrapolated from attested changes to analogous constructions in analogous contexts.

Under the assumption that syntactic reconstruction is indeed possible, two of the three co-editors of this volume proposed a workshop at the 2011 International Conference of Historical Linguistics, entitled Reconstructing Syntax. The goal of this workshop was to address fundamental issues of reconstruction in general and syntactic reconstruction in particular, either via individual case studies of syntactic reconstruction from different languages and language families or via a comparison of how different theoretical frameworks might contribute to improving the methodology for syntactic reconstruction. The workshop attracted a relatively balanced mix of theoretical papers, case studies from Indo-European, and case studies from other parts of the world. Inspired by that workshop, this volume focuses on two primary topics: (i) how to identify syntactic cognates, and (ii) departing from attested syntactic cognates, how to identify the most likely direction of change, which in turn allows for reconstruction from these cognates to the source syntax in a given proto-language. In addition to this introduction, this volume contains four papers presented at that workshop, supplemented by four more papers collected or commissioned afterwards.

The remainder of this chapter is organized as follows. In Section 2, we review some of the problems – beyond identification of cognates and determination of directionality – that have been claimed as obstacles to the reconstruction of syntax. Section 3 surveys several approaches, both early and more current, to syntactic reconstruction. From this background, Section 4 explicitly addresses the viability of establishing, in the domain of syntax, what we believe to be the two critical foundations of comparative reconstruction: identification of cognates (Section 4.1) and determination of direction of change (Section 4.2). Finally, in Section 5, we introduce the chapters of this volume and highlight how each contributes to our understanding of one or both of these primary issues.

2   Some Perceived Limits and Problems of Syntactic Reconstruction

We begin with the observation that, in many general overviews of the comparative method and in comprehensive presentations of proto-languages, no reference at all is found to the possibility of syntactic reconstruction. If, as a representative case study, one probes into the more than two centuries old

tradition of scholarship on Proto-Indo-European, one notices that the attention paid to syntax is scanty in comparison to the great quantity of studies devoted to investigating other aspects of the proto-language. Among the recent presentations of Proto-Indo-European, Szemerényi (1996) and Beekes (2011) do not even have chapters on syntax, while Fortson (2004) devotes significantly less space to syntax than to phonology or morphology, without explicitly discussing the methodology of syntactic reconstruction. Similarly, in his chapter on methods in reconstruction in a recent handbook on Indo-European, Krasukhin (2017) deals with phonological, morphological, and lexical reconstruction, but there is no section on syntactic reconstruction. In contrast, Bowern (2017: 4, 7), in the first chapter of the same handbook, explicitly acknowledges that the comparative method can be applied to syntax and systematic correspondences in various domains (including syntax) are needed in order to prove genetic relatedness among languages.

Even though, as laid out by Harris & Campbell (1995: 16–35), the Neogrammarians did devote some attention to reconstruction of syntax alongside morphology and phonology, in the words of Barðdal & Gildea (2015: 3):

> … there was a major difference in the results of [the neogrammarians'] work, in that for phonology they proceeded to develop and rigorously test a methodology for reconstruction that has remained largely unchanged as the modern Comparative Method. In contrast, for morphosyntax, they did not consolidate their individual works into a coherent, consistent methodology, and so they did not produce similarly large-scale reconstructions of PIE [Proto-Indo-European] syntax.

On the most conservative approach, one might try to reconstruct actual utterances or strings of words that the speakers of a proto-language might have used at a certain point. However, a program of that kind would be doomed to failure: scholars like Keydana (2018: 2195) state that after 200 years of comparative work on Indo-European languages, probably no other phrase can be reconstructed with certainty for Proto-Indo-European (PIE) than VP *$eg^{wh}ent$ $og^{wh}im$ 'he killed (the) snake' (Watkins 1995: 301). Although to this list one might add the well-known formulaic NP *$\acute{k}lewos$ $\mathring{n}d^hg^{wh}itom$ 'unperishable glory' (first pointed out by Kuhn 1853: 467) and perhaps others (for efforts, cf. Clackson 2007: 180–184, with further references), this type of reconstruction would lack much linguistic interest. This is because a grammar, whether diachronic or synchronic, is not a collection of sentences, but a nuanced description of the patterns and rules that characterize the utterances in a language. Daniels (2020) makes this point cogently, framing the task of reconstruction as

one suited for *types* as opposed to *tokens*, a perspective equally necessary for the reconstruction of lexicon and of syntax.

One criticism against syntactic reconstruction has to do with the fact that it is not a "first-order" reconstruction, as Clackson (2007: 158–159) formulates it. As framed by Clackson, in contrast to establishing sets of cognates in phonology, which by itself provides the basis for claiming genetic relationship between two languages, syntactic reconstruction would be a "second-order" operation that relies on previous knowledge of genetic relatedness and seeks to explain the data in the descendant languages starting from the reconstructed pattern in the proto-language, with special attention to archaisms and oldest attested forms. On the one hand, we do not dispute that the reconstruction of abstract syntactic patterns should not be the sole basis for claiming genetic relatedness between two languages and we find strong support for this position in the persuasive discussion in Dybo & Starostin (2008: 124–138).

On the other hand, it is far from a consensus position that the sole basis for claiming genetic relatedness between languages is via establishing lexical cognates with phonological correspondences. Many studies that address controversy in demonstrating genetic relatedness focus on the importance of identifying idiosyncratic morphological behavior of the sort that is unlikely to be borrowed, and that hence can be taken as evidence for shared retention from a common proto-language. In an acrimonious debate on precisely how to determine relatedness among the languages of the Americas, the importance of such evidence is about the only point on which Greenberg and his detractors agree, cf. Greenberg (1957: 37–38, 1987: 30, 1989: 108–109), Campbell (1988, 2003: 268–270), Rankin (1992); Dixon (1997: 22), Campbell & Poser (2008). Similarly, the most convincing argument for the Dene-Yeniseian hypothesis derives less from the limited number of lexical cognates identified so far, with their tentative phonological correspondences, but rather from the fact that "Shared Dene-Yeniseian morphological structures have been identified across every major lexical subsystem." (Vajda 2018: 281). Regardless, this issue is clearly orthogonal to the question of whether syntax can be reconstructed, since at its strongest, Clackson's concern merely restricts the enterprise of syntactic reconstruction to previously established genetic families.

We also take issue with Clackson's view of syntactic reconstruction as being a "second-order operation", i.e. as only existing in order to seek explanation for the data in the daughter languages, starting from the reconstructed pattern in the proto-language. Quite the reverse, this would involve "diachronic interpretation" in the sense of Klimov (1977), as opposed to a proper syntactic reconstruction of the linguistic situation in a proto-stage, based on the syntactic evidence from the daughters. In contrast to Clackson's view, Ferraresi &

Goldbach (2008: 9) point out that syntactic reconstruction has two main functions, namely i) to reconstruct a viable language supposedly spoken in the past, and ii) to better understand the general processes of language change. To this, Eythórsson & Barðdal (2016: 97) add a further goal, iii) to model the grammar of a proto-stage, adequately reflecting the state-of-the-art of the historical linguistics community's knowledge at each time. Otherwise, a reconstruction limited to a selection of the language system, in this case phonology, morphology and lexis, will not produce an accurate picture of the grammar as a whole for the relevant proto-stage.

Moreover, several studies of syntactic reconstruction have seen the light of day in recent years, none of which takes the situation in the proto-stage as its point of departure, aiming to explain the development from the proto-stage to the daughters. Instead, all these studies take the situation in the daughter languages as the point of departure and reconstruct syntactic constructions of the proto-stage, thus being first-order operations in Clackson's terminology, despite the fact that they involve syntactic reconstruction.

A further methodological critique against syntactic reconstruction raised in the literature has to do with the risk of misinterpreting syntactic variation, based on the assertion that syntactic variation can be due, with a significantly higher frequency than in other language domains, to style, register, or sociolinguistic level. For example, Hock (2000) draws attention to the impact of genres and literary styles on syntax, which might lead to the misanalysis of stylistic variation as diachronic variation. From this perspective, when reconstructing syntax, one should consider whether in proto-languages or in earlier unattested stages of languages, topicalization and fronting would have occurred as a way of emphasizing or highlighting a given element, beyond being mandatory in certain constructions (e.g. typically in relative clauses in many languages) or being characteristic of certain impressive modalities (see, among others, Watkins 1976; Hale 1987; Justus 1993, 2000 for Proto-Indo-European).

This criticism may be valid for both traditional and generative approaches to word order. On a constructional approach to syntactic reconstruction, however, criticism of this type is less compelling as, on a constructional view, variation in word order can never be a matter of simple frequencies nor of stylistic or register variation. Instead, variation in word order represents different constructions with different information-structural properties, as such constituting form-meaning pairings of their own, which are by definition the comparanda of the Comparative Method.

A successful syntactic reconstruction of the type described above, indeed carried out for Proto-Indo-European and involving fronting to clause initial position, is Barðdal et al.'s (2013) study on the adversative use of 'woe', based

on data occurring in five different archaic Indo-European daughter languages. Two different word orders are consistently found across the daughters, of which one is analyzed as representing neutral word order and the other as involving focus fronting, more specifically predicate fronting of focused material. As a consequence, Barðdal et al. (2013) reconstruct both neutral word order and a clause-initial focus construction for the relevant proto-stage, in this case, Proto-Indo-European.

A further problem for syntactic reconstruction, pointed out by Lightfoot (1979, *inter alia*), relates to the acquisition of syntax, as opposed to the acquisition of lexis. Lightfoot argues that children do not acquire syntax in the same way as they acquire words in that words are transmitted from one generation to the next, while syntactic rules are not. Instead, according to Lightfoot, syntactic rules are deduced by children on the basis of the grammar that they are exposed to. This means that, on Lightfoot's approach, there is a fundamental difference between words and syntax, which in turn excludes the reconstructability of syntactic structures.

We, as a matter of fact, call this lexis–syntax dichotomy into question. As repeatedly argued by Barðdal and Eythórsson (Barðdal & Eythórsson 2012a; Barðdal 2013, 2014; Eythórsson & Barðdal 2016), this view is a gross oversimplification of how children acquire words. Contrary to Lightfoot's view, words are also abstractions, i.e. they are form-meaning correspondences that are also deduced on the basis of the input children are exposed to (cf. Adger 2003; Tomasello 2003). Therefore, both in the case of acquisition of lexical material and syntactic objects, children have to abstract; either they have to build up their vocabulary, or they have to build up their grammar, in both cases on the basis of the input they are exposed to. There is no doubt, however, that on Lightfoot's conception of syntax as consisting of abstract rules, there is little basis for identifying analogy between lexis and syntax. On a construction grammar account, in contrast, the distinction between words and morphosyntax is not a matter of kind, but more a matter of degree (cf. Section 3.4 below).

A somewhat bigger-picture critique of syntactic reconstruction is that one cannot theoretically aim to reconstruct the whole grammar of a proto-language or of a non-attested, earlier stage of a language. Of course, one will never be able to determine the whole set of syntactic rules for any given proto-language – only a partial knowledge will ever be accessible through syntactic reconstruction. As articulated by Gildea (2002: 320), "when reconstructing grammar, one is limited to reconstructing morphology and syntactic systems on the basis of the surviving morphology and morphosyntactic systems. One can never be confident in reconstructing a complete picture of all the

grammatical resources of a proto-language because one never knows what pieces might have disappeared altogether, leaving no reflexes in any modern language." However, this is a practical limit based on what grammatical patterns have survived in daughter languages, not an *a priori* constraint on the types of grammatical patterns that may be reconstructed. Similar limits to phonological or morphological reconstructions have not prevented scholars from trying to gain insights into the past based on attested phonological and morphological systems; hence, the existence of limitations is no reason to eschew recovering also whatever elements of syntax may be reconstructed.

Hale (2007: 228) defines a reconstructed proto-language as a "set of all (chronologically) anterior grammars which do not differ in recoverable features." This explicitly points out that the unique, specific, complete grammar of the proto-language may not be accessible through reconstruction, but it is possible to approximate it by retrieving a number of features through comparison between the attested grammars of the descendant languages. For example, given that syntagmatic relations are produced *in praesentia*, the occurrence of a given element may trigger, allow or prevent the presence of another element that is syntactically linked to it. On this view, reconstructed syntax will be much more abstract than reconstructed phonemes or meaningful sequences of phonemes, basically consisting only of combination rules between linguistic elements that may or must co-occur with each other, such as agreement, dependency marking, word order, or the like.

From this perspective, syntactic rules are abstract and thus should be fleshed out in perceptible morphological elements or linearity rules in discourse (i.e., word order); they would otherwise fail to surface in the actual linguistic productions and would, therefore, be impossible to grasp. On the one hand, we see this as an epistemological problem that follows from the mechanisms of theories that represent syntax in the form of deep structures, which are themselves abstractions from patterns in the data. It seems clear that one cannot reconstruct such abstractions, but that rather what must be done is to reconstruct the actual pronounceable morphosyntactic patterns; if one then wishes to posit abstract rules from these reconstructed patterns (using the same methodology as for synchronic analysis), there would be a comparable empirical basis for the analysis (Campbell 1987: 91; cf. also Gildea's 2000: 68 articulation of a concrete example of this issue, asserting the difficulty in reconstructing abstractions like "basic word order" or "order in deep structures"). Of course, one might step away from this problem altogether by utilizing theories whose representations are more closely tied to concrete, surface representations. From such a perspective, the compound notion of "morphosyntax",

however formalized (or even relatively unformalized) is readily reconstructable (cf. Section 3.4).

Despite the disagreements above, we largely agree with Clackson's (2007: 159) summary of the relevant requirements in a list of three prerequisites that must be met when attempting syntactic reconstruction – his list validates our concern in this volume, as (a–b) are prerequisites to establishing cognates and (c) is the basis of establishing directionality.

a)   Enough evidence of the construction must be available in the earliest phases of the languages as the basis for the reconstruction and language-specific developments must be ruled out before

b)   Instances of the construction in the sister languages must share one or more comparable element (word-order, agreement, morphology or lexical particles)

c)   The development from the reconstructed construction to the ones actually occurring in the daughter languages can be accounted for in terms of a well-known process of syntactic change

We now turn to a brief survey of approaches that have been taken to syntactic reconstruction, which in turn contextualize our position on resolving the problems of cognacy and directionality.

3   Approaches to Syntactic Reconstruction

We start with a concise recapitulation of the Neo-Grammarian view of syntactic reconstruction (Section 3.1), before proceeding to the typological approach (Section 3.2), the generative approach (Section 3.3) and finally concluding the discussion with a summary of the advantages of a constructional approach to syntactic reconstruction (Section 3.4).

3.1   *The Neo-Grammarian Approach to Syntactic Reconstruction*

The Neo-Grammarians were first and foremost focused on comparative work within the areas of phonology, morphology and lexis, developing a comparative methodology to be used for reconstruction involving data from these subfields (see the overview in Eythórsson & Barðdal 2016). Their proposed methodology has more or less withstood the test of time, well known today as the Comparative Method. This methodology involves several procedural steps, including i) identifying cognates, ii) setting up correspondence sets, iii) choosing between different alternants of the relevant correspondence sets (with the aid of the sound laws), and iv) putting forward reconstructed sounds,

morphological forms or words, based on the chosen alternants of the correspondence sets and the directionality of the assumed change.

Neither the Neo-Grammarians nor the structuralists for that matter had much to say on syntax or syntactic reconstruction. The reason is most likely their lack of adequate tools to deal with syntax and syntactic variation; or in other words, their lack of syntactic theory (Fox 1995: 104). However, the Neo-Grammarians certainly made some important observations on syntactic and morphosyntactic matters, even though their stringent methodology did not carry over to the field of syntax.

One of the major contributions to syntax at the time of the Neo-Grammarians is Wackernagel's (1892) insights concerning the placement of enclitics in the second position of the sentence. The same is true of Delbrück's (1878) work on the position of the verb in Vedic and the other early Indo-European languages. Both Delbrück (1907) and Havers (1911) also worked on the function of the morphological cases in the early Indo-European languages. Some comprehensive comparative work on the syntactic properties of mood and modal categories started with the work of Jolly (1872), Thurneysen (1885) and Delbrück (1893–1900), to mention only a few of the syntactic topics that the Neo-Grammarians were concerned with.

One topic, in particular, of interest to the Neo-Grammarians around the turn of the 19th and the 20th centuries, was the reconstruction of an ergative alignment system for Proto-Indo-European. Early scholars like Uhlenbeck (1901), Pedersen (1907) and Vaillant (1936) argued for such a reconstruction on the basis of what appears to be a split system with nouns lower in the animacy hierarchy, i.e., neuters, lacking a morphological opposition between nominative and accusative, with the only exception of Hittite, which has an ergative marker *-anz(a)* for neuters when they are the subject of a transitive verb. This early reconstruction of Proto-Indo-European being an ergative language was later challenged (Villar 1983; Rumsey 1987a, 1987b), preceding a new wave of scholarship where it has been proposed that Proto-Indo-European was a stative–active language (Gamkrelidze & Ivanov 1995; Lehmann 1993; Bauer 2000, *inter alia*). This debate is indeed still ongoing, cf. recent work by Willi (2018) and Pooth et al. (2019).

In sum, the Neo-Grammarian approach to syntactic reconstruction was that it was "essentially the study of the function of forms" (Penney 2000: 35). This position was based on the precedence of morphology over syntax and any syntactic reconstruction was dependent on an a priori morphological reconstruction. That is, after reconstructing a given morphological category, the next step was to determine which of the meanings or functions associated with those categories were present in the proto-language. In the domain of nominal

morphosyntax, this involves number, gender and, especially, case, while in the domain of verbal morphosyntax, person, number and, especially, voice, tense and aspect. The goal to uncover the meanings and/or functions of morphosyntactic categories was achieved by studying the uses of those categories in the older Indo-European languages, especially Greek and the Indo-Iranian languages (a modern statement of this position would be, e.g., Serzant 2015).

3.2 *Typological, Holistic Approaches to Syntactic Reconstruction*

A radical methodological turn took place in the 1970's when Lehmann (1974) made an attempt to apply Greenberg's (1966) findings on implicational universals of word order to syntactic reconstruction (see overview in Eythórsson & Barðdal 2016). Lehmann was primarily concerned with basic word order and its implications for the order of other elements in the clause. The main goal of this endeavor was to determine whether the proto-language was OV or VO. A number of studies following Lehmann include Friedrich (1975) and Miller (1975), also reconstructing basic word order for Proto-Indo-European.

However, Friedrich and Miller proposed a radically different reconstruction from Lehmann and from one another. Lehmann argued that Proto-Indo-European was an SOV language, Friedrich that it was an SVO language, and Miller that the basic word order was SOV, SVO and VSO. The reason that these reconstructions are 180 degrees different is due to differences in the material on the basis of which these reconstructions are made.

The flaws of holistic typological approaches to syntactic reconstruction, i.e. approaches based on implicational relations between properties, were immediately pointed out in an influential article by Watkins (1976), who expressed severe critique of the typological approaches to reconstruction of Proto-Indo-European syntax, carried out by Lehmann, Friedrich and Miller, going as far as labeling these efforts a "pseudo-problem". Watkins' criticism was later iterated by scholars like Jeffers (1976), Lightfoot (1979) and Winter (1984). Since then, the holistic typological approach to syntactic reconstruction, based on implicational relations between properties, has been more or less debunked by the scholarly community (Hale 1987; Mendoza 1998; Drinka 1999; Gildea 2000; Wichmann 2008; Barðdal & Eythórsson 2012a).

In connection to this, another interesting point of discussion is the use of "residues" or "relic constructions", i.e. archaisms, when reconstructing syntax. Irregular and synchronically unanalyzable patterns have played an important role in reconstructing morphology; similarly, one can detect irregular, marginal syntactic constructions that do not fit so well in the more general rules of a language at a given stage and might therefore be considered a relic of a previous stage at which they were regular. Campbell (1986: 81–86) has an insightful

discussion of this principle, with the examples of Balto-Finnic infinitives and participial subjects. Lehmann (1993, 1994) resorted frequently to alleged residues or relics of the original active structure of PIE that he reconstructed. However, Viti (2015) warns about either applying the anomaly principle or using consistency in the descendant languages as a basis for the reconstruction of a proto-language – both are problematic, she argues, due to the uncertain status of syntactic variation (cf. Section 2).

3.3   *Generative Approaches to Syntactic Reconstruction*

After the Neo-Grammarians and the structuralists, and in the era of modern linguistics, in particular generative linguistics, the focus of research has been on synchronic structures, with historical and comparative approaches to language being disfavored. The first important contribution to historical syntax from this perspective was Lightfoot's (1979) book where he explicitly rejects the possibility of syntactic reconstruction, given that in his view syntactic change is basically a matter of reanalysis, a process lacking any inherent directionality, which, in turn, prevents the recovery of any prior stages (1979: 154–166, 1980, 1999: 255–257, 2002a: 114–130). As we argue below (Section 4.2), directionality can indeed be detected for each case of reanalysis *per se*, debunking the claim (and calling into question the theoretical postulates that informed it).

More recent approaches to historical syntax within the generative paradigm have been positively inclined towards the possibility of reconstructing syntax (see Eythórsson & Barðdal 2016 for an overview). For instance, Hale (1987) and Garrett (1990) reconstruct the position of clitics in the hierarchical Proto-Indo-European clause structure, focusing in particular on the Wackernagel position. Roberts (2007: 363–367) also argues that parameters, which are conceptualized as formal features of lexical entries, can be taken as the comparable units required by the comparative method. Examples are Roberts' reconstructions of the null-subject parameter, the OV/VO basic word-order or *wh*-movement in interrogatives and relatives for Proto-Indo-European. No reanalysis would be needed on this approach.

More recently, scholars from within the generative paradigm, like Willis and Walkden, have successfully carried out syntactic reconstructions (Willis 2011; Walkden 2009, 2013, 2014). A particularly important contribution is Willis's (2011) research on the distinction between "universal directionality" and "local directionality". The former relies on extensive comparisons across languages that allow for identifying widespread tendencies in language change. The latter, in contrast, underlines the weight that should be accorded to the data under each analysis, as these undoubtedly call for specific interpretations, providing the basis for identifying the directionality of change, and thus contributing to

a syntactic reconstruction. Hence, Willis concludes, local directionality is all that is needed for identifying prior stages of languages, even if we never determine any kind of universal directionality.

One major problem for generative approaches to syntactic reconstruction, as pointed out by Eythórsson & Barðdal (2016), is that within the generative paradigm, there is no natural place for semantics, hence strictly speaking no reconstructions can take place on the basis of form–meaning correspondences within that framework. Generative scholars are forced to either base their reconstructions on form alone, or they may indeed take meaning or function into account, despite the lack of natural space for meaning or function in their formalisms. As such, generative scholars have to go beyond the limits of their own frameworks in order to contribute to syntactic reconstructions.

3.4   *Constructional Approaches to Syntactic Reconstruction*

Despite the pessimistic voices of the scholars of the 1970's, such as Watkins (1976), Jeffers (1976) and Lightfoot (1979, 1980) (see Section 3.2 above) a new generation of historical linguists emerged, more specifically, historical syntacticians, eager to address the challenges of historical syntax, including the more-or-less stranded endeavor of syntactic reconstruction. In particular, Harris (1985) and Harris & Campbell (1995) introduced into the field the notion of *syntactic pattern* and developed a rigorous research program based on this concept of how to reconstruct syntax. Gildea (1992, 1998, 2000) soon followed in their footsteps, putting forward a reconstruction of the Proto-Cariban verbal system (including multiple alignment properties), as well as reconstructing the nonverbal origins of six additional clause types found in the modern Cariban languages. Key to these reconstructions were the identification of the cognate structures that composed each innovative *clause type* plus providing local arguments for directionality of change, and thereby deducing the reconstructable source. Independent of this enterprise, Kikusawa (2002, 2003), working on alignment changes in Proto-Central-Pacific, launched the concept of *cognate structures*, applying it to basic word order constructions, as a part of a larger program of establishing correspondence sets in syntax. Through the efforts of these scholars, huge advances have been made in the methodology of reconstruction of syntax.

Inherent in the approaches of Harris, Campbell, Gildea and Kikusawa is the preconception of a construction, i.e. a form-meaning pairing, which indeed is the comparanda of the Comparative Method, as laid out by the Neo-Grammarians (see Section 3.1 above). It has also been pointed out by scholars working within the framework of Diachronic Construction Grammar, i.e. scholars who apply constructional analysis on historical data, that Construction Grammar is more

easily extendible to syntactic reconstruction than other linguistic models, due to the basic status of form-meaning pairings in that framework (cf. Eythórsson & Barðdal 2011, 2016; Barðdal & Eythórsson 2012a–b; Barðdal 2013, 2014; Barðdal & Gildea 2015; Daniels 2015, 2017, 2020).

That is, constructions are assumed to be the basic building blocks of language and are as such form-meaning pairings (Goldberg 1995, 2006; Croft 2001; Michaelis & Ruppenhofer 2001; Boas 2003; Fried & Östman 2005, *inter alia*). On a Construction Grammar approach, constructions exist at all levels of language, at the level of morpheme, word, phrase, as well as at the level of larger syntactic units like argument structures and clause structure, yet maintaining their status as form-meaning pairings. Constructions also range from being concrete lexically filled units, like words, to being partly lexically filled set phrases and idioms, to being almost or entirely schematic (Lakoff 1987; Fillmore, Kay & O'Connor 1988; Nunberg, Sag & Wasow 1994; Goldberg 1995; Jackendoff 1997, *inter alia*), but at all times maintaining their status as form-meaning pairings. In other words, the representational formalism of Construction Grammar, explicitly identifying and linking both form and meaning, can be equally well applied to all linguistic material. From this basic assumption of Construction Grammar, reconstructing syntax not only becomes a practical exercise in historical linguistics, but also a viable and a realistic undertaking when studying the history of languages.

In order to illustrate this point, consider the findings of Barðdal et al. (2013) where the linguistic history of *woe* in the Indo-European languages is investigated. This Indo-European adverb 'woe' builds a part of an argument structure construction found with a compositional predicate involving a dative subject and the verb 'be' confined to 3rd person singular. A comparison of the relevant data, stemming from five different Indo-European subbranches, Germanic, Baltic, Slavic, Italic and Indo-Iranian, reveals three different but clearly related constructions, i) 'woe', ii) 'woe'-DAT and iii) DAT-'is-woe'. The first construction is analyzed as having an exclamative function, while the third construction is analyzed as a predicative construction with the pragmatic function of expressing speaker's dismay (used first and foremost in situations of adversity). The second construction is analyzed as a 'be' less variant of the predicative construction, an analysis that requires an explanation of the difference in word order between the two.

Barðdal et al. (2013) reconstruct both the first construction, 'woe', and the second construction, 'woe'-DAT, as exclamative constructions, but the internal structure of the second is inherited from the predicative construction, DAT-'is-woe' – the verb 'be' is omitted and 'woe' occurs in first position due to

focus fronting of the predicate. This analysis is based on instances in the early daughter languages where focused material indeed occurs in first position, ultimately preceding the subject. As a result, Barðdal et al. (2013) reconstruct neutral word order for Proto-Indo-European (the relative position of the subject and the predicate), as well as a clause-initial focus position for that same language. This research was successfully carried out through a proper analysis of the meaning/function of the three constructions, formalized with the box formalism of Construction Grammar, which includes fields for both form and function.

Further syntactic reconstructions carried out via this formalism in Construction Grammar include the reconstruction of grammatical relations for Proto-Germanic (Barðdal & Eythórsson 2012b; Eythórsson & Barðdal 2016), different levels of schematicity of the ditransitive construction in Proto-Germanic (Vázquez-González & Barðdal 2019), a non-finite (gerundive) modal construction for Proto-Indo-European (Danesi et al. 2017), oblique subject constructions for Proto-Indo-European (Barðdal & Smitherman 2013; Barðdal et al. 2013; Johnson et al. 2019), as well as the reconstruction of a verb-class specific argument structure constructions for verbs of success in Proto-Indo-European and the conceptual metaphor motivating this argument structure in the minds of Proto-Indo-European speakers (Johnson et al. 2019).

We turn now to the questions that motivate this collection.

4 The Comparative Method and Syntactic Reconstruction

In this section, we take the position that the comparative method is equally adequate to the reconstruction of syntax as it is to the reconstruction of phonology and morphology; it is merely different because the domain of investigation has different properties. We suggest that the theoretical definition of cognates (i.e., "what they are") is independent of the operational definition (i.e., "how to know one when you see one"). While the nature of phonological change makes regular correspondences an obvious way to *identify* lexical cognates in data, such correspondences never have been (and should not be) a part of the theoretical definition of *what* a cognate is. Once this point is made clear, one can apply the same theoretical definition of cognacy to syntactic constructions and then, based on one's knowledge of how syntactic constructions change, search for the operational criteria that allow one to identify syntactic cognates, rather than mistakenly trying to apply operational criteria based on how form and meaning evolve in the process of diachronic lexical change.

4.1   *Identification of Cognates*

We begin with the observation that cognates, whether lexical or syntactic, exist because it is possible for linguistic structures (both formal and semantic) to remain consistent between generations, and thus be inherited from source structures that were spoken many generations ago. Such inheritance need not involve change, but as more time passes, change becomes the rule. From this observation we derive our theoretical definition of the notion "cognate", from which any operational definition must follow: two linguistic items are cognate if they descend by direct inheritance from a common ancestor. Defined in this way, the question becomes how one can identify cognates, and especially, distinguish them from units that resemble cognates but are not. Identification of cognates is a methodological question, guided by operationally defined criteria: by what methods can one confidently assert that any two linguistic structures do, in fact, descend from a common ancestor? In particular, can one identify linguistic properties that are arguably unique to direct inheritance, and which can thus serve as circumstantial evidence for that status?

The first line of argument has to do with the plausible causes of similarity between linguistic structures. Before turning to syntax, we begin with the lexicon, asserting that lexical items from different languages that are sufficiently similar in both form and meaning are potential cognates. There is general agreement in the field that such similarity can only be due to one of four situations:

a)   chance (the null hypothesis)
b)   extra-linguistic factors (e.g. iconicity, in which onomatopoeic words resemble an external sound)
c)   borrowing (received via contact)
d)   cognacy (inherited from the same proto-language)

The weaker the degree of similarity, the more likely it is to be due to chance, and thus the more important it becomes to have methodological tools to overcome the likelihood that weak similarities are due to chance. All criteria that argue for cognacy derive their force from how specific sorts of similarity are identified, especially those which strengthen the argument against chance. For example, the more potential cognates one finds the less likely it is that each one, individually, could have arisen by chance. Likewise, when one can further specify regular patterns of similarity internal to the formal and/or semantic components of potential cognates, the statistical plausibility of chance diminishes.[1]

---

1   Although it is a logical possibility, we are aware of no examples of regularities in meaning change that have been used as primary arguments for lexical cognacy. The only role we have seen for semantic change is when scholars argue (apparently on an intuitive basis) about whether two given meanings are "similar enough" to be potential cognates.

Once a similarity is too great to sustain the null hypothesis, then the task turns to identifying the type of common origin responsible for the similarity.

A non-historical explanation for similarity would be when potential lexical cognates have a common origin in that they have an iconic resemblance to something external to language, such as machine or animal noises. Such non-arbitrary forms are usually excluded from the set of cognates. Beyond these sorts of cases, it is generally accepted that for most of the lexicon the relation between sound and meaning is arbitrary. This means that when chance and factors like iconicity are discarded, the only account remaining is that the similarity must be due to a set of lexical items having their origin in a common source, whether transmitted by inheritance (cognates) or by contact (borrowings).

To begin with the simple case, if a lexical item is borrowed from an unrelated language, there is little likelihood that it will be similar enough in form to an inherited lexical item to be considered a potential cognate; this sort of borrowing merely reduces the number of potential cognates to be found. The more serious problem comes when lexical items are borrowed from a related language (cf. Bowern 2008: 199–200). In this situation, attested reflexes actually are descended from a common source, but via different pathways: inherited via a direct line of descent from the proto-language vs. inherited via a detour through another language/branch of the family. In a way, both of these would satisfy one component of the theoretical definition of cognates in that both trace back to the same proto-form; one might call these different subtypes "inherited cognates" and "borrowed cognates". On the one hand, this difference would not necessarily challenge our ability to reconstruct a shared proto-form, but it would call into question the pathways via which the shared proto-form arrived into each attested reflex. On the other hand, to the extent that sound changes may have differed along different pathways, it becomes possible to use such differences to distinguish "inherited cognates" from "borrowed cognates", which should allow a more reliable reconstruction of the sounds of the proto-form.

As such, the independent identification of the kinds of sound changes undergone by phonemes internal to potential cognates is a crucial tool both for arguing against chance and for distinguishing inherited from borrowed cognates. To give an example of the validity of statistically increased similarity, when comparing many potential lexical cognates, if one observes multiple cases in which a given sound (say /p/) in Language 1 corresponds to a specific sound (say /f/) in the same position in each potential cognate word in Language 2, such a structured similarity statistically decreases the plausibility of the null hypothesis. When every sound in each potential cognate in

Language 1 corresponds to another specific sound in the potential cognate in Language 2, and when this holds over a large number of potential cognates, unless there is reason to suspect that there has been a massive relexification event (something akin to creole formation), even the most hardened skeptic is forced to concede a common origin, i.e., that the relevant words are indeed inherited cognates.

However, because of the widespread acceptance of the Regularity Hypothesis (Osthoff & Brugman 1878), if such regular phonological correspondences are *not* found in any potential cognates, the Comparative Method obliges one to assume that the potential cognate did not descend via the same pathway, i.e., that the word with irregular correspondences must be similar due to chance or due to having been borrowed from a related language. Having identified the unexpected sound changes, one can then attempt to identify the donor language as one for which the non-conforming sound changes would have been expected. This method, the identification of regular phonological correspondences (what Walkden 2013: 101 calls the "Double Cognacy Condition") is the gold standard for establishing cognacy between lexical items.

Nevertheless, it is important to be clear that inheritance from a common ancestor is logically distinct from the creation of consistent internal correspondences in either form or meaning. That is, even in the domain of lexical comparison, the identification of correspondences is not the characteristic that *defines* a cognate; rather, it is the characteristic that effectively rules out alternative hypotheses, thus ending debate over *whether* two lexical items are cognate. In this sense, it is a logical error to require the equivalent of "regular correspondences" when seeking to identify cognates in other domains of language, such as syntax. The error is that this requirement entails setting aside the theoretical definition (cognates are linguistic structures inherited from a common source) and substituting for it a particular criterion (regular phonological correspondences/the Double Cognacy Condition) that is derived from a single domain of historical change (phonological change in the lexicon). Since such phonological correspondences are an outcome specific to processes of phonological change in lexical items, there is no reason to expect something identical in other domains of historical change, such as syntax.

This domain error seems to be at the root of the incongruous theoretical claim that cognates cannot be identified in syntax. Based on vast amounts of scholarship, it is widely agreed that the absence of regular phonological correspondences between suspiciously similar pairs of words is sufficient to conclude that they are not (inherited) cognates. *Mutatis mutandis*, the impossibility of identifying similar regular correspondences in potential syntactic cognates obliges the careful scholar to conclude that they, too, are not cognate.

However, syntactic constructions arise and change via mechanisms unlike those that create regular phonological correspondences, so an exact analog to such correspondences will never be found in syntax. As such, it is almost tautological to conclude that it is impossible to identify cognates in syntax and therefore that one should reject claims by those who believe they have done so. When this false equivalency is laid bare, it becomes clear that the lack of such correspondences justifies only the much more limited conclusion that no potential syntactic cognates can qualify as *fully lexical* cognates.

Walkden (2013) makes a valiant effort to expand the validity from phonology and lexicon to syntax by reframing the requirement in more general terms as his "Double Cognacy Condition". This condition characterizes both the phonological correspondences inside lexical items and the lexical items themselves as "cognate", meaning each is independently inherited from a distinct proto-unit. Framed this way, it is the identification of the embedded "double cognate" phonemes inside each potential cognate lexical item that makes it possible to confirm them as cognates. Such "double cognates" are not found in syntax, he argues, because the combinations of units in syntactic constructions contain slots that are not fully determined, which means that they cannot be inherited by children in the same way as lexical items. Unlike words and sounds, which he claims are inherited directly from prior words and sounds, sentences are not inherited directly from prior sentences – rather, they are generated from abstract syntactic rules.

In Section 2 above we have discussed challenges to these empirical claims about the "direct" inheritance of lexical cognates. Here we add that, despite its more general phrasing, the Double Cognacy Condition is nothing more than a restatement of earlier attempts to import into the domain of syntactic change a method that was developed based on the process of regular phonological change in lexical items. It is time to let go of the bias that follows from trying to assess the likelihood of syntactic cognates using a method derived from phonological change in lexical cognates. Obviously one needs methods to assess the strength of potential syntactic cognates, but this should be done on the basis of our understanding of syntax and syntactic change, not other sorts of change.

In this regard, no matter how one interprets it theoretically, historical linguistics has an obligation to model the empirical fact that it is possible for syntactic constructions to remain consistent through multiple generations of speakers. Our methods must have a way to recognize and model such diachronic consistency before our theories can reasonably be expected to explain it. While it may be acceptable to claim that intergenerational transfer is different in the domains of phonology, lexicon, and syntax, it is not reasonable to

insist that intergenerational consistency in phonology and lexicon constitute evidence for direct transfer whereas parallel consistency in syntax does not, as though such consistency in the domain of syntax might be some kind of historical accident. Historical syntacticians, too, would like a fully satisfying cognitive explanation for intergenerational consistency in all domains of language, but at the moment no explanation is in sight that is fully satisfying for any domain.

In the meantime, there is no reason to postpone the study of historical change in syntax while one awaits additional theoretical clarity. It is indeed possible to take what is already known about the facts of intergenerational continuity and discontinuity in syntactic patterns and use this knowledge to identify patterns that are both continuous and discontinuous. Given the theoretical definition of the term "cognate", it is self-evident that this is the appropriate label for constructions or syntactic patterns that do show intergenerational continuity.

As such, the task now is to elaborate a method by which to examine potential cognate constructions, a method that identifies types and degrees of similarity that could not plausibly arise by chance. Having eliminated the null hypothesis, one should also identify criteria that would aid in distinguishing between similarity due to contact (whether direct borrowing, calquing, or pattern replication) or due to language external factors (e.g., iconicity), as opposed to simple inheritance. Until we have elaborated such a method, it is premature to make any claims about the degree of confidence that we should place in these methods. This must be an empirical question, which will eventually need to be tested statistically (along the lines of the calculations made by Ringe 1992).

The first key to identifying cognates is to identify the point at which chance becomes an unreasonable explanation. Obviously, the strongest case against chance will be found when there is identity between constructions; this is the situation in which even the most skeptical scholars already accept that constructions are cognate (e.g., Lightfoot 2002a: 120; Walkden 2013: 98, 107). However, most of what is interesting about historical syntax involves the kinds of changes that take place, whether to entire constructions (in cases of reanalysis) or to the sub-components within constructions (cases of analogical extension, expansion of items that can occur in schematic slots, etc., cf. Section 4.2). As changes accrete to one or another of the attested reflexes of a proto-construction they will become increasingly different from one another, and as these differences become greater the intuitions of experts will reach less consensus about why the relevant reflexes are similar (i.e., whether they are cognate). For this reason, when potential cognate constructions are similar,

but not identical, one needs methods to probe inside the larger constructions for specific properties that can be used as arguments for or against the null hypothesis.

From a Construction Grammar perspective (consistent with the synchronic approaches in, e.g., Goldberg 1995, 2006; Croft 2001; articulated diachronically in, e.g., Traugott & Trousdale 2013; Barðdal & Gildea 2015, *inter alia*), a construction is a combination of morphemes, more schematic "slots" (to be filled by lexical items selected from specific categories), and the syntactic relations between these elements. While it is not absolutely necessary for our argument here, it is the case that the meaning and/or distribution of a construction is not necessarily predictable by combining the meanings and/or distributions of its component subunits. That is, a construction as a whole can have its own meaning beyond the combinatory meanings of the elements within it. In this way, constructions resemble lexical items in being a combination of form (which contains consistent internal structure) and meaning (which may be derivable from the meanings of the subunit forms, but may also be arbitrary). What is crucial for identifying cognates is that the formal component of potential cognate constructions will contain multiple subunits that can be independently determined to be cognate: entire lexical items, bound morphology, constituent relations, etc. As the number of cognate subunits increases, the plausibility of the null hypothesis decreases. This same argument has been used successfully in frameworks that invoke theoretically less developed notions akin to construction, such as cognate "patterns" (Harris 1985, 2008; Harris & Campbell 1995), cognate "verbal systems" (Gildea 1998), "cognate constructions" (in a pretheoretical sense; Gildea 2000), "cognate structures" (Kikusawa 2002, 2003).

As a programmatic suggestion, we list here a few of the most obvious candidates to anchor arguments against the null hypothesis when comparing potential syntactic cognates:[2]

a) The presence and location of cognate morphemes (whether free or bound)
b) The presence and location of schematic slots of specifiable types
c) Identity in linear order and/or constituency relations between cognate morphemes and/or schematic slots

---

2 See Seržant (2015: 125–130) for a somewhat overlapping methodological discussion, specifically what he labels *morphological profile, lexical profile, syntactic profile*, and *semantic profile*. While Seržant's methodological discussion does not characterize types of properties as we do here, most of these properties do appear in the detailed example of the independent partitive genitive that he provides later in that article. Of course, from our more constructional view, we do not compartmentalize such elements into distinct *profiles*, preferring to maintain a holistic view of all components internal to each construction being examined.

d) Identity in other systematic relations amongst internal elements, e.g. argument structure patterns like case-marking, verb agreement, etc.

e) Either identical semantic values or a relation between the semantic values that is consistent with attested reanalyses in syntactic constructions (e.g. constructions containing a resultative participle > perfective aspect > past tense)

Unlike the phonemes that are sub-units in cognate lexical items, there are more varied types of subunits within syntactic constructions. Fixed morphemes are internally complex, being themselves full signs with structured form linked to meaning, capable of being independently identified as internal cognates within the larger cognate construction. Alongside fixed morphemes are schematic slots, some more and some less specified with regard to the sets of forms that can fill them. Even though such slots are less concrete than lexical items, they have both formal and semantic components that may be consistent between generations: formally, the locations of slots relative to each other and to fixed morphemes may themselves form fixed patterns or constituents; semantically, the list of elements (or more abstract properties that characterize such lists) could either be consistent or change over time. More abstract syntactic properties may also be consistent or subject to change, such as: phrase structure relations; overt coding properties of verbal argument structure like case-marking, verb agreement, and constituent order; covert properties of argument structure like control of coreference with reflexive morphemes or elided arguments in complex clauses; etc.

As we observe historical change in syntax, the subcomponents of syntactic structures have more autonomy than mere phonemes, able to change meaning in cases of reanalysis and to change form in cases of analogical extension. Despite these internal changes, the overall construction can retain a high degree of similarity with the source construction, similarity that cannot be explained as anything other than pattern consistency maintained between generations – that is, inheritance. Given the relative freedom to change what characterizes subcomponents of syntactic structures, most potential cognate constructions will not be identical. Examining the properties listed in (a–e) above gives us a way to consider both the similarities and the differences between them.

Critical to this perspective is that the existence of differences does not, in itself, invalidate the hypothesis of common origin – even if not identical, any degree of similarity too great to attribute to chance requires an account, whether due to contact, external factors (such as, in the domain of syntax, functionally-motivated typological patterns, as explicated in Seržant 2015 and

Daniels 2015, 2020), direct inheritance from a proto-source, or some combination of the three. At the same time, even if one accepts the plausibility of shared origin one cannot simply dismiss the relevance of differences – ultimately, a satisfying reconstruction should identify both a source construction and a mechanism of change that could plausibly create each difference by changing elements of one (or each) sister construction after separation. In order to properly address the changes, we turn to the second major issue of this volume: how one can construct valid arguments to decide between competing hypotheses about directionality of change.

4.2   *Determining Directionality of Syntactic Change*

We begin our discussion of directionality in syntactic change by pointing out that in most cases, the kinds of syntactic change that are reconstructable are not global modifications of abstract phenomena like order of core constituents, but are rather the changes seen in more concrete, local constructions. For instance, Jeffers (1976) contrasts phonological reconstruction, where there is a basis for positing the expected evolution, and syntactic change, where he claims that there is no such basis, using the example of global word order: if two languages of the same family show different basic word-order (e.g., SVO vs. SOV), one cannot assess which was the original one or if they both come, in fact, from a different alternative.

However, once a specific set of cognate constructions has been identified, one faces a more limited problem, namely that of how to determine what changes might have created the differences, so that one can reconstruct the source construction from which these changes would have followed (for case studies on construction-specific word order change, cf. Claudi 1994; Gildea 2000; Kikusawa 2002; Barðdal et al. 2013). For example, given a set of cognate constructions, one can reconstruct one or another of the documented constructions as being identical to the source, with the others having changed to get to their attested forms, or one can reconstruct an unattested construction, such that all the documented constructions have undergone some change to arrive at their attested forms. In this way, it may be possible to narrow the problem of determining directionality to identifying what evidence allows for making a choice between specific competing hypotheses (called the 'pool of variants' in Vincent & Roberts 1999; Roberts 2007: 362, 367–368). In this specific task, the following more concrete considerations may function as guidelines:

a)   Identification of the most plausible **mechanisms of change**, which may be recognized based on specific synchronic patterns in the cognate constructions

b)   Within these mechanisms of change, identifying preponderance of evidence for the **direction of attested changes** in parallel situations (this is especially valuable in cases of construction reanalysis)

c)   In cases where the attested record does not provide any guidelines, a weaker metric is available in the form of more general principles that speak to the relative age of morphology found in constructions (Givón 2000: 120–121; Gildea 2002: 319–320)

d)   The weakest of all the metrics is parsimony, in which the directionality is determined based on which proto-form would require fewer changes to arrive at the attested forms

We turn now to each of these considerations in turn.

4.2.1   Mechanisms and Attested Direction of Change

The standard set of mechanisms of change has been identified in multiple works as **reanalysis, analogical extension,** and **borrowing** (cf. Harris & Campbell 1995; Hopper & Traugott 2003; Gildea 1998; Barðdal & Gildea 2015; and many other references contained in these sources). In very abbreviated terms, reanalysis (somewhat reconceptualized by Traugott & Trousdale 2013: 35–37 as **neoanalysis**) is change in the cognitive representation or conceptualization of a construction that does not lead to any immediate observable change in its formal properties. As such, reanalysis is a mechanism that leads to identical formal constructs with sometimes very different meanings, generally also with distinct functional and distributional profiles (cf. also Gildea 1998: 153–155); synchronically, it is not always clear to all analysts, especially those who focus on economy in structural representation, that immediate post-reanalysis constructions are different enough from their sources to merit distinct formal analyses.

In contrast, analogical extension (Hopper & Traugott's 2003 **analogy**; Traugott & Trousdale's 2013: 37–38 **analogization**) is change in individual elements or patterns in the formal component of a construction, for example, a change in which a morpheme or collocational pattern from an independent construction in the language is "borrowed" into the innovating construction. Analogical extension is thus a mechanism that leads to irregular changes of individual morphemes or syntactic patterns, creating visible distinctions in individual cognate constructions. While analogical change also creates problems in the domain of phonology (especially morphophonological change), it is the prevalence of analogical change that makes syntactic cognates so different from lexical cognates, resulting in situations where constituent components differ between sister constructions that come from a common origin. Given this characterization of analogical extension, it is clear that the mechanism of

borrowing is actually another manifestation of analogical extension, in which the independent construction that provides the source component is found in another language.

The relevance of these two mechanisms for determining directionality in reconstruction is particularly addressed in several works by Gildea (1998: 35–44; 2002: 316–320; 2008: 67–71), to which we refer readers for additional detail. When inspecting many examples of directionality in attested historical change, it is clear that reanalyses overwhelmingly proceed in a single direction, primarily by reducing complexity in innovative constructions. For example, there is a multitude of examples of biclausal constructions being reanalyzed as monoclausal, with erstwhile main clause verbs becoming auxiliaries and then inflections; in this process, subordinate inflections/derivations and other elements of subordinate clause grammar (e.g. argument structure) are introduced into main clauses. We have yet to identify any examples of the opposite directionality, i.e. monoclausal constructions being reanalyzed as biclausal, with main clause verbal inflections becoming auxiliaries and then complement-taking verbs, and some other component of the main clause then being reanalyzed as a subordinating morpheme, thereby introducing main clause grammatical patterns (like argument structure) into subordinate clauses. This claim is not theoretical, but empirical, and as such is subject to falsification in databases of attested change. Further, while this claim draws on much of the data brought to bear in the sometimes heated debates about unidirectionality in grammaticalization, we claim no theoretical significance to this observation, merely the instrumental significance that it facilitates identifying the direction of change in those specific instances where the mechanism is clearly reanalysis.

In contrast, attested examples of analogical extension appear to be relatively unconstrained in directionality; for example, specific morphemes from main clauses readily extend into subordinate clauses, specific morphemes from subordinate clauses can equally well extend into main clauses, and morphemes can also move about inside paradigms, which in turn may affect the morphological inventory of a given construction. However, to the extent that one is able to first identify the reanalysis that gave birth to a new construction, and to the extent that this enables one to identify the full set of components in the source construction, it is possible to distinguish which differences between erstwhile cognate constructions represent conservative vs. innovative component elements. Having identified which components are innovative, it is possible to discern where these components occur outside of the construction of interest. In most cases the "donor" construction is readily identifiable based on parallels to the function that the innovative component serves in the construction of interest. So even though the mechanism of analogical extension

is not intrinsically directional, when the source construction and the donor construction can be identified on independent grounds, it becomes possible to identify the direction of change in specific cases of analogical extension.

With regard to borrowing, directionality is clear when the sources for specific morphemes can be traced to unrelated languages, or to related languages whose cognate morphology has undergone distinct phonological changes (cf. Daniels 2017: 303–304). However, when a borrowing takes the form of a calqued structure or a copied usage pattern (e.g. the "overuse" of a native passive construction when third person agents act on first or second person patients, a mechanism identified in Mithun 2006, 2012), the absence of borrowed morphology makes it more difficult to identify the effects of contact. Although contact effects of these kinds are inherently difficult to identify, to the extent that our purposes are limited to reconstructing individual constructions to a proto-language, and to the extent that copied usage patterns do, in fact, operate on constructions that were present in the proto-language, contact-induced change of this sort does not interfere with the identification of the proto-construction.

Despite the manifest differences between lexical/phonological and syntactic reconstruction, they share an important similarity: reconstructions are relatively straightforward when the time depth is shallow and they become increasingly difficult as the proto-language recedes farther into the past. In particular, it is easiest to reconstruct a case of recent reanalysis: (i) the source construction is often still attested, perhaps even in multiple languages; (ii) there have been enough changes in the reanalyzed reflexes to make it clear that the reanalysis has happened, that the modern reflexes are actually analytically distinct entities from their source; and (iii) there have not been so many individual changes in the reanalyzed reflexes as to obscure their shared inheritances from the source construction. As more time passes following the original reanalysis, (ii) becomes gradually more powerful and both (i) and (iii) less so; it is increasingly likely that the pre-reanalysed constructional source will no longer be attested and the accretion of individual analogical changes become themselves inherited patterns. Since the analogical changes are not inherently directional, it becomes difficult or impossible to argue confidently as to which formal elements are conservative (and hence reconstructable to the proto-construction) and which innovative.

For a concrete example of the process described above, in the Cariban family it is relatively straightforward to identify the sources of reanalyses that have relatively shallow time depth (summarized in Gildea 1998: 52–53), whereas not even all of the formal elements of the Proto-Cariban main clause construction are reconstructable, much less their sources (this is illustrated in Gildea 1998:

Chapter 5 *inter alia*; for a more specific illustration, cf. the reconstruction of number-marking on pp. 99–101).

When comparative information becomes less rich, or when it yields fewer cognate elements in particular cognate configurations, one needs other metrics to turn to that might help to separate older from more recent elements. In the next section, we review some of the metrics that have been proposed.

4.2.2   The Role of General Principles in Diagnosing Directionality

We open this section with the explicit caveat that the principles we list here are more heuristic than those in the previous section, following largely from the findings associated with grammaticalization studies. To the extent that they represent frequent outcomes of grammatical change, these are possible clues to the relative age of individual morphemes found in older constructions where the source constructions are no longer recoverable. In cases of internal reconstruction, where comparative evidence does not provide cognates outside of the attested construction, these principles may be the only criteria available to diagnose relative age.

Givón (2000: 120–121) offers the following general principles:
a)   **Phonetic Size:** The smaller a morpheme is, the older it is
b)   **Semantic size:** The more schematic (generic, grammaticalized, semantically opaque) the meaning of a morpheme is, the older it is
c)   **Distance from stem**: Other things being equal, the closer a morpheme is to the stem/root of the word, the older it is
d)   **Morphophonemic irregularity**: The more irregular or variable the allomorphs of a morpheme are, the older it is

While any of these individual properties could develop in more recent morphemes, when a morpheme is characterized by all four, it is hard to imagine a convincing argument that the reanalysis would be recent. Of course, very old morphemes can occur inside relatively recent constructions, so it is important not to confound the age of component morphemes with the age of the construction in which they occur.

In their discussion of directionality in Diachronic Construction Grammar, Traugott & Trousdale (2013: 112–124) identify three components of grammatical constructionalization as being consistently directional: increases in productivity and schematicity occur alongside decreases in compositionality (Barðdal & Gildea 2015: 15, 34, 37–41 have an independent discussion of these factors under the labels collocational **expansion** and increased **schematization**). Many constructions begin life with low type frequency, high coherence, and low schematicity. With the term *low type frequency*, we mean that few lexical items may occupy the open "slots" in the construction, with the term *high coherence*, we

refer to the fact that the lexical fillers of the relevant slot are closely related se-
mantically, and with the term *low schematicity*, we mean that the construction
is used in a limited set of communicative contexts and its meaning/usage is
more concrete, internally coherent, and perhaps componential. At the time of
initial reanalysis, the meaning of a new schematic construction is already less
compositional than that of the source. As a newer construction grows more
productive, a greater number (and variety) of lexical items can occur in its
open slots, leading to more complex lexical coherence among fillers of these
slots and greater schematicity in usage. As such, increased productivity then
drives a still less-concrete, less-componential constructional meaning.

As an illustration of the development described above, early on in the evo-
lution of the English *way*-construction, only the verb *go* appeared; by 1700 the
number had expanded to 16 motion verbs, by 1875 to 38 motion verbs (with
increasingly convoluted path or manner semantics), and by now the modern
construction seems to be a productive venue for non-motion verbs to gain
(sometimes metaphorical) motion semantics (Israel 1996). Similarly, in the
early stages of the English Progressive construction, the verb slot was limited
to activity verbs expressing events that took place in a stereotypical location;
as the construction developed, both of these restrictions gradually eased, lead-
ing to the highly productive Progressive Construction attested today (Bybee,
Perkins & Pagliuca 1994: 136; Heine, Claudi & Hünnemeyer 1991: 214–215). To
the extent that cognate constructions in different languages differ in terms of
semantic compositionality, productivity, and schematicity, this metric would
predict that the less productive and schematic versions of the construction
would be younger, that is, closer to the original reanalysis.

The final directionality metric that we discuss here is parsimony, which
would lead us to reconstruct a source construction or grammatical pattern
by selecting the hypothesized proto-form that requires the fewest changes in
order to arrive at the attested forms. On the one hand, this metric requires
a priori a reasonably solid classification of the languages in question, so as to
plot the number of changes accurately through the branches. On the other
hand, it is important to bear in mind that both cognate morphemes and cog-
nate grammatical patterns are internal components of entire cognate con-
structions. When the goal is to reconstruct a specific grammatical morpheme,
as it often is in grammaticalization studies, or a specific typological pattern,
it is possible to lose sight of the constructional context and consider only the
pattern in question, which in turn may lead to a claim of parsimony that is
divorced from the larger constructional context.

For example, in a simple application of parsimony, when a construction is
found only in one sub-branch, with some competing construction serving the

same function in the rest of the family, it is more economical to reconstruct this as an innovation at the level of that one sub-branch rather than to the proto-language. This reflects parsimony in that the more widespread construction would be inherited, and as such would achieve its broader distribution via fewer innovations. This conclusion would further entail that the construction with more restricted distribution would be a recent innovation, limited to a single branch. In addition to providing an economical account of these creative innovations, such a reconstruction also offers an economical approach to the negative side of these innovations: the older construction has been lost in only one branch and the innovative construction never existed in the rest of the family, and so there is no parsimony cost in its loss.

However, because innovative constructions come from source constructions with source morphology that generally occurs also externally to the source construction, the simple hypothesis does make the implicit claim that the innovative construction is the result of a relatively recent reanalysis compared to its more widespread competitor. This claim can be tested by identifying cognates to the component morphology that could have combined into a plausible source construction; ideally this cognate morphology (and even the source construction) would be found throughout the family, but certainly it should be seen in the branch where the innovative construction is attested. If source morphology is not attested, then one must seriously consider the alternative hypothesis, which is that the minority construction is actually an archaism that reconstructs to the proto-language, but has been lost in the rest of the genetic unit. Gildea (2002: 320) frames this as a general principle:

> **Identifiable Source Forms**: Morphology that has no "cognates" elsewhere in the grammar to serve as possible source forms for reanalysis or extension, or that was not plausibly borrowed into the language from an identifiable source in another language, is more likely to be old.

This principle follows from considerations of parsimony, in recognition of the costs of losing cognate source material in a relatively shallow time period. One more difference between syntactic constructions and lexical items is that constructions generally contain multiple morphemes and sub-constructions, most of which began outside of the construction in question; prior to the reanalysis that created the innovative construction, a given sequence of morphemes and slots could have been simply one collocation among others. At the time of initial reanalysis, the source construction with all of its component morphemes should continue to exist alongside the innovative construction. As time passes, an innovative construction and its source construction will drift apart, each

changing independently and each also independently available to be inherited by subsequent daughter languages.

The loss of the source construction, as well as each independent loss of a lexical or morphological source, must be counted as separate changes; if all traces of the source construction have been lost, this effectively neutralizes the simple argument from parsimony. In the absence of a clear argument from economy, the question shifts to one of how plausible each timeline appears. With a greater time depth, simple lexical replacement makes it increasingly likely that the source lexical items and other morphology would have been replaced by innovative forms, and indeed it is not unusual for an older construction to contain morphology that has no obvious external source forms. However, if one posits that a construction with no external cognate morphology is a relatively recent innovation, then not just the source construction, but all its component source forms must have been lost in every branch of the family; at least the most local of these losses must have taken place in the time period since the initial reanalysis in the one branch where the innovation is found.[3] This invites one to seek source components for the more widespread construction(s) in the other branches of the family; if those are readily identifiable, this would serve to effectively reverse the argument from parsimony, making it more plausible that the widespread construction is the innovation and the construction with more restricted distribution is conservative.

Following this principle, Gildea (1998: 51) uses the lack of identifiable source forms for personal prefixes, TAM suffixes, and the personal number suffix as evidence that a given construction containing those morphemes (his "Set 1 verbal system") reconstructs back to Proto-Cariban. Similarly, partly on the basis of the absence of source forms, Gildea (2002: 322–323) argues that the A/Sa subset of the personal prefixes in Proto-Tupi-Guarani represents an older component of the hierarchical agreement system, in particular prior to the addition of the O/So prefixes, for which sources are still readily identifiable. Pacchiarotti (this volume) provides another nice exposition of this situation in her reconstruction of the source of the case marker for the subject possessor in the Viceitic (Chibchan) alienable possession construction.

With these principles, once we are able to identify cognates, we are well equipped with knowledge about syntactic change that allows for positing a reconstructed construction, pattern, or feature from which the grammars of

---

3 To be fair, given the inevitable limitations of grammatical descriptions, and in particular the preliminary state of description for some languages (and language families) of the world, it is a logical possibility that source forms may still exist, but have simply not yet been described.

the descendant languages are more likely to be explained. These principles are broadly compatible with Givón's (1971: 413) well-known remark that "today's morphology is yesterday's syntax", and indeed these principles owe much to the decades of work in grammaticalization, which has aided us in recognizing lexical cognates by "undoing" the processes that have led to the grammaticalization of elements like clitics, auxiliary verbs, bound morphemes, etc., and which also provide many of the examples of specific constructional changes. Despite the historical disconnect between the intellectual communities dedicated, on the one hand, to the comparative method and, on the other, to grammaticalization studies, we suggest that the principles as articulated here are compatible with both approaches.

This concludes our theoretical arguments, and the methodological principles that follow from them. We turn now to the contributions in the remainder of this volume.

5 Conclusions

The final step of the Comparative Method is to propose an original state of affairs that may account for the outcomes found in the attested languages. With this in mind, Eythórsson & Barðdal (2016: 87) point out that many articles and books dealing with syntactic reconstruction do not actually carry out a reconstruction. Rather, they draw a scenario that allows us to understand the development from the proto-language to the attested descendant languages; in that scenario a reconstruction of the original stage may be implicit, but most times it is not presented in detail and the exact status of the reconstruction is not discussed. The eight chapters that appear in this volume all offer explicit reconstruction of syntax. Since all were written before this introduction was written, each represents an independent solution to the problems discussed in Section 4. Four of the chapters deal with the issue of how to identify cognates: Gildea & Castro Alves, Kikusawa, Luraghi and Barðdal & Eythórsson. The remaining four chapters are concerned with directionality in syntactic change: Pacchiarotti, Lavidas & Kulikov, Pat-El, and Luján & López Chala. We offer a brief summary of each of these contributions below.

Gildea & Castro Alves focus their efforts on reconstructing a specific grammatical pattern that is typologically rare, in which nominative case marking of pronouns co-occurs with absolutive indexation of main verbs. This pattern was originally attested in only the Jê and Cariban language families, both spoken in the Amazonian region of South America. Gildea & Castro Alves do not attempt to reconstruct these patterns in the abstract, but rather they identify

and reconstruct the subordinate source constructions that contain absolutive verbal indexation alongside ergative case marking of a free nominal A. They then identify two kinds of biclausal source constructions that led to the loss of the ergative case-marked A. In the first type, the S/A of the matrix clause is coreferential with the A of the subordinate clause, leading to a complex clause in which the A of the subordinate clause simply does not occur. In the second type, the source construction does not contain such coreference conditions, thus, after reanalysis, the innovative construction contains an ergative A. In subsequent constructional changes, two distinctive patterns emerge. In Timbira, the ergative A is conserved in the past tense, but is replaced by an unmarked topic pronoun/noun in the nonpast. In the Suyá future and negative, the ergative is conserved in pronoun subjects, but for nominal subjects, the ergative marking is lost. None of these represent deep reconstructions: the lexical and morphological cognates within the innovative constructions provide a clear roadmap to reconstructing the source constructions, from which the directionality of additional changes can be identified.

Kikusawa introduces a methodology for carrying out syntactic reconstruction for languages for which no written records exist. In such cases, the comparanda must be extracted from modern languages for which genetic relations have already been established. Kikusawa compares the alignment system of five Austronesian languages which between them show high disparity of alignment patterns. As a first step, abstract clause structures are described and classified on the basis of transitivity and case marking, providing a descriptive representation from which patterns belonging to typologically different languages may be compared. Each clause structure is also marked for the position in which the remnants of earlier genitive pronouns are found; these are traditionally analyzed as ergative markers, having marked the A of transitive clauses. Through this step, cognate clause structures are established. The second step involves comparing and analyzing the different positions found with the genitive across the five daughter languages. This reveals that the genitive is not limited in distribution to transitive clauses, but is also found in both monadic and dyadic intransitives, and also in languages with a synchronic accusative system. A further comparison uncovers an earlier merger of the genitive and the nominative pronominal sets. The third step in this process is to identify the directionality of the relevant change. To this end, further scrutiny of the data reveals that the morphological merger of the genitive/nominative appears to be functionally motivated by a change in word order from verb-subject to subject-verb clause structure. On this basis, Kikusawa reconstructs Proto-Malayo-Polynesian as an ergative system, from which she tracks its development into the different types of alignment systems found in the five daughters under discussion.

Luraghi analyses two different external possessor constructions across the early Indo-European daughter languages, i.e. the dative external possessive construction and the double case construction, both denoting inalienable possession. The two constructions are unevenly distributed across the daughter languages, with the double case construction found only in Homeric Greek, Hittite and Armenian. However, the syntactic and semantic properties documented for the double case construction do not converge in these three languages, thus suggesting that the relevant constructions are innovations in all three branches. In contrast, the dative external possessive construction is found in Greek, Latin, Germanic, Baltic, Slavic, Hittite and Vedic with similar syntactic and semantic properties. The evidence from Hittite is ambiguous, and while it is scanty for Vedic, it still suggests a Proto-Indo-European origin. Luraghi analyses both the syntactic and the semantic properties of the relevant constructions, and only in instances where the syntactic and semantic properties are a match across the daughter languages does she argue that they have a shared origin.

Barðdal & Eythórsson propose a research program for how to identify cognates in syntax, in particular within the realm of argument structure. They base their program on Watkins' (1995) proposal that it may be possible to reconstruct larger units of grammar through identifying morphological flags of larger constructions. This – as they point out – is no insignificant part of grammar: *the whole of morphosyntax*. In more detail, Barðdal & Eythórsson propound that cognates in argument structure constructions may be identified through a) cognate lexical verbs, b) cognate case frames, c) cognate predicate structure and d) cognate case morphology. They then suggest supplementing Watkins' proposal with one further analytical step, namely through e) identifying cognate argument structure constructions with the aid of noncognate, but synonymous, lexical predicates. The rationale behind this addition to the program lies in facts of lexical replacement in general and known patterns of changes in argument structure in particular. This program allows for the identification of cognate argument structure constructions across a deeper time span than corresponding reconstructions based only on cognate lexical verbs.

Pacchiarotti begins with a typologically unusual case-marking pattern in the Costa Rican languages Bribri and Cabécar (from the Viceitic branch of Chibchan), in which the transitive subject is marked with one ergative marker in most constructions, but with a different ergative marker, a mysterious *wã* in the Perfect (a.k.a "Anterior") construction. She identifies the immediate source of the Perfect construction as a resultative participle in a possessive construction (a source well-known from European languages). In the source possessive predicate construction, the possessor is marked by *wã*. Next, Pacchiarotti attempts to find a deeper source for this Possessor marker *wã*. Given that there

is no readily available synchronic source in Bribri or Cabécar, she searches for cognates in Possessive Predicate constructions throughout the Isthmic branch (to which Viceitic belongs), but finds only unconvincing prospects. According to the principle of parsimony, one might conclude that *wã* was an innovation in Proto-Viceitic. However, this conclusion clashes with the fact that there is no synchronically available source in either Viceitic or Isthmic for this alleged 'new' piece of grammar. By expanding her search beyond the Isthmic branch, Pacchiarotti identifies other possible cognates, some quite distant. These ultimately allow her to reconstruct a Proto-Chibchan possessive predicate construction in which the source of the Viceitic possessor marker *wã* reconstructs back to a Proto-Chibchan word meaning 'thing'.

Lavidas & Kulikov's contribution is focused on the directionality of changes in the domains of tense-aspect and transitivity-voice in the history of Vedic and Ancient Greek, a topic falling out from their reconstruction of the linguistic system of Proto-Indo-European, manifested as "split causativity" in the daughters. Lavidas & Kulikov document a correlation between verbal formations of the present system being used transitively or causatively, on the one hand, and being used intransitively, on the other. The evidence for their reconstruction is found through relics in Vedic and innovations in Ancient Greek. The Vedic relics consist of active perfects that show up as intransitives, thus being functionally middles, while the innovations in Ancient Greek are manifested through the rise of new markers of transitivity oppositions. These new oppositions consist of the common distinctions between active vs. passive, on the one hand, and causative vs. anticausative, on the other. Thus, changes in voice are parallel with another development in the history of Greek, namely the separation of tense and aspect. Active morphology thus becomes associated with transitive alternations rather than aspectual properties. The development that Lavidas & Kulikov outline serves as evidence for the directionality of the historical changes that have taken place in the linguistic system of Koine Greek, triggered by the original oppositions of the relevant domains of tense-aspect and transitivity-voice in the proto-language.

Pat-El focuses on the development of adverbial subordination across several Semitic languages; subordination patterns which are parallel in many respects, except that the subordinators are not lexically cognate. Pat-El discusses two different potential scenarios for reconstruction: a) one assuming a proto-structure for these subordination patterns with lexical replacement being responsible for lack of cognates, or b) parallel development motivated by a certain type of nominally headed relative clauses. Case morphology of nouns in nominally headed relatives is typically impoverished due to the status of such nouns as proclitics, leading to a reanalysis of nominal heads with

spatial, temporal and causal meaning as adverbial subordinators. Pat-El argues for the second reconstructional scenario outlined above, proposing that a certain type of relative clause may be responsible for the parallel development in each of the daughter languages. On this analysis it is assumed that parallel development has taken place multiple times and that the impetus for the development has been carried down to the daughters, in fact still being a part of the synchronic grammar of the relevant daughter languages where this adverbial subordination pattern exists. While emphasizing that the relevant structures across the daughters are most likely the result of independent parallel development, Pat-El also highlights the fact that parallel developments are set in motion by shared structures which create the type of coercion needed for analogous evolutionary paths to emerge.

Luján & López Chala focus on the fate of the desinences based on PIE *-b^hi in the ancient Indo-European languages in order to reconstruct the history of this morpheme and the semantic path that it has followed since its PIE origins. Endings continuing PIE *-b^hi occur in several branches of the Indo-European family, both in the nominal and in the pronominal declension. They display a whole array of different semantic roles that range from their use as prototypical Recipients with pronouns to the expression of Instrument, Comitative, Agent, Manner or Place with nouns. The most common use of *-b^hi -endings across languages is as an Instrumental marker and, therefore, *-b^hi has traditionally been reconstructed as the athematic instrumental plural ending of PIE. However, traditional reconstructions were "static" and did not pay explicit attention to the semantic paths of change followed by the ending and to the actual occurrence of different meanings that were not easy to reconcile. In their "dynamic" approach Luján & López Chala argue that the directionality of the change can be reconstructed on the basis of what we currently know about the historical tendencies of change of the markers of semantic roles. This allows for an integrated account of the history of the desinence and the reconstruction of its original Comitative value, which, in turn, must have been the result of a grammaticalization process of a postposition with local meaning ('by, next to').

Acknowledgements

References

Adger, David. 2003. *Core Syntax: A Minimalist Approach*. Oxford: Oxford University Press.

Barðdal, Jóhanna. 2013. Construction-Based Historical-Comparative Reconstruction. *The Oxford Handbook of Construction Grammar*, ed. by Graeme Trousdale & Thomas Hoffmann, 438–457. Oxford: Oxford University Press.

Barðdal, Jóhanna. 2014. Syntax and Syntactic Reconstruction. *The Routledge Handbook of Historical Linguistics*, ed. by Claire Bowern & Bethwyn Evans, 343–373. London: Routledge.

Barðdal, Jóhanna, Valgerður Bjarnadóttir, Serena Danesi, Tonya Kim Dewey, Thórhallur Eythórsson, Chiara Fedriani & Thomas Smitherman. 2013. The Story of 'Woe'. *Journal of Indo-European Studies* 41(3–4): 321–377.

Barðdal, Jóhanna & Thórhallur Eythórsson. 2012a. Reconstructing Syntax: Construction Grammar and the Comparative Method. *Sign-Based Construction Grammar*, ed. by Hans C. Boas & Ivan A. Sag, 257–308. Stanford: CSLI Publications.

Barðdal, Jóhanna & Thórhallur Eythórsson. 2012b. "Hungering and Lusting for Women and Fleshly Delicacies": Reconstructing Grammatical Relations for Proto-Germanic. *Transactions of the Philological Society* 110(3): 363–393.

Barðdal, Jóhanna & Spike Gildea. 2015. Diachronic Construction Grammar: Epistemological Context, Basic Assumptions and Historical Implications. *Diachronic Construction Grammar*, ed. by Jóhanna Barðdal, Elena Smirnova, Lotte Sommerer & Spike Gildea, 1–50. Amsterdam: John Benjamins.

Barðdal, Jóhanna & Thomas Smitherman. 2013. The Quest for Cognates: A Reconstruction of Oblique Subject Constructions in Proto-Indo-European. *Language Dynamics and Change* 3(1): 28–67.

Bauer, Brigitte. 2000. *Archaic Syntax in Indo-European: The Spread of Transitivity in Latin and French*. Berlin: Mouton de Gruyter.

Beekes, Robert S. P. 2011. *Comparative Indo-European Linguistics: An Introduction* (2nd revised ed. by Michiel de Vaan). Amsterdam: John Benjamins.

Boas, Hans C. 2003. *A Constructional Approach to Resultatives*. Stanford: CSLI Publications.

Bowern, Claire. 2008. Syntactic Change and Syntactic Borrowing in Generative Grammar. *Principles of Syntactic Reconstruction*, ed. by Gisela Ferraresi & Maria Goldbach, 187–216. Amsterdam: John Benjamins.

Bowern, Claire. 2017. General and Methodological Issues. *Handbook of Comparative and Historical Indo-European Linguistics*, vol. 1, ed. by Jared Klein, Brian Joseph, Matthias Fritz & Mark Wenthe, 1–7. Berlin: Walter de Gruyter.

Bybee, Joan, Revere Perkins, & William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: University of Chicago Press.

Campbell, Lyle. 1986. Theories of Syntactic Change and Finnish Historical Syntax. *Journal of the Atlantic Provinces Linguistic Association* 8: 72–93.

Campbell, Lyle. 1988. Review of *Language in the Americas*, by Joseph H. Greenberg, 1987. *Language* 64: 591–615.

Campbell, Lyle. 1998. *Historical Linguistics: An Introduction*. Edinburgh: Edinburgh University Press.

Campbell, Lyle. 2003. How to Show Languages Are Related: Methods for Distant Genetic Relationship. *The Handbook of Historical Linguistics*, ed. by Brian D. Joseph & Richard D. Janda, 262–282. Oxford: Blackwell.

Campbell, Lyle & William Poser. 2008. *Language Classification: History and Method*. Cambridge: Cambridge University Press.

Clackson, James. 2007. *Indo-European Linguistics*. Cambridge: Cambridge University Press.

Claudi, Ulrike. 1994. Word Order Change as Category Change. *Perspectives on Grammaticalization*, ed. by William Pagliuca, 191–231. Amsterdam: John Benjamins.

Croft, Willam. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.

Danesi, Serena, Cynthia A. Johnson & Jóhanna Barðdal. 2017. Between the Historical Languages and the Reconstructed Language: An Alternative Approach to the Gerundive + "Dative of Agent" Construction in Indo-European. *Indogermanische Forschungen* 122: 143–188.

Daniels, Don. 2015. *A Reconstruction of Proto-Sogeram: Phonology, Lexicon, and Morphosyntax*. University of California, Santa Barbara, Ph.D. dissertation.

Daniels, Don. 2017. A Method for Mitigating the Problem of Borrowing in Syntactic Reconstruction. *Studies in Language* 41(3): 577–614.

Daniels, Don. 2020. *Grammatical Reconstruction: The Sogeram Languages of New Guinea*. Mouton De Gruyter.

Delbrück, Berthold. 1878. *Die altindische Wortfolge aus dem Çatapathabrāhmana dargestellt* (Syntaktische Forschungen III). Halle: Verlag des Waisenhauses.

Delbrück, Berthold. 1893–1900: *Vergleichende Syntax der indogermanischen Sprachen* I–III. Strassburg: Trübner.

Delbrück, Berthold. 1907. *Synkretismus. Ein Beitrag zur germanischen Kasuslehre*. Strassburg: Trübner.

Dixon, R. M. W. 1997. *The Rise and Fall of Languages.* Cambridge: Cambridge University Press.

Drinka, Bridget. 1999. Alignment in Early Proto-Indo-European. *Language Change and Typological Variation: In Honor of Winfred P. Lehmann on the Occasion of his 83rd Birthday*, Vol II, ed. by C. F. Justus & E. C. Polomé, 464–500. Washington, DC: Institute for the study of Man.

Dunn, Michael, Tonya Kim Dewey, Carlee Arnett, Thórhallur Eythórsson & Jóhanna Barðdal. 2017. Dative Sickness: A Phylogenetic Analysis of Argument Structure Evolution in Germanic. *Language* 93(1): e1–e22.

Dybo, Anna V., & George S. Starostin. 2008. In Defense of the Comparative Method, or the End of the Vovin Controversy. *Aspects of Comparative Linguistics* 3 (Moscow: Russian State University for the Humanities), 119–258.

Eythórsson, Thórhallur & Jóhanna Barðdal. 2011. Die Konstruktionsgrammatik und die komparative Methode. *Indogermanistik und Linguistik im Dialog: Akten der XIII. Fachtagung der Indogermanischen Gesellschaft vom 21. bis 27. September 2008 in Salzburg*, ed. by Thomas Krisch & Thomas Lindner, 148–156. Wiesbaden: Reichert.

Eythórsson, Thórhallur & Jóhanna Barðdal. 2016. Syntactic Reconstruction in Indo-European: State of the Art. *Veleia* 33: 83–102.

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

Fillmore, Charles J., Paul Kay & Mary Catherine O'Connor. 1988. Regularity and Idiomaticity in Grammatical Constructions: The Case of *Let Alone*. *Language* 64: 501–538.

Fortson, Benjamin W. IV. 2004. *Indo-European Language and Culture: An Introduction*. Oxford: Blackwell.

Fox, Anthony. 1995. *Linguistic Reconstruction: An Introduction to Theory and Method*. Oxford: Oxford University Press.

Fried, Mirjam & Jan-Ola Östman. 2005. Construction Grammar and Spoken Language: The Case of Pragmatic Particles. *Journal of Pragmatics* 37(11): 1752–1778.

Friedrich, Paul. 1975. *Proto-Indo-European Syntax*. Washington: Institute for the Study of Man.

Gamkrelidze, Thomas V. & Vjaceslav V. Ivanov. 1995. *Indo-European and the Indo-Europeans: A Reconstruction and Historical Analysis of a Proto-Language and Proto-Culture* I–II [Translated by Johanna Nichols]. Berlin: De Gruyter Mouton.

Garrett, Andrew. 1990. *The Syntax of Anatolian Pronominal Clitics*. Harvard University PhD dissertation.

Gildea, Spike. 1992. *Comparative Cariban Morphosyntax: On the Genesis of Ergativity in Independent Clauses*, University of Oregon PhD dissertation.

Gildea, Spike. 1998. *On Reconstructing Grammar: Comparative Cariban Morphosyntax*. Oxford: Oxford University Press.

Gildea, Spike. 2000. On the Genesis of the Verb Phrase in Cariban Languages. *Reconstructing Grammar: Comparative Linguistics and Grammaticalization*, ed. by Spike Gildea, 65–106, Amsterdam: John Benjamins.

Gildea, Spike. 2002. Pre-Proto-Tupí-Guaraní Main Clause Person-Marking. *Línguas Indígenas Brasileiras. Fonologia, Gramática e História. Atas do I Encontro Internacional do GTLI da ANPOLL*, Vol I, ed. by Ana Suelly Cabral & Aryon Rodrigues, 315–326. Belem: Editoria Universitária U.F.P.A.

Gildea, Spike. 2008. Explaining Similarities between Main Clauses and Nominalized Clauses. *La structure des langues amazoniennes*, ed. by Ana Carla Bruno, Frantomé Pacheco, Francesc Queixalos, & Leo Wetzels. *Amérindia* 32: 57–75.

Givón, Talmy. 1971. Historical Syntax and Synchronic Morphology: An Archaeologist's Field Trip. *Chicago Linguistic Society* 7: 394–415.

Givón, Talmy. 2000. Internal Reconstruction: As Method, as Theory. *Reconstructing Grammar: Comparative Linguistics and Grammaticalization Theory,* ed. by Spike Gildea, 107–159. Amsterdam: John Benjamins.

Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Goldberg, Adele E. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.

Greenberg, Joseph. 1957. *Essays in Linguistics.* Chicago: University of Chicago Press.

Greenberg, Joseph H. 1966. Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. *Universals of Language*, 2nd ed., ed. by Joseph H. Greenberg, 73–113. Cambridge, Mass: MIT Press.

Greenberg, Joseph. 1987. *Language in the Americas*. Stanford: Stanford University Press.

Greenberg, Joseph. 1989. Classification of American Indian Languages: A Reply to Campbell. *Language* 65: 107–114.

Hale, Mark. 1987. Notes on Wackernagel's Law in the Language of the Rigveda. *Studies in Memory of Warren Cowgill*, ed. by Calvert Watkins, 38–50. Berlin: De Gruyter.

Hale, Mark. 2007: *Historical Linguistics: Theory and Method*, Oxford: Blackwell.

Harris, Alice C. 1985. *Diachronic Syntax: The Kartvelian Case.* New York: Academic Press.

Harris, Alice C. 2008. Reconstruction in Syntax: Reconstruction of Patterns. *Principles of Syntactic Reconstruction*, ed. by Gisella Ferraresi & Maria Goldbach, 73–95. Amsterdam: John Benjamins.

Harris, Alice C. & Lyle Campbell. 1995. *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: Cambridge University Press.

Harrison, S. P. 2003. On the Limits of the Comparative Method. *The Handbook of Historical Linguistics*, ed. by Brian D. Joseph & Richard D. Janda, 213–243. Oxford: Blackwell.

Havers, Wilhelm. 1911. *Untersuchungen zur Kasussyntax der idg. Sprachen*. Strassburg: Trübner.

Heine, Bernd, Ulrike Claudi, & Friederike Hünnemayer. 1991. *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press.

Heine, Bernd & Tania Kuteva. 2004. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

Hock, Hans H. 2000. Genre, Discourse and Syntax in Early Indo-European, With Emphasis on Sanskrit. *Textual Parameters in Older Languages*, ed. by Susan C. Herring, Pieter van Reenen & Lene Schøsler, 163–195. Amsterdam: John Benjamins.

Hoenigswald, H. M. 1978. The Annus Mirabilis 1876 and Posterity. *Transactions of the Philological Society* 76: 17–35.

Hopper, Paul & Elizabeth Closs Traugott. 2003. *Grammaticalization*. 2nd ed. Cambridge: Cambridge University Press.

Israel, Michael. 1996. The *Way* Constructions Grow. *Conceptual Structure, Discourse and Language*, ed. by Adele E. Goldberg, 217–230. Stanford: CSLI.

Jackendoff, Ray. 1997. Twistin' the Night Away. *Language* 73: 534–559.

Jeffers, Robert J. 1976. Syntactic Change and Syntactic Reconstruction. *Current Progress in Historical Linguistics: Proceedings of the Second International Conference on Historical Linguistics, Tucson, Arizona 12–16 January 1976*, ed. by William M. Christie, 1–16. Amsterdam: North Holland.

Johnson, Cynthia A., Peter Alexander Kerkhof, Leonid Kulikov, Esther Le Mair & Jóhanna Barðdal. 2019. Argument Structure, Conceptual Metaphor and Semantic Change: How to Succeed in Indo-European without Really Trying. *Diachronica* 36(4): 463–508.

Jolly, Julius. 1872. *Ein Kapitel vergleichender Syntax: Der Conjunktiv und Optativ und die Nebensätze im Zend und Altpersischen*. Munich: Ackermann.

Justus, Carol F. 1993. Dislocated Imperatives in the Indo-European Prayer. *Word* 44: 273–294.

Justus, Carol F. 2000. Word Order and the First Person Imperative. *Stability, Variation and Change of Word-Order Patterns over Time*, ed. by Rosanna Sornicola, Eric Poppe & Ariel Shisha-Halevy, 165–184. Amsterdam: John Benjamins.

Keydana, Götz. 2018. The Syntax of Proto-Indo-European. *Handbook of Comparative and Historical Indo-European Linguistics*, vol. 3, ed. by Jared Klein, Brian Joseph, Matthias Fritz & Mark Wenthe, 2195–2228. Berlin: Walter de Gruyter.

Kikusawa, Ritsuko. 2002. *Proto Central Pacific Ergativity: Its Reconstruction and Development in the Fijian, Rotuman and Polynesian Languages*. Canberra: Pacific Linguistics.

Kikusawa, Ritsuko. 2003. The Development of Some Indonesian Pronominal Systems *Historical Linguistics 2001: Selected Papers from the 15th International Conference on Historical Linguistics, Melbourne, 13–17 August 2001*, ed. by Barry J. Blake, Kate Burridge & Jo Taylor, 237–268. Amsterdam: John Benjamins.

Klein, Jared S. 2010. Review of Principles of Reconstruction, ed. by Gisela Ferraresi & Maria Goldbach (2008), *Language* 86: 720–726.

Klimov, G. A. 1977. *Tipologija jasykov aktivnogo stroja* [The Typology of Active Languages]. Moscow: Nauka.

Krasukhin, Constantin G. 2017. Methods in Reconstruction. *Handbook of Comparative and Historical Indo-European Linguistics*, vol. 1, ed. by Jared Klein, Brian Joseph, Matthias Fritz & Mark Wenthe, 15–20. Berlin: Walter de Gruyter.

Kuhn, Adalbert. 1853. Ueber die durch nasale erweiterten Verbalstämme. *Zeitschrift für vergleichende Sprachforschung* 2: 455–471.

Lakoff, George. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.

Lehmann, Winfred P. 1974. *Proto-Indo-European Syntax*. Austin: University of Texas Press.

Lehmann, Winfred P. 1993. *Theoretical Bases of Indo-European Linguistics*, London: Routledge.

Lehmann, Winfred P. 1994. *Residues of Pre-Indo-European Active Structure and their Implications for the Relationships among the Dialects*. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.

Lightfoot, David W. 1979. *Principles of Diachronic Syntax*. Cambridge: Cambridge University Press.

Lightfoot, David W. 1980. On Reconstructing a Proto-Syntax. *Linguistic Reconstruction and Indo-European Syntax*, ed. by Paolo Ramat, 27–45. Amsterdam: John Benjamins.

Lightfoot, David W. 1999. *The Development of Language: Acquisition, Change and Evolution*. Oxford: Blackwell.

Lightfoot, David W. 2002a. Myths and the Prehistory of Grammars. *Journal of Linguistics* 38(1): 113–136.

Lightfoot, David W. 2002b. More Myths. *Journal of Linguistics* 38(3): 619–626.

Lightfoot, David W. 2006. *How New Languages Emerge*. Cambridge: Cambridge University Press.

Mendoza, Julia. 1998. Sintaxis. *Manual de lingüística indoeuropea* III, ed. by Francisco R. Adrados, Alberto Bernabé & Julia Mendoza, 141–246. Madrid: Ediciones Clásicas.

Mengden, Ferdinand von. 2008. Reconstructing Complex Structures: A Typological Perspective. *Principles of Syntactic Reconstruction*, ed. by Gisella Ferraresi & Maria Goldbach, 97–119. Amsterdam: John Benjamins.

Michaelis, Laura A. & Josef Ruppenhofer. 2001. *Beyond Alternations: A Constructional Model of the German Applicative Pattern*. Stanford: CSLI Publications.

Miller, D. Gary 1975. Indo-European: VSO, SOV, SVO, or all three? *Lingua* 37: 31–52.

Mithun, Marianne. 2006. Integrating Approaches to Diversity: Argument Structure on the *NW* Coast. *Diversity in Language: Perspectives and Implications*, ed. by Yoshiko Matsumoto, David Oshima, Orrin Robinson & Peter Sells, 9–36. Stanford, CA: CSLI.

Mithun, Marianne. 2012. Core Argument Patterns and Deep Genetic Relations: Hierarchical Systems in Northern California. *Argument Structure and Grammatical Relations: A Crosslinguistic Typology,* ed. by Pirkko Suihkonen, Bernard Comrie & Valery Solovyev, 258–294. Amsterdam: John Benjamins.

Nunberg, Geoffrey, Ivan A. Sag & Thomas Wasow. 1994. Idioms. *Language* 70: 491–538.

Osthoff, Hermann & Karl Brugman. 1878. *Morphologische Untersuchungen auf dem Gebiet der indogermanischen Sprachen* 1. Leipzig: Hirzel.

Pedersen, Holger. 1907. Neues und nachträgliches. *Zeitschrift für vergleichende Sprachforschung* 40: 129–217.

Penney, John. 2000. *Review of Berthold Delbrück y la sintaxis indoeuropea hoy. Actas del coloquio de la Indogermanische Gesellschaft*, ed. by Emilio Crespo & José L. García Ramón, Madrid, Ediciones de la UAM & Wiesbaden, Dr. Ludwig Reichert Verlag. 1997. *Kratylos* 45: 29–35.

Pires, Acrisio & Sarah G. Thomason. 2008. How Much Syntactic Reconstruction is Possible? *Principles of Syntactic Reconstruction*, ed. by Gisela Ferraresi & Maria Goldbach, 27–72. Amsterdam: John Benjamins.

Pooth, Roland, Peter Alexander Kerkhof, Leonid Kulikov & Jóhanna Barðdal. 2019. The Origin of Non-Canonical Case Marking of Subjects in Proto-Indo-European: Accusative, Ergative, or Semantic Alignment. *Indogermanische Forschungen* 124: 245–263.

Rankin, Robert. 1992. Review of Greenberg. 1987. *International Journal of American Linguistics* 58: 324–351.

Ringe, Donald. 1992. On Calculating the Factor of Chance in Language Comparison. *Transactions of the American Philosophical Society* 82: 1–110.

Roberts, Ian. 2007. *Diachronic Syntax*. Oxford: Oxford University Press.

Rumsey, Alan. 1987a. Was Proto-Indo-European an Ergative Language? *Journal of Indo-European Studies* 15: 19–37.

Rumsey, Alan. 1987b. The Chimera of Proto-Indo-European Ergativity: Lessons for Historical Syntax. *Lingua* 71: 297–318.

Seržant, Ilja. 2015. An Approach to Syntactic Reconstruction. *Perspectives on Historical Syntax*, ed. by Carlotta Viti, 117–154. Amsterdam: John Benjamins.

Szemerényi, Oswald. 1996. *Introduction to Indo-European Linguistics*. Oxford: Oxford University Press.

Thurneysen, Rudolf. 1885. Der indogermanische Imperativ. *Zeitschrift für vergleichende Sprachforschung* 27(2): 172–180.

Tomasello, Michael. 2003. *Constructing a Language*. Cambridge, Mass: Harvard University Press.

Traugott, Elizabeth Closs & Graeme Trousdale. 2013. *Constructionalization and Constructional Change*. Oxford: Oxford University Press.

Uhlenbeck, Christianus C. 1901. Agens und Patiens im Kasussystem der indogermanischen Sprachen. *Indogermanische Forschungen* 12: 170–171.

Vaillant, André. 1936. L'ergatif indo-européen. *Bulletin de la Société de Linguistique de Paris* 37: 93–108.

Vajda, Edward. 2018. Dene-Yeniseian: Progress and Unanswered Questions. *Diachronica* 35(2): 277–295.

Vázquez-González, Juan G. & Jóhanna Barðdal. 2019. Reconstructing the Ditransitive Construction for Proto-Germanic: Gothic, Old English and Old Norse-Icelandic. *Folia Linguistica Historica* 40(2): 555–620.

Villar, Francisco. 1983. *Ergatividad, acusatividad y género en la familia lingüística indoeuropea*. Salamanca: Universidad de Salamanca.

Vincent, Nigel, & Ian Roberts. 1999. *Remarks on Syntactic Reconstruction.* A paper delivered at the Annual Meeting of the Deutsche Gesellschaft für Sprachwissenschaft, University of Konstanz.

Viti, Carlotta. 2015. Historical Syntax: Problems, Materials, Methods, Hypotheses. *Perspectives on Historical Syntax*, ed. by Carlotta Viti, 3–31. Amsterdam: John Benjamins.

Wackernagel, Jacob. 1892. Ueber ein Gesetz der indogermanischen Wortstellung. *Indogermanische Forschungen* 1: 333–436.

Walkden, George. 2009. *The Comparative Method in Syntactic Reconstruction*, University of Cambridge M. A. thesis.

Walkden, George. 2013. The Correspondence Problem in Syntactic Reconstruction. *Diachronica* 30: 95–122.

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

Watkins, Calvert. 1964. Preliminaries to the Reconstruction of Indo-European Sentence
      Structure. *Proceedings of the IX International Congress of Linguists*, ed. by H. G. Lunt,
      1035–1045. The Hague: Mouton.

Watkins, Calvert. 1976. Towards Proto-Indo-European Syntax: Problems and
      Pseudo-Problems. *Papers from the Parasession on Functionalism*, ed. by S. B. Steever,
      C. A. Walker & S. S. Mufwene, 205–326. Chicago: Chicago Linguistic Society.

Watkins, Calvert. 1995. *How to Kill a Dragon: Aspects of Indo-European Poetics*. Oxford:
      Oxford University Press.

Wichmann, Søren. 2008. The Study of Semantic Alignment: Retrospect and the State
      of the Art. *The Typology of Semantic Alignment*, ed. by M. Donohue & S. Wichmann,
      3–23. Oxford: Oxford University Press.

Willi, Andreas. 2018. *Origins of the Greek Verb*. Cambridge: Cambridge University Press.

Willis, David. 2011. Reconstructing Last Week's Weather: Syntactic Reconstruction and
      Brythonic Free Relatives. *Journal of Linguistics* 47: 407–446.

Winter, Werner. 1984. Reconstructional Comparative Linguistics and the Reconstruction
      of the Syntax of Undocumented Stages in the Development of Languages and
      Language Families. *Historical Syntax*, ed. by J. Fisiak, 613–625. The Hague: Mouton.

PART 1

Cognacy

∵

Reconstructing the Source of Nominative-Absolutive Alignment in Two Amazonian Language Families

Spike Gildea and Flávia de Castro Alves

Abstract

In this article, we take the strong position that syntactic constructions can be reconstructed, first by identifying constructional cognates, then by identifying evidence for the directionality of constructional change that best explains the modern distribution of the cognate constructions from the hypothesized source construction. Further, we argue that the grammatical properties of the resultant constructions are often best explained by a combination of their etymological source(s) and the evolutionary pathways by which they arise. We illustrate these larger theoretical claims by reconstructing a typologically unusual set of constructions in the Jê and Cariban families, which present a rare ergative alignment pattern we call **nominative-absolutive**. Prior to 2010 this alignment pattern, which combines nominative free pronouns and absolutive verbal indexation, was held to be impossible, and it remains attested in very few language families. In the Jê and Cariban languages, this alignment type always occurs as part of ergative splits conditioned by TAM, which are again counter to previously claimed universals in that they are conditioned by future tense, imperfective aspects, and agent-oriented modalities. We reconstruct the sources of these nominative-absolutive constructions and then argue that the unusual formal properties and functional distributions of the nominative-absolutive clause types are both best understood as combinations of typologically unusual source constructions that follow well-established diachronic pathways of tense-aspect-mood renewal.

1 Introduction

As part of a volume on the topic of how syntax can be reconstructed successfully, this article offers multiple case studies that highlight the methodology by which cognate constructions can be identified in related languages.[1] As a

---

precondition to assessing the relative plausibility of competing reconstructions, we also examine arguments that make the case for specific, local directions of change. We assert that this sort of methodology could be used for any language family in which cognates are identifiable.

As part of the community that studies alignment typology, both synchronic and diachronic, we focus our reconstructions on an unusual subtype of ergative alignment that we (Gildea & Castro Alves 2010) have named nominative-absolute, in which pronouns have a distinctive nominative form while verbal prefixes index the absolutive. This pattern of alignment was previously thought to be impossible (Dixon 1994: 95), which raises two interesting historical questions: First, how did this (apparently rare) pattern come to be, and second, does anything about its origins help us to understand why it appears to be so rare?

It is also notable that the Cariban and Jê constructions with nominative-absolute alignment constitute over half of the attested cases in which constructions with ergative alignment patterns occur on the wrong side of TAM-conditioned splits. Many have observed that in such alignment splits, the ergative is typically found in the past tense and perfective aspect, to which Dixon (1994: 101) adds non-agent-oriented modalities and positive polarity. However, these examples of nominative-absolute alignment are conditioned also by nonpast tense, imperfective aspect, agent-oriented modalities, and negative polarity. By applying our methodology to exactly these typologically unusual constructions, we reconstruct not only the sources of the grammatical

patterns that are attested in the modern constructions, but also the conditions that allow them to arise on the "wrong" side of all the expected alignment splits. As such, this article also constitutes a contribution to the growing field of diachronic typology, specifically focused on the genesis of unusual alignment patterns in main clauses.

The main body of this article identifies the source of absolute cross-referencing in both the Cariban and Jê language families as possessive prefixes on nominalized verbs; the sources of tense-aspect-mood morphology are specific nominalizers, together with, in some cases, adpositions or reduced forms of older main verbs that took these possessed nonfinite verbs as arguments in the source constructions. Crucially, in both families, nominalized transitive verbs already have the option of expressing their notional A arguments in an oblique phrase. This source is attested as creating main clause ergative-absolute case-marking in the past tense in Timbira (Castro Alves 2010) and in multiple tenses and aspects in Cariban (Gildea 1998, 2012). However, in the modern nominative-absolute constructions, the uniquely case-marked A of the source nominalization is lost. In a subset of the source constructions that we reconstruct, it was suppressed in the source due to coreference conditions; we reconstruct the remainder of the source constructions with the ergative-marked A, but then we suggest that it is lost uniquely in the nonpast tenses, replaced by former topic pronouns that were reanalyzed as nominative pronouns.

The article is organized as follows. Section 2 introduces the Cariban and Jê language families, as well as the unusual nominative-absolute alignment pattern. Section 3 briefly introduces our theoretical framework, Diachronic Construction Grammar (DCxG), and situates our methodology within that framework. Sections 4–6 contain the details of the reconstructions: Section 4 characterizes the reconstructed nominalizations and their argument structure, Section 5 reconstructs source complex clauses in which the ergative A is suppressed due to coreference to the S/A of the matrix verb, and Section 6 reconstructs source complex clauses in which the ergative A was retained, but then replaced by topic pronouns and (unmarked) NPs in focus position. In the conclusions in Section 7, we suggest that the reconstructions help us to make sense of the typologically anomalous patterns identified in Section 2.

2   Nominative-Absolutive Alignment as a Puzzle for Reconstruction

To set the stage for the larger arguments, we first briefly introduce the Jê and Cariban language families (2.1), then some crucial properties of the nominative-absolute alignment pattern found in these languages (2.2).

2.1 *The Jê and Cariban Language Families*

The Jê language family consists of nine languages, many with named dialects in the literature, all spoken in Brazil (cf. Gildea & Castro Alves 2010: 174); these are in turn often linked to a larger set of languages (whose membership remains disputed) called the Macro-Jê stock (Rodrigues 2009; see classification of Jê in Castro Alves 2010: 439–40 and of Macro-Jê in Ribeiro & van der Voort 2010: 547). In this article, we focus on four of the six Northern Jê languages: Apinajé, Timbira (Canela Apãniekrá, Canela Ramkokamekrá, Krahô, Parkatêjê, Pykobjê, Krĩkatí), Mẽbêngôkre (Kayapó, Xikrín), and Kĩsêdjê (Suyá). We leave aside the fifth Northern Jê language, Panará, as it does not present the nominative-absolutive pattern and the sixth, Tapayúna, whose published work (Camargo 2015) we became aware of after this article was written. Our examples for these languages come from Castro Alves' field notes, and also from a number of published sources: for Apinajé, Oliveira (2003, 2005); for the Canela Apãniekrá dialect of Timbira,[2] Castro Alves (2004, 2008, 2010); for Mẽbêngôkre, Stout & Thomson (1974), Reis Santos (2001), Castro Alves & Reis Silva (2007), Salanova (2007, 2008); and for Kĩsêdjê, Santos (1997, 1999) and Nonato (2014). For additional data on these languages, cf. references in Castro Alves (2010) and Gildea & Castro Alves (2010: 175).

The Cariban language family consists of some 25 extant languages spoken in northern South America, primarily north of the Amazon in Brazil, Venezuela, and the three Guianas, with outliers spoken in Colombia and south of the Amazon in Brazil. In this article, we focus on two languages that belong to different genetic subgroupings (Gildea 2012: 442–446): Panare (Panare-Pemón Macro Group; Venezuelan Branch), spoken in Central Venezuela and Katxuyana (Parukotoan Group), spoken north of the Amazon in the Brazilian state of Pará and northern Brazil. Gildea has field notes for both languages, but the data we rely on here can all be found in prior publications: for Panare, Gildea (1998), Mattéi Muller (1994, 2007) and Payne & Payne (1999, 2013); for Katxuyana, Gildea & Castro Alves (2010).

For the purposes of the argument in this article, the details of the classifications are not important except in that they limit the depth of the reconstructions to the most restricted genetic unit that contains all of the languages in focus. All four of the Jê languages in focus belong to a sub-unit within Jê, which limits the scope of our reconstructions to the proto-language of that

---

2   Since all the Timbira data in this article come from a single dialect, Canela Apãniekrá, for the rest of the article we avoid the collective term, Timbira, in favor of the individual term, Canela. The reader should keep in mind that other Timbira dialects may present important differences to the patterns described here.

sub-unit, Proto-Northern Jê. In contrast, there is no reliable classification of the Cariban family that combines the Parukotoan Group and the Panare-Pemón Macro-Group in any smaller unit than the entire family; this allows us to utilize prior reconstructions of Proto-Cariban morphology and syntax (especially Gildea 1998, 2012) as a resource for identifying the source constructions in question in this article. We turn now to a characterization of the alignment pattern to be reconstructed.

2.2   *Alignment Typology and Nominative-Absolutive Alignment*

Alignment typology is the study of how languages code the basic clause-level semantic information of who did what to whom, observing how the single core argument of an intransitive clause (the S) aligns with one or both of the core arguments of a transitive clause (A & O).[3]

Basic morphosyntactic properties that languages generally use to distinguish core arguments are (nominal) case-marking, (verbal or auxiliary) indexation of person/number, and order of core argument constituents vis-à-vis the verb. Additional syntactic patterns are commonly added to this basic list, including constituency of core arguments *vis-à-vis* the verb, coreference with reflexive (4th person) morphology, coreference restrictions between core arguments of one clause and core arguments of another (either conjoined or subordinated to the clause in question), and analogical (often considered derivational) relationships between clause types (e.g. active versus passive, or main versus relative clauses). Having identified these grammatical properties for the S of an intransitive clause and the A and O of a transitive clause, we can then seek out the ways in which the properties of S **align** with those of A and/or O.

The label *nominative-accusative*, commonly called just *accusative*, describes the situation where S and A pattern together (the nominative) in opposition to the O alone (the accusative). In contrast, the label *ergative-absolutive*, commonly called just *ergative*, describes the situation where S and O pattern together (the absolutive) in opposition to the A alone (the ergative). There are a number of languages in which one or more of the core arguments for a lexically specified subset of verbs do not present the same grammatical patterns. In particular, a subset of A or O might be marked differently from the normal, or canonical patterns (e.g., dative-subject or locative-object), a pattern sometimes

---

3   We use Dixon's (1979, 1994) terms, S, A, and O, but without endorsing Dixon's assertions regarding their theoretical status as "semantico-syntactic primes" (cf. Queixalós & Gildea 2010, Haspelmath 2011 for discussion of problems with these labels). For us, they are merely a convenient way to exposit patterns that link transitive subject (A) and/or transitive object (O), with intransitive subject (S).

labeled *noncanonical marking* (Aikhenvald et al. 2001) or *semantic alignment* (Donohue & Wichmann 2008). This is also quite frequently found with the single core argument of intransitive verbs, where a subset of S might be marked like A and another subset like O, a situation labeled variously *Active-Stative, Active-Inactive, Agent-Patient, Split S, split intransitive* or, again, *semantic alignment.* These types are almost universally recognized in typological surveys and textbooks (e.g. Dixon 1979, 1994; Comrie 1989; Payne 1997; Givón 2001; Croft 2003; Dryer 2007; etc.).

Each of these labels describes a **pattern** that can be observed and described in any given **construction** in any given **language**. It is at the level of pattern that each of these alignments is a descriptive label: for example, it is relatively common for a nominal or pronominal case-marking pattern in a given construction, say, a past-tense clause type to be ergative-absolutive while verbal indexation is nominative-accusative. However, the label to be applied to the entire construction is not neutral: based on the case-marker, such a construction is labeled **ergative**, with no sub-label to recognize the accusative verbal indexation. The opposite combination, in which an accusative case-marking pattern co-occurs with an ergative or absolutive verbal indexation pattern, was not attested at the time this convention was established, cf. this quote from Dixon (1994: 95):

> Cross-referencing systems are thus basically pronominal (with the affixes having developed from free-form pronouns, in some earlier stage of the language). We would expect them to be on a nominative-accusative pattern, since this characterizes pronouns at the extreme left of the hierarchy ... What we can predict is that, if there is a 'split' of this kind, then bound prefixes will be accusative and case-marking on free forms will be ergative. This is exactly what is found.

However, we have identified multiple languages in which there are constructions characterized by Nominative-Accusative patterns everywhere except in verbal indexation, which is absolutive: this is the pattern we call *nominative-absolutive* (Gildea & Castro Alves 2010). As an initial illustration of the nominative-absolutive type, consider these clauses from Panare (1a–b) and Canela (2a–b).

(1) Panare nominative-absolutive (examples from Mattéi Muller 1994)[4]

   s-V   s.AUX   S
a. *y-u-të-sejpa*   *këj*   *kën*
   3S-S$_A$-go-FUTURE   3S.ANIM.COP   3.ANIM.DIST
   's/he will go'

   o-V   a.AUX   A
b. *y-ama-sejpa*   *këj*   *kën*
   3O-throw.away-FUTURE   3A.ANIM.COP   3.ANIM.DIST
   's/he will throw away it/him/her.'

(2) Canela nominative-absolutive (Castro-Alves 2004)[5]

   S   TAM   s-V   AUX
a. *wa* *ha*   *i-wrɨk*   *narɛ*
   1   IRLS   1S-descend.NF   NEG
   'I will not descend'

   A   TAM   o-V   AUX
b. *wa* *ha*   *iʔ-pɨr*   *narɛ*
   1   IRLS   3-grab.NF   NEG
   'I will not grab it (e.g., the knife).'

In both (1a) and (2a), S occurs as a free pronoun (postverbal in Panare, clause initial in Canela) and in Panare S controls auxiliary agreement; in both clauses, S is also indexed on the verb as a person-marking prefix. In both (1b) and (2b), A occurs as a free pronoun in the same clause location and case-form as the S pronoun, and in Panare, A controls auxiliary agreement; in both clauses, O is indexed on the verb as a person-marking prefix. In this initial illustration, A and O show clearly distinct patterns in the case form of pronouns (Canela), order and constituency (both Panare and Canela), auxiliary indexation (Panare), and/or verbal indexation (both Panare and Canela). S aligns with A in terms of pronominal case form (Canela), order (Panare and Canela), and auxiliary agreement (Panare), whereas S aligns with O only in the verbal index-ation forms. As such, S and A share a nominative pattern of pronominal case,

---

4 Orthographic symbols in Panare follow their IPA values with the following exceptions: *ë* [ə], *j* [h], *y* [j], *ch* [tʃ], *'* [ʔ], and vowel length indicated by doubling the vowel rather than a colon diacritic: *aa* [aː].

5 Orthographic symbols for Jê languages cited in this article take their IPA values.

constituent order, and auxiliary agreement, while S and O share an absolutive pattern of verbal indexation. These patterns are explored in some detail in Gildea & Castro Alves (2010) for the Jê languages Canela, Kĩsêdjê, and Apinajé, and for Cariban languages Panare and Katxuyana.

We turn now to the problem of labeling the alignment type of such constructions: for the previously attested mixtures, i.e., ergative case-marking alongside nominative agreement, the presence of ergative case-marking is sufficient for labeling a construction "ergative". In order to apply this precedent, we must determine whether the most important criterion for choosing the label is that the case-marking is ergative, or whether any ergative pattern anywhere in the construction would be sufficient. If the construction is to bear the label of the case-marking pattern, then the nominative versus accusative pronouns would force us to label these constructions "nominative-accusative". However, given the tradition of labeling certain constructions in Mayan languages as ergative even though case-marking is neutral, we conclude that the existence of any marked ergative morphological pattern justifies the use of the term "ergative", and as such, the existence of absolutive verbal indexation would require us to consider these constructions as representing a subtype of ergative (cf. Gildea & Castro Alves 2010: 161). Yet there is something odd about either solution in that (i) both patterns are present, and (ii) the patterns that are morphologically marked are exactly those that we expect to be unmarked: the nominative (unique pronominal case forms and auxiliary agreement) and the absolutive (verbal indexation prefixes). To prevent carelessly lumping these constructions in with either the accusative or the ergative type, and to index the categories that receive explicit morphological realization, we have adopted the hybrid label nominative-absolutive.

If we assume, as we did in Gildea & Castro Alves (2010: 192ff), that the presence of absolutive indexation makes these constructions a subtype within the ergative category, then we encounter a second typologically rare pattern in the tense-aspect-mood-polarity values that condition this subtype of ergative construction. In each language where it is attested, the nominative-absolutive alignment pattern is conditioned by specific values of tense, aspect, mood, and polarity, whereas other alignment patterns are conditioned by the remaining tense-aspect-mood-polarity values. This phenomenon is known as TENSE-ASPECT-BASED SPLIT ERGATIVITY, which itself possesses a characteristic pattern, as observed by Dixon (1994: 101):

> An ergative system is less likely to be employed when the clause refers to something that has not yet happened (in future tense), or is not complete

(imperfective aspect) or did not happen (negative polarity), or where the emphasis is on the agent's role (imperative or hortative moods).[6]

However, Gildea & Castro Alves (2010:191) show that the nominative-absolutive constructions in Cariban and Jê are mostly conditioned by exactly the wrong tenses, aspects, moods, and polarity, as indicated in Table 2.1.

The only tense value that conditions the nominative-absolutive alignment in our data is future, two of the four aspects are progressive and imperfective (the completive value is consistent with Dixon's expectation and the Panare "nonspecific" aspect is arguably neutral, cf. Payne & Payne 2013: 212–213), the desiderative mood certainly puts emphasis on the agent's role (whereas the "evaluative" moods emphasize the speaker's evaluation rather than the agent's role), and all negative clauses in the three Jê languages must occur in the nominative-absolutive alignment. It is remarkable that the nominative-absolutive construction is on the wrong side of every single one of the semantic values expected to condition ergative alignment.

These facts require some explanation. The nominative-absolutive pattern itself is sufficiently rare that its existence violates a proposed universal; one wonders why the pattern should be so rare, or, as its corollary, why it should exist at all. Similarly, one might ask why, in terms of typological correlations, the nominative-absolutive clause type in Cariban and Jê better matches the predicted distribution of a non-ergative alignment type. We believe that richer understanding may be gained from considering how an alignment type comes

TABLE 2.1 Tense-aspect-mood distinctions that condition nominative-absolutive

	Canela	Apinajé	Kĩsêdjê	Panare	Katxuyana
Tense			FUTURE	FUTURE	
Aspect	PROGRESSIVE	PROGRESSIVE	PROGRESSIVE	NONSPECIFIC	IMPERFECTIVE
	CONTINUOUS	CONTINUOUS	CONTINUOUS		
	COMPLETIVE		COMPLETIVE		
Mood	EVALUATIVE			DESIDERATIVE	
Polarity	NEGATIVE	NEGATIVE	NEGATIVE		

6 Dixon (personal communication) indicates that the one case of a polarity-based split-ergative has been disconfirmed in subsequent research, and as such, he would no longer include negative polarity in this passage.

into being. One would expect that very rare patterns arise from conditions that are equally rare, whether from unusual properties of expected source constructions, or from complex scenarios of change (cf. Givón 1979, 2009; Malchukov 2010; and for the same observation in the domain of phonology, Blevins 2004). Counter-examples to typologically general patterns also offer us the possibility of new perspectives on the forces that create the more typical patterns. We return to the question of explanation in Section 7, after having reconstructed the genesis of this alignment pattern in the two families where it is most widely attested, Cariban and Jê. The next section introduces a preliminary sketch of our methodology for reconstructing morphosyntactic patterns, which depends crucially on the notion of construction.

3   Diachronic Construction Grammar

We begin with the assertion that the relevant comparanda for reconstructing morphosyntax are not so abstract as schematic patterns, but are rather the much more concrete combination of morphemes and patterns that are packaged together in individual **constructions**. The framework of Construction Grammar (CxG) offers multiple ways to formalize the notion of construction (cf. Goldberg 2006, Ch. 10), but for our purposes, neither the details of the formalism nor the theoretical postulates associated with the framework are central.[7]

For these reconstructions, it is sufficient that constructions have both formal and semantic properties, and that these are arbitrarily linked to one another in a way analogous to the link between lexical form and meaning. Crucial to our work is that the formal and semantic properties of an innovative construction are not readily reducible to a combination of the formal and semantic properties of their component parts. Once we recognize the construction as the relevant synchronic unit, we can identify the formal and semantic properties of individual constructions in individual languages in search of similarities that are too consistent to be explained as products of chance. That is, we can identify cognate constructions.

As explained in the introduction to this volume, the identification of syntactic cognates is controversial, both in theory and in practice. Rather than enter into the theoretical debate, we choose to explain our method carefully, then show that it works. In brief, our theoretical definition is: Morphosyntactic

---

7   See Traugott & Trousdale (2013) and Barðdal & Gildea (2015) for detailed discussions of the diachronic implications of the formal architecture of CxG models.

cognates are constructions that are so similar in form and meaning that they logically must come from a common origin (the etymological meaning of the term *cognate*). Operationally, we seek to identify constructional cognates via their formal correspondences: we seek cognate grammatical morphemes, including personal pronouns, verbal indexation prefixes, case-markers, and tense-aspect-mood-polarity markers, both those found bound to verbs and those that are free words, such as the auxiliaries in Cariban and Jê. We also identify less concrete grammatical patterns (such as constituency or coreference restrictions) that are found in the syntactic constructions where these morphological cognates occur, identifying in particular those patterns that are identical or extremely similar across the distinct constructions in which the cognate morphology appears.

In addition, the semantics of candidate constructions must be either identical or relatable by known pathways of semantic change. Comparing candidate constructions, when we detect parallel morphology and syntax plus semantic identity or plausible semantic connection, we then consider the possible explanations for how they might have come to exhibit such similar form and meaning. The logical possibilities are coincidence, contact, or shared origin. When we can rule out coincidence or contact as plausible explanations, we conclude that the shared form and meaning must come from a common origin, and hence that the constructions are cognate.

In Section 4, we seek out the constructional cognates in both language families that contain the grammatical morphology and syntactic patterns that coalesce into the attested nominative and absolutive properties found in the various nominative-absolutive constructions.

The next step is then to explain the modern distributions of the relevant cognate constructions: did one modern construction give to rise to the others, or did all of them arise from some older construction that is no longer attested? This is the question of directionality, not in the absolutist sense of the unidirectionality debates in grammaticalization literature, but rather in the concrete details of each individual construction.[8] Here, we rely on three metrics.

---

8   As pointed out by a reviewer, the idea of "local directionality" (Willis 2011: 414–416) seems to be an intellectual antecedent to the kind of more concrete directionality that concerns us here. Our ideas reflect Willis' in that we are less concerned with possible vs. impossible changes and more with relative plausibility of specific hypothesized changes. Our ideas differ in how we argue for the relative plausibility of competing hypotheses about direction of change – in particular, we do not consider logical arguments about which source could have more readily created alternative constructions that might facilitate reanalysis during first language acquisition.

– First, which of these constructions is more widely attested in each family? To the extent that some constructions are widely attested in the family, we are forced to reconstruct them farther back than we are forced to reconstruct the others. This does not necessarily mean that they are older than constructions with more limited distribution, just that they are too widespread to be a relatively recent innovation, whereas constructions with more limited modern reflexes might be old, but they might also be quite recent.

– Second, in which constructions do the cognate morphemes in question occur with consistent meanings? To the extent that morphemes occur with consistent meanings in multiple, heterogeneous constructions, but have distinct meanings when they occur together in a single construction, this single construction is most likely to be an innovative reanalysis (cf. Gildea 1998: 39–41).

– Third, when the set of cognate constructions can be aligned with sets of constructions that have gone through an attested historical change in unrelated languages (e.g. locative constructions > progressive, cf. Heine 1994, Bybee et al. 1994, etc.), then we can appeal to the record of attested change to motivate positing a parallel change in the languages in question. Especially important to this final metric is that some changes appear to be unattested (e.g. a progressive suffix becoming a sequence of a nominalizing suffix followed by a locative postposition), and as such, no reconstruction should posit such a change without exceptionally good arguments of the first two types.

We turn now to the details of the reconstructions.

4   The Source of the Absolute Cross-Referencing: Possessed Nominalizations

Subordinate clause structures for Proto-Cariban and for northern Jê have been reconstructed in Gildea (1998) and Castro Alves (2004, 2010), respectively. These reconstructions are not controversial, as basically the same subordinate clause structures can be found synchronically in nearly all the modern languages of both families, making it fairly automatic to reconstruct them to the two proto-languages. In the remainder of this section, we briefly summarize the evidence in favor of reconstructing subordinate clause grammar to Proto-Cariban (Section 4.1) and Proto-Northern-Jê (Section 4.2).

4.1 *The Proto-Cariban Possessed Nominalization Construction*

Based on cognates found in 24 of the 25 extant languages, Gildea (1998: 119–128) reconstructs five distinct nominalizing suffixes, of which two play a role in this article: *-rɨ* 'ACTION NOMINALIZER' and *-ne* 'A NOMINALIZER'. Based on cognates found in Panare, Kapóng and †Tamanaku, Gildea (2003: 17) reconstructs the suffix *-cetɨ* 'NOMINALIZER' to the Proto-Venezuelan Branch of Cariban (for full classification, cf. Gildea 2012: 445). In addition, Gildea demonstrates that verbs bearing these nominalizations are all inalienably possessed by their notional absolutive (S of intransitive and O of transitive), expressed as a preceding free NP possessor in a tight possessor-possessed constituent or as a bound possessive prefix/proclitic on the nominalized verb. The notional A of the transitive nominalized verb need not occur (indeed, it cannot occur when the verb bears *-ne* 'A NZR'), but if it does, it can only be in an oblique phrase, marked in most Cariban languages by a modern reflex of the dative postposition *wɨja* 'AGT'.

Consider the illustrative examples of intransitive nominalized verbs in (3). In (3a), the intransitive verb *ehito* 'begin' bears the nominalizing suffix *-rɨ* 'NZR' and is possessed by the preceding NP *konoho* 'rain', its notional S; this entire NP is then the object of the postposition *me* 'ESSIVE'. In (3b), the intransitive verb *wu-të* 'INTR-GO' bears the cognate nominalizing suffix *-n* 'NZR' and is possessed by the preceding noun *waiki* 'deer', its notional S; this NP is then the object of the postposition *mënkai* 'like'. In both cases, the possessor and the possessed form a clear constituent, a tight NP.

(3) Nominalized intransitive clauses in Katxuyana (a) and Panare (b)

    [[PSSR   V-NZR]   P]PP
  a. [*konoho*   *ehitor*]   *me*]   *wahtawɨ,*   *tiihɨra.*
     konoho   e-hito-rɨ   me   wahtawɨ   t-iri-hɨra
     rain   DETR-begin-NZR   ESS   when.being   3-make-NEG
     'During **the beginning of the rains**, it is not done.'

                                                   (DV.Imoho yiitohu-003)
     (lit. 'when being **like the beginning of the rain**, it is not done')

    V   S   ADV [[PSSR V-NZR   ] P   ]PP
  b. *të-po-'ka'*   *kërënëpën tañe* [*waiki* *wu-të-n*   ] *mënkai*]
     go-ABIL-NEG dog   fast deer   INTR-go-NZR like
     'The dog does not go fast like the deer goes.' (Payne & Payne 2013: 100)
     (lit. 'The dog does not go fast **like the going of the deer**')

The prefix *wu-* 'INTR' in Panare is a reflex of a prefix found across the Cariban family on the S$_A$ subclass of nominalized intransitive verbs (Meira 2000: 205–206);[9] in Katxuyana this prefix has been lost with nominalizations that are possessed by the third person (cf. its absence in *ehitor* 'beginning' in 3a).

Now consider the representative examples of transitive nominalized verbs in (4), where the possessor is the notional O and the notional A (when it occurs) bears the dative postposition. In (4a), the transitive verb *y-ii-tohu* 'the making (of it)' bears the circumstantial nominalizing suffix *-tohu* and is possessed by the preceding noun, *imoho* 'field', which is the notional O; the notional A is the PP *katxuyana wïya* 'by the Katxuyana'. This entire NP is then the subject of the predicate PP *soro wara* 'like this'. In (4b), the transitive verb *y-uku-n* 'the eating of it' bears the nominalizing suffix *-n* and is possessed by the preceding noun *paaru* 'banana'; the notional A is expressed as the animate in visible third person pronoun *kën*, which occurs at the end of the sentence bearing the dative postposition *úya*.

(4)   Nominalized transitive clauses in Katxuyana (a) and Panare (b)
       [[PSSR V-NZR]NP   [NP   P]PP]SUBJ [[NP P]PP]PRED
    a. *Imoho yiitohu,*   *Katxuyana wïya,*   *soro wara.*
       imoho y-iri-tohu   katxuyana wïya   soro wara
       field   REL-make-CIRC.NZR Katxuyana AGT   this like
       'The making of (their) fields, by the Katxuyana, is like this.'

       [PSSR   V-NZR]NP   V   PRO$_{SUBJ}$ [NP- P]$_{PP}$
    b. *paaru   y-uku-n   pí-mpëj   chu   kën-úya*
       banana TRN-eat-NZR want-IMPRF.T 1SG   3.ANIM.INVIS-DAT
       'I want him to eat bananas.' (Payne & Payne 2013: 393)
       (lit. 'I want **the eating of bananas by him**.')

In the absence of the absolutive NP, the absolutive prefixes are identical to the prefix paradigms found on possessed nouns and postpositions, as seen in Table 2.2. Different allomorphs are conditioned by the initial segment of the possessed noun, consonant-initial (___-C) and vowel-initial (___-V). In Panare, forms possessed by the first person or by a preceding NP show leftward stress

---

9   Payne & Payne (2013: 234, 237) consider the underlying form of the prefix *wu-* to be *u-*, which
    they sometimes write with an asterix **u-*, as though it were a reconstructed form. They posit
    that it becomes *w-* when word-initial (i.e., when not preceded by a personal prefix). Given
    the widespread occurrence of *w-* across the family, we endorse Meira's (2000) reconstruction
    of **w-* rather than **u-*.

TABLE 2.2   Possessive prefixes in Katxuyana and Panare

	Katxuyana		Panare	
	___-C	___-V	___-C	___-V
1SG	∅-	y-	∅^-	y^-
2SG	o-	oy-	a-, o-	ay-, o-
3SG	i-	∅-	i-, yï-	y-, ty-
1+2	kï-	k-	–	–
preceding NP	NP ∅-	NP y-	NP ∅^-	NP y^-

shift (indicated by ^), thereby disambiguating first and third person when both show the same *y*- prefix (cf. Payne & Payne 2013: 75; Gildea 2012: 453–454). Note that many nominalized intransitive verbs take the prefix *w*- 'INTR' (Panare) / 'S$_A$' (Katxuyana), and as such take the allomorphs for consonant-initial stems; the exception is the second person prefix *o*- '2', which is a coalescence of the prefix *a*- with a following *u* or *w*.

Constructions with cognate morphology and identical argument structure patterns are found across nearly all modern Cariban languages (Gildea 1998: 104–127), making it vanishingly unlikely that the similarity is due to coincidence. Given the geographic dispersal of these languages, as well as the many differences in phonology and vocabulary, contact is an equally unlikely explanation. These very concrete constructions could only be so parallel because they come from a common origin. Accordingly, Gildea (1998) reconstructs this type of subordinate clause to Proto-Cariban. Before demonstrating the relevance of this reconstruction to the origins of the Cariban nominative-absolutive clause types, we turn to the parallel reconstruction of nonfinite clauses in Northern Jê.

4.2   *The Proto-Northern-Jê Possessed Nonfinite Verb Construction*

The key morphological element of the Jê reconstruction is the formal distinction between two forms of the verb, a shorter form that is generally considered the finite verb versus a longer form that is generally considered the nonfinite verb. This distinction is marked idiosyncratically in each verb, such that sometimes the verb pairs differ in initial consonant or vowel, sometimes by a final consonant or vowel, sometimes via entirely suppletive forms, and for a few verbs, there is no phonological distinction. To illustrate, we present the finite and nonfinite forms for a selection of verbs in Canela (Table 2.3), Kĩsêdjê (Table 2.4), Apinajé (Table 2.5), and Mẽbêngôkre (Table 2.6).

TABLE 2.3   FINITE versus NONFINITE verb forms in Canela (Castro Alves 2010: 448–449)

FINITE	NONFINITE	'gloss'	FINITE	NONFINITE	'gloss'
tʃa	tʃãm	'get up'	tʃa	tʃãm	'build'
ape	ʒpen	'work'	kura	kuran	'kill'
wrə	wrək	'descend'	tɔ	tɔn	'make'
aktʃa	piktʃar	'laugh'	apro	apror	'take'
amti	pimtir	'dream'	ŋɔ̃	ɔ̃r	'give'
amrã	ŋkwər	'cry'	tʃet	tʃet	'roast'
ŋɔ̃r	ɔ̃t	'sleep'	akʰɛp	akʰɛp	'cut'

TABLE 2.4   FINITE versus NONFINITE verb forms in Kĩsêdjê

FINITE	NONFINITE	'gloss'	FINITE	NONFINITE	'gloss'
tẽ	tẽm	'leave'	ku	kuru	'eat'
ŋgrɛ	ŋgrɛre	'dance'	pĩ	pĩrĩ	'kill'
rwə	rwək	'descend'	kõ	kõm	'drink'
nɔ̃	nɔ̃rɔ̃	'lay'	ŋɔ̃	ŋɔ̃tõ	'give'

TABLE 2.5   FINITE versus NONFINITE verb forms in Apinajé

FINITE	NONFINITE	'gloss'	FINITE	NONFINITE	'gloss'
tẽ	tem	'go'	ku	kur	'eat'
tʃa	tʃəm	'stand'	pɨde	pɨden	'capture, arrest'
amitɨ	pimtɨr	'dream'	pubu	pubuɲ	'see, look at'
jako	jakor	'smoke'	kura	kuran	'hit; batter; break'

TABLE 2.6   FINITE versus NONFINITE verb forms in Mẽbêngôkre

FINITE	NONFINITE	'gloss'	FINITE	NONFINITE	'gloss'
tẽ	tẽm	'go'	kurwa	kurwaj	'hit'
re	rere	'swim'	pumũ	pumũɲ	'see, look at'
ŋrɛ	ŋrɛrɛ	'sing'	arẽ	arẽɲ	'tell'
mua	mɤrɤ	'cry'	krẽ	krẽn	'eat'

At this point, two things should be clear. First, the distinction between finite and nonfinite verbs is robust and pervasive in these four Northern Jê languages: nearly all transitive and intransitive verbs have two distinct forms, readily distinguishable from one another. Second, there is no easily reconstructable morphological form that marks this difference – both forms of each individual verb will need to be reconstructed one by one to the common ancestor of these four languages, proto-Northern-Jê.[10]

For our purposes in this article, the data from these four languages is sufficient to establish that the morphological categories of FINITE VERB and NON-FINITE VERB are both reconstructable as morphological categories, even in the absence of a reconstruction of each verb form.

Having established the reconstructibility of the morphological category of nonfinite verbs in Northern Jê, we turn to the question of the syntactic environments in which we can encounter the category in each modern language. The term 'NONFINITE' has been chosen because in all four languages, this form occurs as a base for the further derivation of deverbal nouns and it also occurs as the nucleus of subordinate clauses, in particular those that function as an NP in a matrix clause.[11]

Castro Alves (2010: 458–463) presents synchronic examples from Canela, Apinajé and Mẽbêngôkre, and given that all three present virtually the same argument structure, she reconstructs that argument structure to Proto-Northern-Jê. In an almost exact structural parallel to the Proto-Cariban nominalized verb, the Proto-Northern-Jê nonfinite verb is inalienably possessed by its notional absolutive, which may be either a free NP that forms a tight constituent with the possessed nonfinite verb, or simply a bound pronoun, represented as either a pronominal clitic or a prefix. In contrast, the notional A, if it occurs at all, must be marked by a modern reflex of the postposition *tɛ* 'GENITIVE'.

We illustrate this structure with examples from Apinajé and Kĩsêdjê. In (5a), the intransitive nonfinite verb *tẽm* 'go.NF' is possessed by its notional S, the first person proclitic *ic-* '1'. In (5b), the cognate nonfinite verb *thẽm* 'go.NF' is possessed by the preceding free noun *i-nã* 'my mother'; the two form a constituent, as indicated by the brackets.

---

10   In fact, the FINITE-NONFINITE distinction is also attested in Southern Jê languages Shokleng (Urban 1985), and Kaingang (Wiesemann 1972, 2002), so it will surely reconstruct all the way back to Proto-Jê. Ribeiro & van der Voort (2010: 553) reconstruct a Proto-Amazonian-Jê "abstract nominalizer" *-r*, which may be implicated in the history of the nonfinite form, and they also identify a possible cognate *-r-* 'nominalizer' in the Macro-Jê language Karajá.

11   Santos (1999) goes so far as to argue that in Kĩsêdjê, the "long form" of a verb is not just a nonfinite verb, but could actually be analyzed as a lexical noun.

(5)   Nominalized intransitive clauses in Apinajé (a) and Kĩsêdjê (b)

                              (s-V)
                              [PSSR-PSSD]

a. *ịn-mɔ̃* [[*akunĩ   kot   ic-tem*] *ja*]   *pu-ba*   APINAJÉ
    1-DAT   woods   DIR   1-go.NF   DEF   RP-fear
    'I'm afraid of walking in the woods.' (lit. '**my walking**') (Oliveira 2005: 236)

                        (S   V.NF)
                        [PSSR   PSSD]

b. *hẽn*   ∅   [*i-nã   thẽm*]   *khãm*   *s-õmu*   KĨSÊDJÊ
    FACT   3   [1-mother   go.NF]   in   3-see
    'He/she saw my mother going.' (lit. '**my mother's going**')
                                          (Nonato 2014: 134)

In (6a), the transitive nonfinite verb *nõr* 'lie on' is preceded by its possessor, the notional O [*pikap ja*] 'the earth'. In (6b), the nonfinite verb *khuru* 'eat.NF' is possessed by the third-person prefix ∅- '3', which refers to the notional O, the thing that is eaten. In both (6a–b), the notional A of the nonfinite transitive verb is an oblique argument, labeled 'ERGATIVE' by the authors in question: *atɛ* '2.ERG' in (6a) and *i-nã=re* 'my mother=ERG' in (6b). Castro Alves (2010) reconstructs this ergative postposition to a marker of genitive.

(6)   Nominalized transitive clauses in Apinajé (a) and Kĩsêdjê (b)

                        (A   O   V.NF)
                        PP   [Possr   Possd]

a. *pa*   *na*   *pa* ∅-*ịnmɔ̃ atɛ*   *pikap ja*   *nõr*   APINAJÉ
    1.EMPH   RLS   1   1.DAT   2.ERG   earth   DEF   lie.on.NF
    prɔ̃m   *ket*
    want   NEG
    'It is I who don't want you to lie on the ground.' (Oliveira 2005: 86)
    (lit. '**the lying on of the earth by you**')

                        (A   o-V.nf)
                        [PP   PSSR-PSSD]

b. *hẽn*   ∅   [*i-nã=re   ∅-khuru*]   *khãm*   *s-õmu*   KĨSÊDJÊ
    FACT   3   [1-mother=ERG   3-eat.NF]   in   3-see
    'He/she saw my mother eating.' (Nonato 2014: 134)
    (lit. '**the eating of it by my mother**')

So for Northern Jê, we find in all four languages that the morphological category of nonfinite verb is robust and that the nonfinite verb is always possessed by

its notional absolute, leaving the notional A to occur as an oblique, marked by a modern reflex of the genitive postposition *tɛ*. Such parallel constructions, with cognate morphology and identical argument structure patterns, could not have arisen independently in closely related languages due to coincidence; while it is always possible that such a construction could have spread via contact, given the phonological differences between the cognate morphemes, contact is also not a plausible explanation. The Northern Jê nonfinite verb constructions could only be so parallel because they come from a common origin, which Castro Alves (2010) reconstructs to Proto-Northern-Jê.

At this point, the observant reader should note that, for both families, so far we have only reconstructed the morphological forms and/or categories of nonfinite/nominalized verbs and the grammatical treatment of their notional arguments. We have not yet demonstrated that these Proto-Cariban nominalized clauses or the Proto-Northern-Jê nonfinite clauses are in any way related to the nominative-absolutive clauses. The next two sections of the article demonstrate that a cognate to each of these nonfinite/nominalized constructions is found in every single distinct nominative-absolutive clause found so far in any Jê or Cariban language.

However, note that the notional A of these reconstructed nonfinite / nominalized constructions bears a distinctive postposition, which actually does yield an ergative case marker in other innovative constructions in each family (cf. Castro Alves 2010 for Jê; Gildea 1998, 2012 for Cariban). If the same subordinate construction is at the heart of the nominative-absolutive clauses, then either that oblique A must have simply not been realized in the biclausal source constructions or it must have been eliminated as a part of some subsequent change. In Section 5, we explore three biclausal source constructions that maintained the absolutive prefixes on the erstwhile nominalized verbs but, from the outset, arguably did not contain an independent expression of the A in its oblique form; instead, the matrix clause of each construction aligned the notional S and A, creating a nominative category. In Section 6 we explore biclausal source constructions in which the oblique A did occur originally, but was subsequently replaced by a topic pronoun in A and S roles, creating an innovative nominative category.

5 Eliminating the Oblique A in a Biclausal Source Construction

We now proceed to investigate the clearest cases, those in which a source biclausal construction already contains the nominative-absolutive pattern, just spread between two clauses rather than co-occurring in a single main clause predicate. In each case, the absolutive property of verbal indexation is inherited from

the source subordinate clause grammar reconstructed in Section 4.2. However, in these specific constructions, the option of expressing the oblique A in the subordinate clause is not exercised because it is coreferential with the subject of the matrix clause. In these same constructions there is also coreference between the main clause subject and the notional S of the subordinate clause, however, the morphological realization of the S is as an inalienable possessor, which cannot be suppressed. As such, there is either a lexical or a morphological form expressing S in both the matrix and the subordinate clauses.

In the most straightforward source construction, found only in the Jê languages, the matrix verb is a transitive verb of completing an action, the subordinate clause is a direct object complement of that verb, and the same participant is inherently the A of the matrix verb and the A/S of the complement clause (Section 5.1). A bit more complicated is the case, found in both families, in which an intransitive main clause takes an adverbial adjunct that contains the subordinate clause structure as the complement of an adposition, with the combination expressing aspectual notions like inceptive and completive, or modal notions like desiderative or intentional. In all these cases, the matrix clause beginner, finisher, desirer or intender is also the A/S of the desired/intended action, yielding the necessary coreference conditions to create a nominative pivot (Section 5.2). The third construction is quite idiosyncratic, and is limited to the two Cariban languages: there are actually two morphologically distinct source constructions, each readily reconstructable as independent, which then collapse into a single construction with a suppletive verb inflection and nominative-absolutive alignment (Section 5.3). In all three of these biclausal source constructions, there is an A/S pivot that creates a nominative pattern, conditioning both a single form of the S/A pronoun (in the matrix clause) and, in Cariban, main verb agreement with A/S. When the biclausal construction is reanalyzed as monoclausal, the new construction inherits its nominative patterns – pronominal case and/or auxiliary agreement – from the erstwhile matrix clause, while retaining the absolutive indexation from the erstwhile subordinate clause.

5.1   *The Phasal Matrix Verb Source*

This source is well-known from grammaticalization studies, where a phasal verb like 'begin', 'start', 'finish,' or 'stop' takes a nonfinite clause as its O; the A of the matrix phasal verb is always coreferential with the notional A or S of the nonfinite complement clause. After reanalysis, the erstwhile nonfinite verb is the main verb and the erstwhile aspectual verb becomes an aspectual auxiliary. Consider first these examples of cognate nominative-absolutive completive clauses from Canela (7a), Mẽbêngôkre (7b–c), and Kĩsêdjê (7d).

	A		o-V	AUX	
(7) a.	*wa*	*ha*	*iʔ-kʰrẽn*	*par*	CANELA
	1	IRLS	3-eat.NF	COMPL	

'I will eat it all'

			S	S-V	AUX	
b.	*arəp*	*nẽ*	*ba*	*i-jʌpeɲ*	*pa*	MĔBÊNGÔKRE
	already	NON.FUTURE	1	1-work.NF	COMPL	

'I already finished (my) working' (Stout & Thomson 1974: 71)

	A		o-V	AUX	
c.	*bir*	*nẽ*	*∅-krẽn*	*pa*	MĔBÊNGÔKRE
	Bir	NON.FUTURE	3-eat.NF	COMPL	

'Bir ate it all' (Stout & Thomson 1974: 71)

		A	[O	V	] AUX	
d.	*hẽn*	*wa*	*arə*	*i-t-ʌ*	*hwen*	*hwa*	KĨSÊDJÊ
	ASP	1	PAST	1-RP-thing	do.NF	COMPL	

'I already finished doing my work.' (Santos 1997: 91)

The following patterns should look familiar by now: in all four examples, the A and S are expressed as unmarked free (pro)nouns preceding the VP, which contains a nonfinite verb form immediately preceded by its absolutive argument, either as a bound prefix (7a–c) or a free O NP (7d).[12] Synchronically, the completive particle is the last element in the clause; it is analyzed as a particle because it is able to occur after the main verb, which usually is the final element of a clause, and because it takes a single invariant form.

In only one language, Kĩsêdjê, have we found a lexical verb that is cognate to this particle, the verb *hwa* 'kill, finish', used in the sense of 'kill all of' (8). Unfortunately, there is only a single example, and so we do not have evidence of a nonfinite form something like *hwar* 'finish.NF'.[13]

---

12   We added the gloss 'NF' to the non-finite form in (7c), which Stout & Thompson apparently overlooked.

13   In Canela the form *partu* is attested in alternation with the simple form *pa* – it is speculative, but not impossible, to imagine that this alternant contains an archaic reflex of a nonfinite form *par* followed by some unidentified element *=tu*.

(8) *ɲʌkɔ-n*   *kukwəy*   *hwa*   KĨSÊDJÊ
     Nháko-TOP   monkey   kill/finish
     'Nháko finished/killed the monkeys (until they were all gone)'
                                                                              (Santos 1997: 92)

Despite the absence of this form as a lexical verb in the other languages, it
is relatively easy to find examples of complement-taking verbs with phasal
meanings, such as the modern examples from Canela with the main verbs *iku*
'stop' (9a–c). The monoclausal main verb use of *iku* is illustrated in (9a), where
it is preceded by the O *pɜrkrɛ* 'boat', and its use as the matrix verb of a biclaus-
al phasal verb construction is illustrated in (9b–c). Note that the grammar of
the two biclausal constructions is exactly parallel to that of the nominative-
absolutive clauses: the subject of the transitive phasal matrix verb occurs first
as an unmarked pronoun; the phasal matrix verb is final, preceded by its ob-
ject, the nonfinite verb plus its absolutive argument (S in (9b), O in (9c)). The
absolutive is the same obligatory possessor that we reconstruct in Section 4,
expressed either as a possessive prefix (9b) or a free noun possessor imme-
diately preceding the nonfinite verb (9c). Given the condition of coreference
with the matrix clause A, the oblique notional A of the nonfinite verb is not
(and cannot be) expressed. Thus, the visible argument structure of the entire
construction is as follows: The notional A of the nonfinite verb is expressed
only via the explicit A of the phasal matrix verb, the notional O of the nonfinite
verb is expressed only as the possessor of the nonfinite verb, and the notional S
of the nonfinite verb is expressed twice, once as the A of the phasal matrix verb
and once as the obligatory possessive prefix on the intransitive nonfinite verb.

(9)   Canela aspectual verb example:
        A   [N$_O$ V$_{TR}$]$_{VP}$
   a. *wa*   *pɜrkrɛ*   *j-iku*   CANELA
        1   boat   RP-stop
      'I stop [the boat].'

        A   [[POSSR-V.NF]$_O$ V$_{TR}$]$_{VP}$
   b. *ka*   *ha*   [*a-j-ɜpən*]   *j-iku*   CANELA
        2   IRLS   2-RP-eat   RP-finish
      lit. 'you will stop [your$_i$ eating]$_O$' (Castro Alves 2004: 134)

        A   [[POSSR V.NF]$_O$ V$_{TR}$]$_{VP}$
   c. *ka*   *ha*   [*hĩ*   *kur*]   *j-iku*   CANELA
        2   IRLS   meat   eat.NF   RP-stop
      lit. 'you will stop [the meat's eating].' (Castro Alves 2004: 134)

At this point, we assert that this combination of cognate morphology, order of elements, and argument structure patterns could not have arisen via chance. As seen in examples (7a–d), the grammar of the nominative-absolutive completive clause is identical across the three northern Jê languages, and the postverbal particle is still attested as a lexical verb meaning 'kill/finish' in one of these languages (8). As seen in examples (9b–c), the grammar of biclausal constructions with phasal matrix verbs is virtually identical to the nominative-absolutive completive clause, the only difference being the lexical status of the final element, a transitive verb in (9a–b), but an aspectual particle in (7a–d). It is straightforward to make the argument that the constructions in (7) are too similar in both form and meaning for the parallels to have arisen by chance, so we must posit that they have arisen from a common source. The parallels between the grammar of (7) and the synchronically biclausal phasal construction in (9b–c) are also too similar to be due to chance, suggesting parallel source constructions. The completive particle is cognate to a lexical verb (8) and it occurs in exactly the same clausal location as the phasal matrix verb in (9a–b). As such, we now have sufficient evidence to posit a biclausal proto-construction (10), parallel to that in (9a–b), which gave rise to the modern completive constructions in (7).

The example in (10a) is exactly parallel to the one in (9b), with the nonfinite intransitive verb *i-jʌpeɲ* 'my working' occurring as the object complement of *pa* 'finish'. Although the 'finisher' and the 'worker' in both clauses is the same referent, the prefix in the subordinate clause cannot be deleted or suppressed due to coreference because the nonfinite verb is obligatorily possessed. In (10b), parallel to (9c), the nonfinite transitive verb *iʔ-kʰrẽn* 'its eating' is the object complement of *pa* 'finish'. The oblique notional A argument of the nonfinite clause does not occur due to the coreference with the A of the main clause verb. The result is a clause in which the sole reference to the A of both verbs is the unmarked pronoun that belongs grammatically to the main verb 'finish', whereas the possessive prefix on the transitive nonfinite verb references only its O. We repeat the examples in English calques in (11a–b), to illustrate the coreference conditions in a form that will be more familiar to most readers.

(10) A reconstructed source for the completive clause in Proto-Northern Jê

			A	[s-V-NF]ₒ	V
a.	*arəp	nẽ	ba	i-jʌpeɲ	pa
	already	NON.FUTURE	1	1-work.NF	finish

'I already finished my working' (based on 7b above)

	A		[o-V-NF]$_O$	V
b.	*wa	ha	iʔ-krẽn	par
	1	IRLS	3-eat.NF	finish

'I will finish eating it' (lit. 'its eating') (based on 7a above)

(11) a. I$_i$ finished [my$_i$ working]$_O$ (matrix A = Sbrd S; Sbrd possessor = S)
      b. I$_i$ will finish [∅$_i$ its$_j$ eating]$_O$ (matrix A = Sbrd A; Sbrd possessor = O)

After reanalysis, the argument structure has both a nominative subject (< A of the matrix complement-taking verb) and an absolutive verbal prefix (< obligatory possessive prefix on the nonfinite V); the matrix transitive verb becomes the auxiliary particle, *pa*, and the erstwhile subordinate verb – still in its morphologically nonfinite form – becomes the new main verb. The combined argument structures in the source clearly provide both the nominative S/A and absolutive S/O patterns, so after reanalysis, no further adjustment is necessary to create the nominative-absolutive alignment.

As a postscript to this section, we present the less consistent cognate construction in Apinajé as a likely case of constructional innovation in progress. Oliveira (2005) does not explicitly describe the grammar of the completive construction, but her dictionary in Appendix C lists the form *pa* 'CONCLUSIVE; COMPLETIVE', and there are 39 examples scattered throughout the grammar and the appended texts. Of these, 11 examples contain five transitive verbs, two intransitive verbs, and four descriptive verbs that have no morphological distinction between finite and nonfinite. Also, four examples contain three verbs (one of which occurs twice) for which Oliveira does not indicate whether the form is finite or nonfinite and they occur nowhere else in the work. Of the remaining 24 examples, 13 examples contain eight transitive verbs (two of which occur twice) and three intransitive verbs in the expected cognate nonfinite form, accompanied by the nominative-absolutive grammar, as seen in (12a–b).

					S-V.NF	AUX	
(12) a.	*nẽ*	*əbri*	*prɛ*	*me*	*i-j-akrɛn*	*pa*	APINAJÉ
	CNJ	then	PAST	PL	1-RP-go.by	CNCL	

'then they passed us by' (Oliveira 2005: 330)

		A	O	V.NF	AUX			
b.	*kɔt*	*paj*	*kagə*	*n-ipetʃ*	*pa*	*ri*	*kɔtmã apku*	APINAJÉ
	IRLS	1.IRLS	mark	RP-make	COMPL	LOC	still	eat.INTR

'I will eat upon/when I have finished studying' (Oliveira 2005: 291)

However, the remaining 11 examples are surprising in that the main verb occurs in the finite form instead of the expected nonfinite form preceding *pa* 'COMPL'. Ten of these unexpected examples are of transitive verbs and only one is an intransitive verb. We illustrate these surprising patterns via the examples in (13a–b): in (13a), the finite form of the intransitive verb *tʃa* 'stand' precedes the particle *pa* and in (13b), the finite form of the transitive verbs *krẽ* 'eat' and *tʃi* 'put' each precede the particle *pa*.

(13) a.
			∅	V	AUX			
anigrɔ	*mã*	*ɲum*	∅	*tʃa*	*pa*	*nẽ*	*əgɨw*	APINAJÉ
daylight	DAT	DS	∅	stand	COMPL	CNJ	starch	

nõ
lie
'let it sit under the sun so that the tapioca will go all down to the bottom.'
(Oliveira 2005: 355)

b.
O		V	AUX					
miti		*krẽ*	*pa*	*nẽ*	*kir*	*kamã*	*kə,*	*krã,*	APINAJÉ
alligator		eat	COMPL	SS	roast	LOC	skin	head	

i	*tʃi*	*pa*
bone	put	COMPL

'(they) ate the alligator and put its skin, head, bones all into the roasting place.'
(Oliveira 2005: 260)

At this point, we must give some historical account of the completive construction in Apinajé, which has main verbs preceding *pa* in both the expected nonfinite form and in the surprising finite form. We can imagine three possible theories, of which the only one that seems plausible departs from our reconstruction above. First, we might posit a proto completive construction in which the main verb was finite; in this case, the other three Northern Jê languages would have entirely replaced the finite forms with nonfinite forms, and Apinajé has nearly completed this replacement as well, but still retains the original finite form in some (currently undefined) contexts. Second, we might posit a proto completive construction in which the main verb could be either finite or nonfinite, as currently attested in Apinajé; in this case, Apinajê would conserve the proto system and the other three languages would have converted the subset of finite forms to nonfinite. We find both scenarios unlikely because the completive construction is clearly a type of main clause in the synchronic grammar of all four languages, with no obvious drift towards becoming subordinate; as such, we see no motive for the spontaneous replacement of the finite form with a nonfinite form.

The only plausible hypothesis is that the matrix verb *pa* 'finish' became the completive particle *pa* already by Proto-Northern Jê (and perhaps even earlier, although testing this hypothesis would require expanding the comparative data to include the other branches of the family); the nonfinite form of the main verb and its accompanying nominative-absolutive argument structure were retained without change in Canela, Mẽbêngôkre, and Kĩsêdjê. In Apinajé, a subset of the completive clauses continues to conserve the proto-structure, but a new construction has arisen in which a finite verb may now occur preceding *pa* 'COMPL'. Our *post hoc* explanation for this innovation is that (i) speakers no longer think of *pa* as a verb (presumably this would be true in all four languages) and (ii) they have subconsciously drawn an analogy between the nonfinite main verbs of this construction and the finite main verbs that occur in many other constructions. This condition of analogy then allows the extension of the finite verb forms (with their accompanying argument structure) into the completive construction, creating a new completive construction that we assume must be distinct in use from the original construction, whether semantically, pragmatically, or stylistically. Obviously, this would be a fascinating topic for further synchronic research in Apinajé.

5.2   *The Adverbial Predicate Sources*

Our second source is also well-known from grammaticalization studies, in which the matrix clause predicates a location or other adverbial property of the subject, and in which the adverbial predicate consists of the subordinate (nonfinite / nominalized) clause embedded in a postpositional phrase. This is a common source of progressives (Heine 1994; Bybee et al. 1994), and is implicated here in the Panare DESIDERATIVE and FUTURE, as well as the northern Jê PROGRESSIVE, CONTINUATIVE, and INGRESSIVE.

5.2.1   In Northern Gê

We begin with the progressives, which are based on two models. The first model, found in Canela (14a–c), Apinajé (15a–b) and Mẽbêngôkre (16a–b) takes an intransitive main verb of motion, *mõ/mɔ̃* 'go', preceded by the cognate form *tɔ* 'PREVERB (PV)' (Canela), *ɔ* 'INSTR' (Mẽbêngôkre), or *ɔ* 'do' (Apinajé). Note the expected argument structure: the main clause S and A are not required to occur (cf. 14a, 15b), but if they do, they occur early in the clause as free pronouns (14b, 15a, 16b) or as unmarked nouns (16a). The main verb appears in its nonfinite form, preceded by either the absolutive prefix (14a–b, 15a–b, 16a–b) or a free NP O (14c).

	S-V.NF	[AUX]			A	O-V.NF	[AUX]	
(14) a.	*iʔ-ŋkrə*	*tɔ=m�*		b.	*ka*	*i-kakʰwin*	*tɔ=mɔ̃*	CANELA
	3-dry	PV=go			2	1-hit.NF	PV=go	
	'It is drying (the cashew tree).'				'You are hitting me'			

	A	O	V.NF		[AUX]		
c.	*i-kʰra*	*ko*	*tɔjkʰɔ̃m*		*tɔ*	*mɔ̃*	CANELA
	1.child	water	PV:3:drink.NF		PV	go	
	'my son is drinking water.'						

		S		S-V.NF	[AUX]		
(15) a.	*na*	*pa*	*ra*	*ic-tik*	*ɔ*	*mɔ̃*	APINAJÉ
	RLS	1	ASP	1-die.NF	do	go	
	'I'm dying' (Oliveira 2005: 293)						

			S-V.NF		[AUX]		
b.	*na*	*ra*	*∅-kʌ*		*ɔ*	*mɔ̃*	APINAJÉ
	RLS	ASP	3-mature.NF		do	go	
	'It's getting ripe already.' (Oliveira 2005: 294)						

	A		O-V.NF	[AUX]		
(16) a.	*bir*	*nẽ*	*∅-krẽn*	*ɔ*	*mɔ̃*	MẼBÊNGÔKRE
	Bir	N.FUT	3-eat.NF	do	go	
	'Bir is eating it (moving)' (Stout & Thomson 1974: 71)					

	A	O-V.NF		[AUX]		
b.	*ba*	*kwũɲ*		*ɔ=mɔ̃*		MẼBÊNGÔKRE
	1NOM	3.break.SG.NF		INSTR=go.PL		
	'I am breaking it' (Salanova 2007: 55)					

The synchronic analysis of the auxiliary complex differs between the languages not because the grammatical behavior is different, but because three different analyses have been proposed for the cognate form *tɔ/ɔ* preceding *mõ/mɔ̃* 'go'. In all three languages, the existing descriptions distinguish between the instrumental postposition *tɔ/ɔ* 'INSTR' and the homophonous preverbal causativizer *tɔ/ɔ* 'CAUS'. To these two meanings, Castro Alves (2015) adds the homophonous form *tɔ* 'PREVERB (PV)' and both Stout & Thomson (1974) and Oliveira (2005: 415) consider this specific construction to contain the homophonous transitive verb *ɔ* 'do', which immediately precedes other verbs (such as *mõ* 'go' in these examples) in a serial verb sequence. Regardless of the preferred

synchronic analysis, looking at these examples, the similarity in morphological form and syntactic location cannot be an accident – these constructions clearly all must be modern reflexes of a single source construction.

We suggest that this historical source construction most likely contained its invariant morphemes with the consistent values that they currently show when not combined in this construction: the main verb must have been *mɔ̃* 'go', the form *ɔ* must have been the instrumental/locative postposition, the nonfinite verb form must have been the object of this postposition, and the unmarked subject (pro)noun must have been the subject of the verb *mɔ̃* 'go'. This source is modeled in (17a–b), giving the etymological analysis of the Canela examples in (14a–b). Note that once again, given the condition of coreference between the S of the matrix clause and the notional A of the nonfinite verb, this notional A does not occur explicitly, thereby leaving the superordinate S as the only grammatical form that references the notional A.

	S	[PSSR-V]$_{NP}$	P$_{PP}$	V		S	[PSSR-V]$_{NP}$	P$_{PP}$	V
(17) a.	∅	*iʔ-ŋkrə*	*tɔ*	*mɔ̃*	b.	*ka*	*i-kakʰwin*	*tɔ*	*mɔ̃*
	∅	3-dry.NF	PV	go		2	1-hit.NF	PV	go
	'(It) goes with/to its drying'					'You go with/to my hitting (the hitting of me)'			

Clearly, this has become a new construction (in the sense of Goldberg 2006) in both Apinajé and Canela in that the semantics of the construction are not derivable from the semantics of the component words & morphemes: specifically, there is no longer any element of translational movement in either the Canela or the Apinajé examples, and the instrumental/locative semantics are completely absent in all three languages. After reanalysis, the argument structure has both a nominative subject (< S of the matrix verb of motion) and an absolute verbal prefix (< obligatory possessive prefix); the matrix intransitive verb becomes the auxiliary *mɔ̃*, and the erstwhile subordinate verb – still in its morphologically nonfinite form – becomes the new main verb. In the source, the argument structure of the matrix verb provides the nominative S/A patterns and the argument structure of the nonfinite verb provides the absolute S/O pattern, so after reanalysis, no further adjustment is necessary to create the nominative-absolutive alignment.[14]

---

14   The same structure except with a different verb of motion, *tɛ̃* 'go', gives rise to an alternative progressive in both Canela and Kĩsêdjê, illustrated in passing in Gildea & Castro Alves (2010: 177, 185).

We turn now to the continuative/progressive constructions based on posture verbs, found in all four languages. We illustrate this with the auxiliary verbs *ta/tʃa/dʒa* 'be standing' (18) and *ɲĩ* 'be sitting' (19). In the now-familiar pattern, the nominative argument is expressed via an unmarked noun or pronoun occurring in initial position (although it can be preceded by initial TAM particles, cf. 19a, d, and also by left-dislocated topic pronouns, cf. 19c), whereas the absolutive argument is either a free noun O (18b, d, 19c–d) or an absolutive prefix on the verb (18a, c, 19a–b).

(18) Based on the auxiliary *ta/tʃa/dʒa* 'be standing'

	S	s-V.NF	[	AUX	]		CANELA
a.	*wa*	*i-tʃwər*		*tɔ=*	*tʃa*		
	1	1-bathe.NF		PV=	stand		

'I am bathing myself'

		A	[O	V.NF]	[AUX]		KĨSÊDJÊ
b.	*kaomi*	*ra*	*aŋgro*	*pĩrĩ*	*mã*	*tɔ*	*ta*
	Kaomi	SM	pig	kill.NF	?	do	be.standing

'Kaomi is killing the pig' (Santos 1997: 514)[15]

	S		s-V.NF	[AUX]		MẼBÊNGÔKRE
c.	*mẽnire*	*nẽ*	*∅-tɔrɔ*	*ɔ*	*dʒa*	
	woman	RLS	3-dance.NF	do	be.standing	

'The woman is dancing' (Castro Alves & Reis Silva 2007)

	A	[O	V.NF]	[AUX]		MẼBÊNGÔKRE
d.	*ga*	*tɛp*	*krẽn*	*ɔ*	*dʒa*	
	2	fish	eat.NF	do	be.standing	

'You are eating fish (standing)' (Castro Alves & Reis Silva 2007)

(19) Based on the auxiliary *ɲĩ* 'be sitting'

	S	s-V	[Aux]		KĨSÊDJÊ
a.	*hẽn*	*wa*	*i-mbərə*	*rɔ*	*ɲĩ*
	ASP	1	1-cry.NF	do	be.sitting

'I am crying' (Santos 1997: 87)

---

15   We maintain Santos' gloss 'SM' (Subject Marker) on the particle *ra*, but we question this analysis (Section 6.3). Its similarity to the demonstrative pronoun *ta* in other northern Jê languages suggests a possible analysis as a definite marker.

     S   S-V   [AUX]   MĚBÊNGÔKRE

b. *mẽnire*   *nẽ*   ∅-*tɔrɔ*   *ɔ*   *ɲũ*

   woman   RLS   3-dance.NF do be.sitting

   'The woman is dancing (sitting)' (Castro Alves & Reis Silva 2007)

     TOP   A   [O   V]   [AUX]   KĨSÊDJÊ

c. *pa-n*   *wa*   *tɛp*   *kuru*   *rɔ*   *ɲĩ*

   1-TOP   1   fish   eat.NF   do be.sitting

   'I who am eating the fish' (Santos 1999: 232)[16]

        A   [[O   ]   V   ]_{VP} [AUX] APINAJÉ

d. *na*   *pa kɔt=mɔ̃ i-ɲ-õ*   *pĩ*   *katprɛ*   *ɔ*   *ɲĩ*

   RLS 1   still/yet 1-RP-GEN wood fasten.NF do sit

   'I'm still fastening my wood' (Oliveira 2005: 294)

     A   O-V.NF   [AUX]   CANELA

e. *humrɛ*   *iʔ-kʰrẽn*   *tɔ=* *hɨ̃r*

   man   3-eat.NF   PV= be.sitting

   'The man is eating (something) (sitting)'

Once again, we reconstruct the sources of these clause types beginning with the morphological cognates: clearly the auxiliaries are cognate with the main verbs for 'stand' and 'sit', plus the postposition *tɔ/rɔ/ɔ* 'INSTRUMENTAL/ LOCATIVE' with its object, the nonfinite form of the main verb possessed by its absolutive argument. In the source construction, the subject of the posture verb is coreferential with the notional S/A of the nonfinite verb, leading to coreference with the obligatory notional S possessor of the intransitive nonfinite verb, but conditioning the absence of the optional oblique notional A of the transitive nonfinite verb. After reanalysis, the former S of the intransitive posture matrix verb becomes the S/A of the innovative clause; the former possessor of the nonfinite verb becomes absolutive indexation on the main verb. While the source is clear, in this construction, the semantic evidence for reanalysis is somewhat scant, given that the semantics of the entire construction

---

16   The translation is a calque of the Portuguese translation given by Santos: 'Eu que estou comendo o peixe.' In colloquial Brazilian Portuguese, this construction is used to place focus on the subject of a clause, a meaning consistent with the initial topic-marked pronoun, which is coreferential with the following A pronoun.

could be derived by combining the semantics of the component morphemes (i.e., the translations indicate that the subject is understood as being in the posture indicated by the verb, that is, seated or standing, while doing the action that is in progress).

5.2.2 Predicate Adverbials in the Cariban Family

In the Cariban family, predicative adverbials have given rise to the future tense in both Panare and Yukpa and to the desiderative mood in Panare, each with the same nominative-absolutive argument structure. Consider first the future tense, as seen in (20a–b). In (20a), the intransitive main verb *ariki* 'end' comes first, followed by the copular auxiliary and the unmarked S pronoun. The verb bears the future tense suffix *-sejpa*, the intransitive verb class marker *w-* 'S_A', and the third person absolutive prefix *y-* '3'; the nonverbal auxiliary agrees with the third person S for animacy. In (20b), the transitive main verb *ama* 'throw' comes first, followed by the unmarked free pronoun A *yu* '1SG'. This verb also bears the future suffix *-sejpa* plus the third person absolutive prefix *y-* '3'. It is standard for there to be no copula with first and second person subjects, so the absence of an auxiliary in (20b) is not surprising.

	s-V		s.AUX	S		o-V	A	
(20) a.	*yuri'chejpa*		*kë'*	*kamënton*	b.	*yamasejpa*	*yu*	PANARE
	y-w-ariki-sejpa		kë'	kamënton		y-ama-sejpa	yu	
	3-INTR-end-FUT		3.COP	they		3-throw-FUT	1SG	
	'they (= their family line) will be					'I will throw it.'		
	finished.' (Mattéi Muller 1994: 21)					(Mattéi Muller 1994: xxxii)		

To identify the source grammar of the Panare future tense, we first identify the individual morphemes in the construction. The auxiliary is clearly identical to the nonverbal copula (cf. Gildea 1993a–b) and the prefixes on the main verb are those expected on a nominalization: the verb class marker *w-* 'INTR' is only found on nominalizations or participles in multiple Cariban languages (cf. Meira 2000: 205–208) and the absolutive person prefixes are identical to the possessive prefixes. The suffix *-se'pa* represents three etymological morphemes. The first, **-ceti* 'NOMINALIZER' (Gildea 2003: 17) has only a few reflexes in modern Cariban languages: only in Tamanakuit is it attested as a nominalizer without additional suffixal material; it is attested preceding the adverbializing suffix *-pe* 'ESSIVE' in the Tamanaku adverbial suffix *-chetpe* 'PURPOSIVE' (Gilij 1965[1780–1784]/III: 30) and the Panare medial clause suffix *-sehpe* 'HC.

AGT.SEQ' ('high participant continuity, agent-oriented, sequential', Payne & Payne 2013: 404–405).[17]

The composite suffix *-cetï-pe* is also attested as being nominalized by means of the suffix *-ano* 'NZR', which causes the final vowel of *-pe* to shift to *a*: Tamanaku *y-are-chet-pa-no* 'one with the purpose of taking' (Gilij 1965[1780–1784]/III: 264). The cognate nominalized form in Panare has undergone two regular phonological changes: first, final vowels of nouns generally delete, creating consonant-final forms (Gildea 2003: 18–20); second, the nasal component of the nominalizing suffix is lost, leaving the vowel shift as the only morphological indication of nominal versus adverbial status, As such, a final *-e* indicates adverbial forms and a final *-a* indicates nominal forms.[18] Thus, the Panare suffix *-sehpa* 'FUTURE' represents the reduction of an earlier form *-ceti-pa-no* 'one with the purpose of V-ing'.

With all these morphological cognates in hand, it is relatively straightforward to combine them into the source construction for the Panare future: the matrix clause was a predicate nominal construction with a nominalization as the predicate noun. This predicate nominalization has as its core a verb root; this root is nominalized, making it obligatorily possessed by the notional absolutive. This possessed nominalization bears the essive suffix (itself a reduced form of the reconstructable essive postposition), resulting in a form that is attested synchronically as the medial/adverbial suffix *-sehpe*. This adverbial form is itself nominalized, creating a form with an etymological meaning something like 'one who is with/like the purpose/destiny of doing the act described by V', or more briefly, 'one that is destined to V'. Such a meaning is eminently compatible with the interpretation that the individual will undertake the act in the future, so both the form and the meaning of the future tense construction in Panare follow naturally from the forms and meanings in the cognate nonfinite

---

17   An anonymous reviewer suggests that our *-cetï* could also be cognate with *-ketï* 'S.NZR', another nominalizer with very limited distribution in the family (attested in only Tiriyó, Apalaí, and Makushi). We hope to consider this possibility more carefully in future research, as it would be most welcome to identify cognates to this form outside the inner circle of Venezuelan languages in which it is currently found.

18   This same vowel alternation characterizes adverbial versus nominalized forms of other adverbializing suffixes, including *-ke* 'PROPRIETIVE', and *-re/-ye* 'ADVERBIAL'. Payne & Payne (2013: 124–125) illustrate these patterns, but assume that the nominal *a* form is basic, from which the *e* form adverb is derived (although they acknowledge the alternate possibility on p. 118). Mattéi Muller (2007) argues that, for certain verbal suffixes, this vowel alternation has become an indicator of a temporal and epistemic distinction, with the *e* forms generally indicating less certain / temporally stable meanings and the *a* forms indicating more certain / temporally stable meanings.

constructions attested in the other languages. This proposed etymological analysis is laid out in (21a–b).

(21) a. [[[PSSR-V-NZR]$_{NP}$ P]$_{PP}$-NZR]$_{NP}$ S.COP   S
     yuri'chejpa   *kë'*   *kamënton*
     y-w-ariki-sej-pe-a   kë'   kamënton
     3-S$_A$-end-NZR-ATTR-NZR   3.COP   they
     'they (= their family line) is one that is like destined to end.'

<div align="right">(Mattéi Muller 1994: 21)</div>

     PSSR-V-nzr]$_{NP}$ P]$_{PP}$-NZR]$_{NP}$ S
  b. *yamasejpa*   *yu*
     y-ama-sej-pe-a   yu
     3-throw-NZR-ATTR-NZR   1SG
     'I (am) one who is like destined to throw it.' (Mattéi Muller 1994: xxxii)

Following the reanalysis, we have the familiar matrix clause S > nominative argument and subordinate clause possessor > absolutive argument, with the copula (when one occurs) agreeing with the former matrix clause S > nominative.

A similarly complex set of morphemes coalesces into the desiderative suffix, which is shown in (22) to occur in a standard nominative-absolutive construction.

(22) a. S-V   S-AUX
     witëjtëpa   *waasɨn*   PANARE
     ∅-w-të-jtë-pa   w-a'-sɨn
     1-S$_A$-go-DESID-DUR   1S-AUX.PAST-REL
     'I wanted to go.' (Mattéi Muller 1994: 76)

  b. O-V   A.AUX   A
     atyajtépe   *këj*   *kën*   PANARE
     a-tya-jté-pe   këj   kën
     2-hear-DESID-TEMP   3.COP   3.ANIM
     'He wants to hear you (right now)' (Payne & Payne 1999: 123)

There are actually several allomorphs of the Panare desiderative whose synchronic distribution (and possible meaning difference) remain poorly understood: *-jtë, -jtëpa, -jtëpe* 'DESIDERATIVE' (from Mattéi Muller 1994) and *jtépe, -jtépi, -jtépa* 'DESIDERATIVE' (from Payne & Payne 2013: 217–220). The fourth set of allomorphs, *-jtë'ka/-jtéka* 'NEGATIVE.DESIDERATIVE' (from Mattéi

Muller 1994 and Payne & Payne 2013, respectively) is clearly negative. There appear to be four distinct etymological units here, with variation in the pronunciation of the first and final vowels perhaps representing dialectal variation: the root of every variant of the desiderative suffix is *-jtë/-jté* 'one that wants', which is attested as occurring alone (in Mattéi Muller's corpus), or (in both sources) followed by the essive suffix *-pe/-pi*, the nominalized essive suffix *-pa*, or the negative suffix *-'ka*.

The multiplicity of forms makes it somewhat more difficult to reconstruct the actual morphology of the source construction, not least because only two of these segmented forms, *-pe* and *-pa*, have readily recognizable cognates in other Cariban languages. That said, both *-jtë/-jté* and *-'ka* might be recognizable as cognate to attested forms given irregular syllable reduction in the Panare reflexes (common in contexts of grammaticalization and also not particularly unusual in Panare).

First, the desiderative root suffix, *-jtë*, could come from a reduction of either the Proto-Cariban desiderative postposition **te* (> Pan. *se*) plus the (lexically conditioned) allomorph of the nominalizer **-to* (> Panare *-të*), or it could come from a reduction of the verbal root *pi* 'want', attested in other Venezuelan languages, plus the agentive adverbializer **të*, itself followed by the nominalizer **-no*, a combination which would regularly become *-to* in Panare.[19] In favor of the first hypothesis are the cognates *sa-to* (Apalaí; Koehn & Koehn 1986: 96), *xa-tî* (Waiwai; Hawkins 1998: 96), and *sa-to* (Tiriyó; Meira 1999: 189).[20]

The reduction of the penultimate syllable would be irregular because the nominalizer should have conditioned lowering of the postposition vowel to *a*, which generally is not a reducing vowel: *-sa-të > -stë > -jtë*. The correspondence in the nominalizing suffix of Tiriyó *o* to Panare *ë* is also irregular, but its plausibility is enhanced by the irregular correspondence with Waiwai *î* as well. Against the first hypothesis is the fact that nominalized forms of the desiderative postposition are not attested in languages more closely related to Panare, nor is it common for the vowel *a* to delete in reducing syllables. In favor of the second hypothesis, the verb *pi* 'want' is attested nearby to Panare, the adverbializing suffix is attested in Panare, and the loss of the nominalizing suffix **-no* (after changing word class and the preceding vowel) is also well attested in Panare. Against the second hypothesis is the absence of a motivation for the composite form **-jto > -jtë/-jte*. At this point, we are not convinced that either hypothesis is compelling, but either is marginally plausible and both have the

---

19   Our thanks to an anonymous reviewer for this suggestion.

20   In Tiriyó, the desiderative postposition also accepts the non-cognate nominalizer *-n(o)*, as in *-sa-n* (Meira 1999: 189), and only a reflex of **-no* is found nominalizing *se* (> *sa*) in Kari'na *sa-n* (Hoff 1968: 314) and Wayana *sa-no* (Tavares 2005: 171).

benefit of providing the argument structure that becomes the nominative-absolutive pattern in the resultant desiderative clause.

The final suffix -'ka 'negative' is also an inflectional suffix in Panare (Payne & Payne 2013: 229–230), but the only plausible cognate forms from other Cariban languages are in the phrasal negative particle *taike 'not₁' (Proto-Taranoan; Meira 2000: 104–105) and *taki* 'NEG' (Katxuyana, Gildea's field notes), which presumably would have been itself nominalized to form *taka* 'one that is not'; after affixation, the first syllable would then have idiosyncratically reduced to produce -'ka. The attested forms and their proposed etymologies are all listed in (23).

(23) *se-ato   >   *-jtë*   'one that wants'
      *se-ato-pe   >   *-jtë-pe/-jté-pe* 'like one that wants'
      *se-ato-pe-ano   >   *-jtë-pa/-jté-pa* 'one that is like one that wants'
      *se-ato-take-ano   >   *-jtë-'ka/-jté-'ka* 'one that does not want'

Although there is much more descriptive and comparative work to be done before we can be certain about the details of the ultimate sources for these specific morphemes, we know that the argument structure in the source was identical to that of the future tense: the S of the (copular) nonverbal predicate was coreferential with the A/S of the embedded (nonfinite) clause, and as such, the oblique A of the subordinate clause did not occur; the embedded clause was possessed by its absolutive argument. To get the flavor of this construction, consider the English calques in (24b–c):

(24) a. I am [like [one who wants [a drink]]]$_{PP}$
      b. I$_i$ am [like [one who wants [my$_i$ walking]]]$_{PP}$ >
          'I want to walk'
      c. I$_i$ am [like [one who wants [$\varnothing_i$ the book's writing]]]$_{PP}$ >
          'I want to write the book'

In (24b) the subject of the matrix copula is coreferential with the notional S of the subordinate clause, but the subordinate S also must occur, like a resumptive pronoun, as the obligatory possessor of the nominalized intransitive verb. In (24c), the subject of the matrix copula is coreferential with the notional A of the subordinate clause, but as that A is an optional oblique, it is readily elided, leaving the notional O as the obligatory possessor of the nominalized transitive verb. After reanalysis, the argument structure has both the nominative subject (< S of copula) and the absolutive verbal prefix (< obligatory possessive prefix); the matrix copula becomes an auxiliary (agreeing with A/S) and the nominalized verb is the new main verb.

(25) The proposed etymological analysis of the Panare Desiderative

      [[PSSR-V-NZR   ]   POSTP]$_{PP}$ S-COP   S

  a. *witënëjté*   *pi*   *maj*   *yuto*

      ∅-w-të-në-jté   pe   m-aj   yuto

      1S-S$_A$-go-1+2S-DESID.NZR   ATTR   2/3-COP.PAST 1+2

      lit. 'We (dual inclusive) were (like) wanters of our going.'

      [PSSR-V-NZR]   POSTP]$_{PP}$ S:COP   S

  b. *atyajté*   *pe*   *këj*   *kën*

      a-tya-jté   pe   këj   kën

      2O-hear-DESID.NZR   ATTR   3.COP 3.ANIM

      lit. 'He is (like) a wanter of the hearing of you'

So for both the Panare future tense and desiderative mood, the argument structure in the biclausal source construction becomes the nominative-absolutive alignment after reanalysis.

5.3   *The Mixed Nominalization Sources*

This third kind of source of the A/S pivot, mixed nominalization, is attested in both Panare (for the nonspecific aspect inflection) and Katxuyana (for the imperfective inflection). In each case, two distinct source constructions have merged, with one source construction providing the morphosyntax of the intransitive clauses and a different source construction providing the morphosyntax of transitive clauses. The reconstructions are secure (cf. Gildea 1998:184–186, 213–216), so we will only reprise briefly here. In both cases, intransitive verbs occur in the simple action nominalization, whereas for transitive verbs, they differ: in Panare, the transitive V uses the agent nominalization, whereas in Katxuyana, the transitive V uses the action nominalization in a postpositional phrase. First we address the Panare case, beginning with the English calque given in (26).

(26) The source for the Panare Nonspecific Aspect: S/A pivot

    a. I$_i$ am [my$_i$ walking]   > I walk/am walking/will walk.

    b. I$_i$ am [∅$_i$ the book's writer]   > I write/am writing/will write the book

The action nominalizer (which can occur on both transitive and intransitive verbs) is *-n* (< *-rɨ*), and the agent nominalizer (which can occur only on transitive verbs) is *-ñe* (< *-ne[mɨ]*). When each occurs as the complement of a copula, the combination establishes an A/S pivot (the copular subject = S/A), with the result that the copular subject becomes the nominative of the

reanalyzed clause. Because both nominalizations are obligatorily possessed (the action nominalization by the subordinate S and the agent nominalization by the subordinate O), the structure of the absolutive is retained as well. The examples in (27a–b) show the presumed etymologies for modern nonspecific aspect clauses.

(27) a.   [PSSR-V-NZR]PRED S.COP S
      yutën   *këj*   *kën*
      y-w-të-n   këj   kën
      3S-INTR-go-NZR   3.COP   3.ANIM
      lit. 'he$_i$ is his$_i$ going' (> 'he goes / is going / will go')

      b.   [PSSR   V-NZR   ]PRED S.COP S
      osowantënë   *yaarɨkañe*   *kë'*   *i'yan*
      as-awantë-në   y-aarɨka-ñe   kë'   piyan
      DETR-make.ill-INF RP-remove-A.NZR 3.COP   shaman
      lit. 'The shaman is illness' remover.'
      (> 'The shaman removes / is removing/ will remove the illness')

The Katxuyana imperfective pattern is illustrated with the English calques in (28):

(28) The sources for the Katxuyana Imperfective Aspect
      a. I$_i$ am   [my$_i$ walking] > I am walking/habitually walk.
      b. I$_i$ am [on [∅$_i$ the book's writing]] > I am writing/habitually write the book.

The intransitive imperfective allomorph is simply the action nominalizer *-rɨ*, with no unique phonological changes to distinguish it. The transitive imperfective allomorph is the same action nominalization, but made of the object of a locative postposition: *-rɨ hoko* (< *-rɨ pôkô*) which is then subject to idiosyncratic phonological reduction:[21] initially to *-rhoko*, then to *-:roko* and even to the extreme of *-rko* on the high frequency lexical item *ka* 'say'. Thus, Katxuyana collapses a predicate nominal construction and an adverbial predicate

---

21   Gildea (1998: 198ff) reconstructs the postposition *poko 'on the surface of', and shows that a modern reflex is attested in many Cariban languages with the meaning 'occupied with'; Meira & Franchetto (2005) offer good evidence of the need to distinguish between *o and *ô (mid back/central unrounded) in Proto-Cariban, and Gildea, Hoff & Meira (2010) show that * pôkô is the more plausible reconstruction.

construction into a single suppletive inflection in a nominative-absolutive imperfective construction. As already seen, the action nominalization in the postpositional phrase provides a ready S/A pivot, and as such, this combination of constructions provides precisely the necessary conditions for the creation of the nominative-absolutive alignment. The examples in (29) show the reconstructed sources for the two Katxuyana imperfective clauses.

(29) a. [PSSR-V-NZR-COL]$_{PRED}$ S-COP
          "owohɨrkum   *tahaye etxko,"*   *kamo*   *tɨ*
          o-wɨ-ohɨ-rɨ-kumu   tahaɨe etxi-ko   ka-mo   tɨ
          2-S$_A$-come-IMPRF.I-COL always COP-IMPER   say-DIST.PAST HSY
          lit. ' "Be always [your coming]," he said (HEARSAY).'

      b. [[PSSR-V-NZR]   POSTP]$_{PRED}$ S-COP
          onoo   *roko*   *ahkɨmɨ*
          ∅-ono-rɨ   hoko   ∅-ah-kɨmɨ
          3O-eat.meat-NZR   about   3A-COP-DISTANT.PAST
          lit. 'He$_i$ was [about (occupied with) [∅$_i$ its (meat food's) eating].'

Here, for the first time we see a typologically unusual pathway for the creation of an innovative construction, in which speakers of two different languages have selected different subordinate clause structures and merged them into a single suppletive inflection/construction. Such a merging of two distinct source constructions has also given rise to a progressive attested in six modern Cariban languages, including Panare (reconstructed in Gildea 1998: 197–217). A relatively recent issue of *Studies in Language* was dedicated to the topic of multiple source constructions in syntactic change, and in the lead article, Van de Velde et al. (2013) suggest that the phenomenon is much more common than most realize, simply because we as a field have not focused much attention on it.

   To summarize, all three types of biclausal sources presented in Section 5 are similar in that the morphological and syntactic properties of their source constructions lead to the creation of a nominative category (the S/A of the matrix clause in the source construction) and the absence of an explicit marked A in the subordinate clause structure, giving rise in a fairly mechanical way to the nominative-absolutive alignment we encounter in the various synchronic constructions. In the next section, we turn to the final type of source, which originally retained the marked A, and in which the attested synchronic nominative-absolutive alignment is created by later replacement of the marked (ergative) A with an unmarked A.

6   Creating Nominative-Absolutive through Loss of Ergative Case-Marking

There are two different construction types, both in the Jê family, in which the nominative-absolutive pattern is created through loss of the ergative case-marker, or perhaps more correctly, through replacement of the marked A from the source construction with an unmarked A. For each of these examples, there are also contexts in which the ergative-marked A is not lost, thereby creating a split, with one construction maintaining the ergative-absolutive alignment and another shifting to nominative-absolutive alignment.

In these constructions we find the combination of properties we expect to accompany innovative Tense-Aspect-Modality (TAM) in northern Jê: the main verb is in the nonfinite form (often followed by an auxiliary) and the absolutive argument, which has no case-marking, immediately precedes the main verb or is expressed as a bound pronoun/prefix on the verb. The innovative splits are only in case-marking of the A. In the Canela evaluative (30) and negative constructions (31), the past tense is identifiable by the fact that the A bears ergative case (30d, 31a) and the S is internal to the VP (30c). The nonpast evaluative and negative construction has a different treatment of A/S: in the intransitive nonpast, the S NP is external to the VP and the nonfinite verb bears an S prefix (30e); in the nonpast transitive, the A lacks the ergative case-marker (30f, 31b).[22]

(30) Canela evaluative modes: 'well' and 'a little'

	S	V		A	o-V	
a.	*kahãj*	*krɛ*	b.	*humrɛ*	*iʔ-kura*	CANELA
	woman	sing		man	3-kill	
	'the woman sings'			'the man kills it'		

	S	V.NF	AUX		[A]	o-V.NF	AUX	
c.	[*kahãj*	*ŋkrɛr*]ᵥₚ	*mpɛj*	d.	*humrɛ* *tɛ*	*iʔ-kuran*	*krirɛ*	CANELA
	woman	sing.NF	good		man ERG	3-kill.NF	little	
	'the woman sang well'				'the man killed it a little'			

---

22   This argument structure difference actually led Popjes & Popjes (1986: 180) to analyze the ergative case-marker *tɛ* (which only occurs with the "long form" of the verb) as the morphological marker of past tense; this unfortunate decision led them to consider intransitive clauses as lacking a means to indicate past tense.

	S	s-V	AUX	A	o-V.NF	AUX
e.	*kahãj*	*iʔ-ŋkrɛr*	*mpɛj*	f. *humrɛ*	*iʔ-kuran*	*krirɛ*	CANELA
	woman	3-sing.NF	good	man	3-kill.NF	little
	'the woman sings well'			'the man kills it a little'		

(31) Canela negative past (ergative-absolutive) and nonpast (nominative-absolutive)

	[A	ERG]	[O		V.NF]	AUX
a.	*aʔkrajrɛ*	*tɛ*	*nẽ*	*hĩ*	*krẽr*	*narɛ*	CANELA
	child	ERG	NEG	meat	eat.NF	NEG
	'The child didn't eat meat.' (Castro Alves 2004: 129)					

	A		o -V.NF	AUX
b.	*wa*	*ha*	*iʔ-pɨr*	*narɛ*	CANELA
	1	IRLS	3-grab.NF	NEG
	'I will not grab it'			

In the Kĩsêdjê future (32) and negative (33), all pronouns must occur in the ergative case (32b, 33b), whereas a full NP A is obligatorily unmarked (32a, 33a). For more detail and illustration of these synchronic patterns, cf. Castro Alves (2010), Gildea & Castro Alves (2010) and the references therein.

(32) The Kĩsêdjê future with full NP and pronominal A

	A		[O	V.NF]	AUX
a.	*ludu*	*ra*	*tɛp*	*kuru*	*mã*	KĨSÊDJÊ
	Ludo	SM	fish	eat.NF	FUT
	'Ludo will eat fish' (Santos 1999: 232)				

	A-ERG		[O	V.NF]	AUX
b.	*i-rɛ*		*hwĩsɨ*	*ren*	*mã*	KĨSÊDJÊ
	1SG-ERG		fruit	pick.NF	FUT
	'I will pick fruit' (Santos 1997: 56)				

(33) The Kĩsêdjê negative with full NP and pronominal A

	A		[O	V]	AUX
a.	*rɔtʃi*	*ra*	*mĩtʃi*	*pĩrĩ*	*kere*	KĨSÊDJÊ
	anaconda	SM	caiman	kill.NF	NEG
	'The anaconda did not kill the caiman' (Santos 1997: 165)				

A-ERG	[O	V]	AUX	
b. *i-rɛ*	*hwĩŋgrɔ*	*j-antoro*	*kere*	KĨSÊDJÊ
1-ERG	firewood	RP-hang.NF	NEG	

'I didn't hang the firewood. (Santos 1997: 56)

Given that Castro Alves (2010) has already presented the cognate ergative constructions for all four languages (pp. 463–466) and reconstructed the ergative constructions in Canela (pp. 466–473), our primary task in this section is to describe the situations in which the ergative case-marker is lost. We briefly reiterate Castro Alves' (2010) reconstruction of the source constructions that provide the grammatical morphology and source ergative-absolutive patterns (Section 6.1), then we argue that the topicalizing constructions in Kayapó, which create a (nominative, or unmarked) topic A alongside the prior ergative A, represent an intermediate stage that existed in these constructions for both Kĩsêdjê and Canela (Section 6.2); and finally (Section 6.3), we show how Kĩsêdjê and Canela have split each construction, selecting either the ergative or the nominative A according to person (Kĩsêdjê) or tense (Canela).

6.1 *The Nominalized Clause Is the S of an Intransitive Predicate*

Castro Alves (2010) reconstructs the negative and the evaluative modes as biclausal constructions in which an intransitive verb is the nucleus of the matrix clause and the nonfinite verb is its subject. To this, we add the reconstruction of the Kĩsêdjê future as a (nonverbal) predicate locative matrix clause with the nonfinite verb as the object of its postpositional (locative) phrase predicate. We begin with Castro Alves' reconstructions.

In the case of the intransitive matrix clause, the subordinate clause in its entirety serves as the S of the matrix verb, as illustrated with the English calques in (34–35). The cognate matrix verbs are *mpẽj* 'be good', *kʰɛat* 'be bad', *tɔʔhi* 'be a lot', *ŋkri=rɛ* 'be a little', *na=rɛ* 'not exist' (< **na=rɛ* 'finish/end up'), and *ket/kere* 'not exist'.[23]

---

23   The negative auxiliary *narɛ* is found only in Canela, but we can find the cognate forms *inorɛ* 'finish' in Mẽbêngôkre (Castro Alves 2010: 468–469), and *nõ ~ rõ* NEG in Panará (Dourado 2001: 117–118); *ket/kere* is the negative form found in the other three languages: Kĩsêdjê *ket* before another predicating element (e.g., Santos 1997: 94) or the clitic *=re* '?' (Santos 1997: 96) and *kere* sentence-finally (e.g. Santos 1997: 95), Apinajé *ket* to negate nonverbal predicates and ket=*nẽ* to negate verbal predicates (Oliveria 2005: 249), and Mẽbêngôkre *ket* for verbal negation, but *keti* also attested in unknown contexts (Reis Silva 2001: 63 note 2).

(34) a. *There is no [salt] ~ [(the) salt] finished up/does not exist.*

      b. [[*My running*] *finished up/does not exist*]   > I am not
                                                                              running.
      c. [[*The meat's cooking by me*] *finished up/does not exist*] > I am not
                                                                           cooking the meat.

(35) a. [[The *salt*] *is good/bad/much/little*].
      b. [[*My running*] *is good/bad/much/little*]   > I am running well/badly/
                                                              a lot/a little.
      c. [[*The meat's cooking by me*] *is good/bad/much/little*]
                                            > I am cooking the meat well/badly/a lot/a little.

This source reflects an S/Sentential pivot (Gildea's 1998, 2000 "pleonastic pivot"), in which the entire subordinate clause is (or is coreferential with) the sole participant in the main clause, i.e., the existence of the event expressed by the dependent clause is itself the sole semantic argument of the matrix clause. After reanalysis, the ergative-absolutive alignment pattern of the nonfinite clause surfaces intact into the new main clause, with an ergative-marked A NP or pronoun and the unmarked preverbal absolutive NP in alternation with an absolutive pronominal prefix on the verb. This is the source construction for the evaluative modes in Canela, plus negation in both Canela and Kĩsêdjê (and Mẽbêngôkre, as we will see in the next section).[24]

   In the second source construction, which gave rise to the Kĩsêdjê future tense, an allative postposition becomes the future auxiliary. This source has not been reconstructed before, and it may be controversial – the outcome is certainly remarkable. We begin presentation of our hypothesis with the cognate set for the dative postposition *mã*, which is identical to the future auxiliary in Kĩsêdjê (36). Next, we present the simple verbless existential / predicate locative clause in Kĩsêdjê (37a), a construction with precise analogues in Canela (37b), Mẽbêngôkre (37c), and Panará (37d). In each case, the subject occurs first, followed by the (locative) predicate PP.

(36) Kĩsêdjê: *mã* 'DATIVE, BENEFACTIVE, DIRECTIONAL'
                                                          (Santos 1997: 97)
      Canela: *mã* 'DATIVE, BENEFACTIVE'   (Castro Alves 2004: 8)

---

24   Castro Alves (2008:17; 2010:470–471) suggests that the recent past construction is derived historically by the loss of an earlier auxiliary. However, another hypothesis under exploration is that the recent past is a further development of a possessive perfect construction, which would have had no auxiliary in the source (cf. the cognate 'perfect' construction in Mẽbêngôkre, analyzed in great semantic detail – albeit in a very different framework – by Salanova 2007).

Apinajé: *mə̃* 'DATIVE, BENEFACTIVE, LOCATIVE'

(Oliveira 2005: 141–142)

Mẽbêngôkre: *mĩ* 'DATIVE, ALLATIVE, PROSPECTIVE'

(Reis Silva 2001: 43)

Panará   *mã* 'DATIVE, BENEFACTIVE, ALLATIVE'

(Dourado 2001: 55)

| | S | [[Possr Possd]ₙₚ | P]ₚₚ / pred |

(37) a. | *aket* | *ra* | *kosʌ* | *krɛ* | *kãm* | KĨSÊDJÊ |
| the.rest | SM | basket | interior | LOC | |

'The rest (is) in the basket' (Santos 1997: 140)

b. | *kwər* | *pur* | *kʰãm* | CANELA |
| manioc | garden | LOC | |

'There is manioc in the garden.' (Castro Alves 2004: 124)

c. | *mrɯ* | *(nẽ)*²⁵ | *bʌ* | *kam* | MẼBÊNGÔKRE |
| game | NFUT | woods | in | |

'The animal is in the woods.' (Salanova 2007: 110)

d. | *kuɨ* | *puu* | *yamã* | PANARÁ |
| manioc | garden | LOC | |

'There is manioc in the garden' (Dourado 2001: 99)

To derive the source construction from this matrix clause, we simply replace the possessed lexical noun in the predicate PP with a nonfinite verb and its complements, giving us clauses like the English calques in (38b–c). In (39), we see a clause from Mẽbêngôkre that is virtually identical to the Kĩsêdjê future construction, presented by Salanova as a synchronic predicate locative construction, i.e., as exactly the analysis that we reconstruct for the antecedent of the Kĩsêdjê future.

(38) a. It (is) [in [the basket].
        b. It (is) [to [my walking]]   > 'I (ABS) will/am about to walk'
        c. It (is) [to [the meat's cooking (by me)]]   > 'I (ERG) will/am about to cook the meat'

---

25   We interpret this *() to mean that the *nẽ*, which indicates that the preceding element is in focus position, is obligatory (ungrammatical if optional), cf. Salanova's prose before the example: "But to predicate of a locatum focalization seems to be required."

(39) *arɔm*   *kutɛ*   *tɛp*   *krẽn*   *mɑ̃*   MĚBÊNGÔKRE
     already   3ERG   fish   eat.NF   to
     'He's already about to eat fish.' (Salanova 2007: 56)
     (lit. 'Already (it is) to the eating of fish by him')

In this particular source construction, neither the existential subject nor a copula occurs explicitly, leaving only the ergative-absolutive argument structure of the erstwhile nominalization to express the core arguments.

Having reconstructed all of these split systems to constructions that represent the A argument as an oblique and the O/S arguments as possessors of the nonfinite verb form, we turn next to the elaboration of these constructions that introduced the option of an unmarked A.

6.2   *Subject Pronoun Doubles*

In the grammar of all the Northern Jê languages, we find descriptions of a focus (sometimes called "cleft") construction, in which a focused noun or (nominative) pronoun occurs in first position, followed by a normal clause of whatever type, often with a resumptive pronoun that indicates the grammatical role of the preceding topicalized NP. We also find descriptions of a "subject doubling" rule, in which a noncanonical (i.e., non-nominative) subject may be doubled via a nominative pronoun, which is often described as though it were identical to the focus pronoun, but which may be distinguished in several ways from a pronoun in the focus position. First, the subject copy pronouns do not appear to give additional emphasis or "focus" to the clause, but rather they simply mark the TOPIC with some frequency (in some cases, obligatorily) before certain tense-aspect particles, e.g. *ha* 'irrealis' in Canela (Castro Alves 2004: 95).

	TOP		A		
(40) a.	*ke*	*ha*	*hũmrɛ*	*rɔpti*	*j-ahe*	CANELA
	3	IRLS	man	jaguar	RP-hunt

'The man will hunt a jaguar.' (Castro Alves 2004: 95)

	TOP		A	
b.	*wa*	*ha*	*i-mɑ̃*	*h-ŭpa*	CANELA
	1	IRLS	1-DAT	3-fear'

'I will fear him.' (Castro Alves 2004: 100)

	TOP		S		
c.	*ke*	*ha*	*hũmrɛ*	*iʔ-ŋkrɛr*	*narɛ*	CANELA
	3	IRLS	man	3-sing.NF	NEG

'The man will not sing.' (Castro Alves 2004: 111)

    TOP   A

d. *wa*   *i-tɛ*   *ku-mã*   *rɔpti*   *j-akrɛ*   Canela

   1   1-ERG   3-DAT   jaguar   PR-show

   'I showed him the jaguar.' (Castro Alves 2004: 118)

Santos (1997: 58–60) also discusses the "construction with topicalized pronouns", and while he does not specify anything about the grammatical role of these nominative forms that immediately precede the core of the clause, his examples are consistent with the claim that the topicalized pronoun is always coreferential with the subject (cf. 41a–b).[26] A similar set of examples is available for Mẽbêngôkre (41c–d) – about these topic pronouns, Salanova (2007: 34–35) asserts:

> The unusual characteristic of nominative pronouns is that, in main clauses, they can duplicate a subject that is already expressed lower in the clause by an ergative, dative or absolutive pronominal form. These pleonastic nominative pronouns, unavailable in embedded clauses, seem not to indicate any emphasis ... We take the ability to be "duplicated" by a nominative pronoun in the position between tense and aspectual particles as the primary diagnostic for subjecthood in Mẽbêngôkre main clauses. (Salanova 2007: 34–35)

(41)   Nominative topic pronouns in Kĩsêdjê (a–b) and Mẽbêngôkre (c–d)

                                                Kĩsêdjê

   FOC$_i$ TOP$_i$ S$_i$-V.NF   AUX   FOC$_i$ TOP$_i$ A$_i$-DAT   O-V

a. *pa-n*   *wa*   *i-mbərə*   *kere*   b. *pa-n*   *wa*   *i-mã*   *a-kĩ*

   1SG-TOP   1SG   1SG-cry   NEG   1SG-TOP   1SG   1SG-DAT   2SG-like

   'I didn't cry' (Santos 1997: 69)   'I like you' (Santos 1997: 133)

   TOP$_i$ S$_i$-V.NF   TOP$_i$ A$_i$-ERG   O-V

c. *ba*   *i-tẽm*   d. *ba*   *ijɛ*   *ir*   Mẽbêngôkre

   1NOM   1-go.NF   1NOM   1ERG   3.put.down.NF

   'I go.' (Salanova 2007: 35)   'I put it down.' (Salanova 2007: 35)

While sometimes appearing identical (the pronominal forms are the same), a focused element differs in that (i) it is pragmatically marked, e.g. for contrast, (ii) it may be either a full NP (42a–b) or a pronoun (42c–d, f–g), (iii) it is not limited to the subject role (42b, e, g), and (iv) in at least Mẽbêngôkre, it occurs

---

26   Note that (41a–b) also contain initial focused pronouns (marked with -*n* 'TOP'), which we discuss in the next paragraph.

in a different position vis-à-vis the tense and aspect particles, e.g. preceding *nẽ* 'NONFUTURE' or *dʒa* 'IRREALIS/FUTURE' (Salanova 2007: 34; 110, note 24; 131, note 50); in Kĩsêdjê, it is consistently marked with the suffix/particle *-n/ na* 'FOCUS' (42e–f). While it is not obligatory, a focused element may be core-ferential with the nominative pronoun (42a, c–d, f), the ergative A pronoun (42d, f), or the absolutive S prefix (42f). Example (42f) is remarkable for having four different forms that refer to the same subject participant: from the left, the first element is the FOCUS form *pa-n* '1SG-TOP'; the second is the nomina-tive TOPIC form *wa* '1SG'; the third is the ergative form *i-rɛ* '1SG-ERG', which is allowed uniquely in future, progressive, and negation when a postpositional goal phrase occurs between this S pronoun and the verb (cf. Gildea & Castro Alves 2010: 188–189 for the argument that this is evidence for reanalysis of the Kĩsêdjê future and negative); and finally, the only one of these that is obliga-tory, the S prefix on the nonfinite main verb *i-tẽm* '1SG-go.NF'. Such sequences of coreferential pronouns are found in elicitation in several Northern Jê lan-guages, but in Castro Alves' Canela text database, no more than two such core-ferential pronouns are found in a single clause. It is possible that such extreme redundancy is an artifact of the elicitation situation.

(42) NPs in focus position in Canela (a,f), Mẽbêngôkre (b–d), Kĩsêdjê (e)

FOC*ᵢ*   A*ᵢ*   TAM   O   V
a. *intuw   ke   ha   rɔpkrɔr   pupu*   CANELA
   youth   3   IRLS   jaguar   see
   'The youth, he will see a jaguar.' (Castro Alves 2004: 95)

FOC*ᵢ*   TAM   O*ᵢ*-V
b. *mru   nẽ   ku-bĩ*   MẼBÊNGÔKRE
   animal   NFUT   3ACC-kill.SG.V
   'He killed *an animal* (focus).' (Salanova 2007: 34)

FOC*ᵢ*   TAM   TOP*ᵢ*   S*ᵢ*-V.NF
c. *ba   nẽ   ba   i-tẽm*   MẼBÊNGÔKRE
   1NOM   NFUT   1NOM   1-go.NF
   '*I* go.' (Salanova 2007: 35)

FOC*ᵢ*   TAM   TOP*ᵢ*   A*ᵢ*-ERG   O-V
d. *ba   nẽ   ba   ijɛ   ir*   MẼBÊNGÔKRE
   1NOM   NFUT   1NOM   1ERG   3.put.down.NF
   '*I* put it down.' (Salanova 2007: 35)

e.
[FOC]		TOP$_i$	S$_i$-V.NF	AUX	
a-kot	*na*	*wa*	*i-tẽm*	*mã*	KĨSÊDJÊ
2-COMIT	FOC	1SG	1SG-go.NF	FUT	

'I will go **with you**.' (Santos 1997: 98)

f.
FOC$_i$	TOP$_i$	S$_i$-ERG	[LOC]		[LOC]	
pa-n	*wa*	*i-re*	*akatʃi*	*ɲĩ*	*ŋgo kot*	
1SG-FOC	1SG	1SG-ERG	tomorrow	LOC	river ALL	

S$_i$-V.NF	AUX	
i-tẽm	*mã*	KĨSÊDJÊ
1SG-go.NF	FUT	

'Tomorrow I will go fishing.' (Santos 1997: 57)

g.
FOC$_i$		TOP$_j$	A$_j$	O$_i$-V	
ta	*mã*	*wa*	*i-tɛ*	*∅-kakʰwĩn*	CANELA
3	TOP	1	1-ERG	3-hit	

'It was him I hit (and no other).' (Castro Alves 2004: 127)[27]

We are now able to return to the innovative ergative constructions that we reconstructed in Section 6.1 in order to illustrate the relevance of subject doubling, whether via the TOPIC or the FOCUS slots at the beginning of the clause: it creates a context in which the subject is frequently expressed twice (or three times), once closer to the verb via the ergative argument inherited from the nominalized clause in the source, but also closer to the beginning of the utterance via forms that do not bear ergative case: either a focus NP/pronoun, the nominative subject doubling TOPIC pronoun, or both. In (43), we show examples from Mẽbêngôkre: (43a) is the simple case, the source of the negative with only the argument structure of the embedded nominalization (i.e., with the TOPIC and FOCUS slots unfilled); (43b) is the same type of clause, except with the TOPIC slot filled by a nominative pronoun. The reverse order of marking is also attested, with an unmarked initial A NP, presumably in FOCUS

---

27 Example (42g) was produced in answer to the question 'Who did you beat?', making it clear that the pronoun *ta* '3' is a focused element. This focused pronoun is followed by the particle *mã*, which Castro Alves (2004) glosses as 'TOPIC'. Here we do not change the original gloss, although we do recognize that additional research into Canela discourse will be necessary to resolve the conditions in which *mã* marks focused elements (for example, one hypothesis might be that a focused element followed by *mã* signals a new topic in the subsequent discourse).

position (given that TOP IC position is only filled by pronouns), followed by a coreferential ergative pronoun (43c).[28]

(43) a.

TOP	[[OBL]	POSSR-V.NF]$_S$	V$_{INTR}$	
_____	*kaitirɛ*	*tɛ*	*∅-mari*	*ket*	MĚBÊNGÔKRE
	Kaitire	ERG	3-know.NF	NEG	

'Kaitire doesn't know (it).' (Reis Silva 2001: 48)
(lit. 'The knowing of it by Kaitire does not exist')

b.

TOP$_i$	[[OBL$_i$]	POSSR-V.NF]$_S$	V$_{INTR}$	
ga	*a-jɛ*	*∅-mari*	*ket*	MĚBÊNGÔKRE
2SG	2SG-ERG	3-know.NF	NEG	

'You don't know it' (Reis Silva 2001: 47)
(lit. 'You, the knowing of it by you does not exist')

c.

FOC$_i$	TOP	[[OBL$_i$]	POSSR-V.NF]$_S$	V$_{INTR}$	
kaitirɛ	∅	***kutɛ***	*∅-mari*	*ket*	MĚBÊNGÔKRE
Kaitire		3ERG	3-know.NF	NEG	

'Kaitire doesn't know (it).' (Reis Silva 2001: 48)
(lit. 'Kaitire, the knowing of it by him does not exist')

Similar examples are readily available for Mẽbêngôkre cognates to the Canela evaluative moods (44a) and to the Kĩsêdjê future (44b). In these two examples, we see the reconstructed ergative-absolutive alignment inside the nominalized clause that is the S of the verb *kumɛj* 'be many' (44a) and inside the predicate PP (44b); alongside the ergative A, in both cases we see an unmarked A double, the nominative TOP IC pronoun in (44a) and the unmarked FOCUS NP in (44b).

(44) a.

TOP$_i$	[OBL$_i$	[PSSR-V.NF]]$_S$	V	
ba	*ijɛ*	*∅-krẽn*	*kumɛj*	MĚBÊNGÔKRE
1SG	1ERG	3-eat.NF	be many	

'I eat (something) a lot.' (Castro Alves field notes)
(lit. I, the eating of it by me is a lot.')

---

28 In response to a query by an anonymous reviewer, there are no published intonational data for Mẽbêngôkre – our analysis does predict a strong likelihood of distinctive intonation.

FOC$_i$ [[OBL$_i$]   POSSR   V.NF]$_{NP}$ P]$_{PP}$

b. *kubẽ*   *kutɛ*   *tɛp*   *krẽn*   *mã*   MẼBÊNGÔKRE

   non.Indian   3ERG   fish   eat.NF   PROSP

'The non-Indian is about to eat fish' (Reis Silva 2001: 62)

(lit. 'The non-Indian$_i$, it is [to [the fish's eating by him$_i$]]')

We have now identified a commonplace construction in Mẽbêngôkre, Canela and Kĩsêdjê, that adds an unmarked nominative pronoun (in TOPIC position) or an unmarked A noun or pronoun (in FOCUS position) to any clause that has non-nominative A or S. Crucially, these forms are attested in the source constructions that become the innovative evaluative moods and the negative construction in Canela and the negative and future in Kĩsêdjê. The stage is now set for the final step, whereby the selective loss of the ergative A in some contexts creates the attested split alignment patterns in Canela and Kĩsêdjê.

6.3   *Selective Loss of the Oblique A in Canela and Kĩsêdjê*

In the grammars of the northern Jê languages, it is commonplace to see examples with both a pronoun and an NP – or two different pronouns – jointly referring to the A/S of a clause (Section 6.2), and it is not uncommon to see examples with three different pronominal forms indexing A (cf. 42c–d, f). Clearly this is a situation ripe for reinterpretation and simplification. While the unmarked topic (pro)noun can always co-occur with simple ergative-absolutive clauses in both Canela and Kĩsêdjê, whenever the topic pronoun occurs but the ergative pronoun does not, the result is a nominative-absolutive clause. In this section, we argue that the synchronic splits that have been described in Canela and Kĩsêdjê come from conservative ergative-absolutive constructions – in the past tenses in Canela and with pronominal A arguments in Kĩsêdjê – that are now in opposition with innovative nominative-absolutive constructions, which were created by suppression of the ergative A and retention of the doubled A TOPIC pronoun or FOCUS NP.

Considering first the tense-based split in Canela, the nonpast constructions are derived from the topic doubling construction via two changes. First, the topic pronoun becomes obligatory in the slot preceding second-position TAM particles. In the intransitive (45a), the subject doubling TOPIC pronoun becomes the obligatory nominative S pronoun, alongside the obligatory S prefix inherited from the absolutive possessor prefix on the nonfinite verb. In the transitive (45b), the subject doubling TOPIC pronoun becomes the nominative A, which precedes the absolutive O (whether NP or prefix) inherited from the possessor of the nonfinite verb. Second, the marked ergative A form is lost

(its former position is represented via the ∅ in 45b), leaving behind only the nominative A.

(45) Canela nonpast: Topic > Nominative A becomes obligatory, ergative A is lost

      S   [s-V.NF]   V_{INTR}

a. *wa*   *ha*   *i-wrɨk*   *narɛ*   CANELA

    1   IRLS 1-descend.NF NEG

    'I will not descend'

      A   [o-V.NF]   AUX

b. *wa*   *ha*   ∅   *iʔ-pɨr*   *narɛ*   CANELA

    1   IRLS   3-grab.NF NEG

    'I will not grab it'

Alongside these changes that create the nominative-absolutive pattern in the nonpast constructions, the past tense clauses remain conservative. In the past tense intransitive clause (46a), the TOPIC/FOCUS pronoun optionally co-occurs alongside the absolutive S (46a), which now leads to ambiguity in interpreting the tense of the clause. In the past tense transitive clause, the inherited pattern is also still allowed, in which a TOPIC/FOCUS pronoun optionally co-occurs alongside the ergative A (46b). However, there is also an innovative pattern, in which the unmarked position for the ergative A NP is at the beginning of the utterance, in the position etymologically reserved for the nominative pronoun (46c).

(46) Canela past: Oblique > Ergative A (doubling w/ Topic pronoun is allowed)

      (TOP)   s-V   AUX

a. *ka*   *a-j-ʒʔkukʰrẽn*   *narɛ*   CANELA

    2   2-RP-run.competing.NF NEG

    'You didn't / don't run.'

      (TOP) erg   [O   V.NF] AUX

b. *(ka)* *a-tɛ* *hĩ*   *kʰrẽr* *narɛ*   CANELA

    2   2-ERG meat eat.NF NEG

    'You didn't eat meat.' (Castro Alves 2008: 17)

      TOP/FOC   ERG   [O   V.NF] AUX

c. *aʔkrajrɛ* *ku-tɛ* *nẽ*   *hĩ*   *kʰrẽr* *narɛ*   CANELA

    child   3-ERG NEG meat eat   NEG

    'The child, she/he didn't eat meat.'

The simplification in Kĩsêdjê is not along the lines of tense-aspect, but rather along the lines of person: topic doubling (and even tripling) is attested with negative and future intransitive clauses (47a–c), but it is not attested in transitive clauses (48–49). In transitive clauses in the negative (48a) and the future (49b), the innovative nominative-absolute pattern is created by (i) obligatorily placing a full NP A referent in focus position, and (ii) dropping the third person ergative pronoun that once co-occurred (cf. the examples in Section 6.2) with that focused element. The particle *ra* 'SUBJECT MARKER' often follows both A and S NPs in Santos' (1997) examples, but it is not obligatory (cf. p. 82, where in otherwise identical sentences the A is unmarked, marked with *ra*, and marked with *-n* 'topic'). Moreover, it is not entirely restricted to subjects (cf. the discussion on pp. 129–130, which includes examples of O marked with *ra*). As such, despite the presence of *ra* in most of these examples, we do not posit the innovation of a new nominative case-marker as a part of the creation of the nominative-absolute pattern.

(47) The Kĩsêdjê topic doubling pronouns (Santos 1997: 119–120)

   a. *i-ŋgɛre*   *mã*   b. *wa*   *i-ŋgɛre*   *mã*
      1SG-dance.NF FUT   1SG   1SG-dance.NF FUT
      'I will dance'   'I (who) will dance'

   c. *pa-n*   *wa*   *i-ŋgɛre*   *mã*
      1SG-TOP 1SG   1SG-dance.NF   FUT
      'I (who) will dance'

(48) The innovative nominative-absolute: when A is an NP, use only TOPIC/ FOCUS position

      FOC   *A-ERG   [O   V.NF]   AUX
   a. *rɔtʃi*   *ra*   ∅   *mĩtʃi*   *pĩrĩ*   *kere*   KĨSÊDJÊ
      anaconda   SM   caiman   kill.NF   NEG
      'The anaconda did not kill the caiman' (Santos 1997: 165)
      (originally: 'The anaconda, (it) did not kill the caiman')

      FOC   *A-ERG   O   V.NF]   AUX
   b. *ludu*   *ra*   ∅   *tɛp*   *kuru*   *mã*   KĨSÊDJÊ
      Ludo   SM   fish   eat.NF   FUT
      'Ludo will eat fish' (Santos 1999: 232) (originally: 'Ludo, (he) will eat fish')

Alongside the construction that lost the ergative A NP (and even the ergative resumptive pronoun), Kĩsêdjê has retained the ergative marking only when

the A is a pronoun (49a–b). From the related languages, we know that coreferential topic and/or focus pronouns would have been able to co-occur with the ergative pronoun and we see no reason that this should not still be possible in modern Kĩsêdjê; with that said, we encountered no examples of topic or focus pronouns doubling the ergative A pronouns in Santos' (1997, 1999) examples. If topic or focus pronouns cannot co-occur with a coreferential ergative A pronoun in modern Kĩsêdjê, this would be one more example of a quirky pattern that is limited to the innovative negative and future constructions.

(49) The conservative Ergative-Absolutive: when A is a pronoun, retain the ergative

	TOP/FOC	A-ERG	O-V	AUX	
a.	∅	*ko-rɛ*	*i-kaken*	*kere*	KĨSÊDJÊ
		3-ERG	1SG-scratch.NF	NEG	

'He didn't scratch me' (Santos 1997: 132)

	TOP/FOC	A-ERG	O	V.NF	AUX	
b.	∅	*i-rɛ*	*hwĩsɨ*	*ren*	*mã*	KĨSÊDJÊ
		1SG-ERG	fruit	pick.NF	FUT	

'I will pick fruit' (Santos 1997: 56)

With this, we have now completed all of the actual reconstructions. We turn now to our conclusion, in which we summarize the reconstructions and show how they address the questions posed in the introduction.

7   Discussion: Reconstruction, Diachrony and Universals

We begin our discussion by returning to the outstanding questions from the introduction, after which we summarize our reconstructions and show how they address these questions.

First, it is typologically rare for argument marking to be nominative-accusative while verbal indexation is ergative-absolutive – how did this rare pattern develop in these languages? Second, considering the typological tradition of identifying a construction (and even an entire language) as ergative given just one robust ergative alignment pattern, the absolutive prefixation is sufficient to identify these constructions as belonging to the "ergative" type. However, each construction participates in an alignment split that is conditioned by tense, aspect, mood, and polarity values that are expected to condition the non-ergative side of such splits. Our reconstructions actually explain both of these unexpected outcomes as resulting from normal

grammatical changes that departed from a typologically unusual source construction.

We begin by observing that most of these reconstructions are quite robust, and require no unusual or surprising mechanisms of syntactic change. All of the reconstructions in Section 5 are simply additional examples of well-trodden pathways by which speakers recruit innovative tense-aspect-mood distinctions from existing biclausal constructions. So far as we know, innovations like these are attested in every language that has been recorded for longer than a few hundred years, creating a sufficiently robust body of examples to allow typologies of the sources (e.g., Bybee et al. 1994; Heine & Kuteva 2002). The reconstructions of the innovative ergative-absolutive patterns in Section 6 are similarly robust, reflecting the same kinds of reanalyses seen elsewhere in the literature.

What is unusual about these language families is that the biclausal source constructions that feed into these innovative tense-aspect-mood-negation constructions have very limited resources for expressing nonfinite clauses – in these cases, all of the constructions in question are based on nominalizations that are inalienably possessed by their notional absolutive argument. In her detailed survey of alignment patterns in nominalizations, Koptjevskaja-Tamm (1993: 223–228) finds that it is unremarkable (25 of the 70 languages in her sample) to have a nominalization that is possessed by its absolutive, with the A expressed distinctly, usually as an oblique. However, 22 of these 25 languages have multiple other subordination strategies (they are "complement balancing", in her terms); the three that have more limited subordination strategies include one Cariban language, Hixkaryana, and two NE Caucasian languages, Abkhaz and Georgian. While Hixkaryana is typical of the Cariban family in basing almost all subordination on such nominalizations, both Abkhaz and Georgian have a range of non-ergative subordination strategies available for both adverbial and relative clauses (Hewitt 1987). In sum, all the other languages attested as having ergatively organized nominalizations also have other (non-ergative) subordinating structures available for the creation of innovative tense-aspect-mood constructions; in contrast, such nominalizations are the dominant strategy in the Cariban and Jê language families, which have almost no alternative structures available.[29]

Because the source constructions all contain this possession > absolutive alignment pattern, every innovative tense, aspect, mood, or negation

---

29 Clearly, it would be desirable to conduct additional typological research to further study this claim: how many of the languages that have ergatively-organized nominalizations use them frequently in the sort of constructions that develop into innovative TAM? If our explanation for the relative rareness of the nominative-absolutive alignment pattern is correct, then this number should be quite low.

construction whose source contains one of these possessed nominalizations will, perforce, end up as a construction with absolutive verbal indexation. So in fact, there is no need to invoke any unusual functional factor to explain the creation (and hence the synchronic existence) of the absolutive indexation in these innovative tense-aspects. The absolutive indexation pattern was not in competition with other patterns, and so was not selected for some particular reason that we might capture in a typology: it is simply the basic building block of dependent clauses.

Although we no longer need to provide a functional explanation for the presence of absolutive indexation, we do need to explain the existence of the nominative patterns, as the same nonfinite source construction contains an oblique notional A, which, in both families (and in some cases, in the same languages), does surface as an ergative A in other innovative main clause constructions. Once again, the reconstructions are robust and follow well-attested patterns:

– Transitive phasal verbs predicate the initiation or termination of activity by an agent (Section 5.1), thereby creating the A/S pivot in the innovative completive construction
– Source constructions that locate an agent in the midst of an activity (Section 5.2.1) create an A/S pivot that becomes a nominative-accusative progressive construction (Heine 1994; Bybee et al. 1994, chapter 5; Gildea 1998: 36–37), which then evolves semantically to become a more general imperfective
– Source constructions that predicate the desire of or an obligation imposed upon an agent similarly create an A/S pivot with a subordinate construction (Section 5.2.2), thereby leading to a nominative pattern associated with agent-oriented modalities (which then can easily evolve into future tense)

Each of these source constructions has a matrix clause that selects as its subject the notional A/S of the nonfinite clause, which leads to (i) nominative grammatical patterns (case-marking, order and auxiliary agreement) in the resultant construction and (ii) the suppression of the optional oblique A in the nonfinite clause, thereby removing the source for an ergative A. Not coincidentally, the semantics of these source constructions are similarly agent-oriented, and thus lead to precisely the (agent-oriented) tenses, aspects, and moods that we expect to see associated with the non-ergative side of TAM-based splits. In other words, the innovation of these nominative patterns is motivated, and in the process, has conveniently removed a prior ergative case-marker from the source construction. If it were not for the retention of the conservative absolutive indexation pattern, these agent-oriented tense-aspect-moods would now fall where they belong typologically, on the non-ergative side of a tense-aspect-mood based split ergative system.

Hence, our reconstructions from Section 4 explain the presence of absolutive indexation and our reconstructions from Section 5 explain how innovative nominative case-marking (and other patterns) were added to that older absolutive pattern, thereby accounting for the counter-universal combination of alignment patterns. It is similarly clear how the innovative ergative-absolutive patterns in Section 6 were created, but we have no similarly robust explanation for the two different patterns of selective loss of ergative marking in Canela and Kĩsêdjê. On the one hand, it is well attested that case-markers can simply erode or stop being used, and many linguists have identified languages that appear to have lost an ergative case-marker within a given construction (e.g., Givón 1980, Estival & Myhill 1988, Dixon 1994, Harris & Campbell 1995, etc.). On the other hand, we are aware of only one other carefully documented case where the loss of the ergative case-marker is not complete, but is rather selective, leading to a specific pattern of split ergativity.

This change, showcased in a recent study by Guillaume (2015), documents the creation of a counter-universal case-marking split in Tacana (Takanan, Bolivia), where the ergative is obligatorily retained only for 1SG and 2SG A pronouns, is optionally retained for all third persons, and is completely lost for 1PL and 2PL pronouns. While there are multiple surveys in the literature about how ergative patterns are created diachronically (e.g., Anderson 1977; Dixon 1994; Harris & Campbell 1995; Gildea 2004), there is not a similarly rich research tradition about the processes by which alignment patterns in general – and ergative case-marking in particular – are lost. Combined with Guillaume's (2015) findings, our findings in this article suggest that attrition might create interesting transitional patterns of splits, either in the universal direction, as in Canela, or in the counter-universal direction, as in Kĩsêdjê and Tacana. This might be a fruitful direction to explore in future studies in diachronic alignment typology.

At this point, we can only speculate about the motives of those speakers who made the changes inherited by modern Canela and Kĩsêdjê speech communities. In Canela, it appears that they chose to always use the topic/focus position for the nominative argument uniquely in nonpast situations, so A/S topic/focus constructions were overused in nonpast situations and eventually evolved into indexical markers of nonpast tense. In contrast, agent focus/topic must have been rare in past situations, leading to a situation in which the absence of an A/S noun or pronoun in topic/focus position (and, for transitive clauses, retention of the ergative case-marker) came to index past tense situations. Similarly, the existing case-marking patterns in Kĩsêdjê associate ergative marking uniquely with pronouns, which means with continuing topics, whereas a free A NP must have occurred frequently in focus position, presumably originally with a resumptive (ergative) pronoun. Over time, speakers

appear to have stopped using an ergative-marked A NP inside the clause, expressing the A NP only in focus position; when they also stopped using the resumptive ergative pronoun, the result would be the attested split alignment pattern. It would be most interesting to enquire if similar statistical asymmetries in focus constructions are attested in texts from closely related languages (such as Mẽbêngôkre and Apinajé) that have not made them obligatory grammar as they are in Kĩsêdjê. Alternatively, similar information structure asymmetries might be observable in completely unrelated languages for which we have larger, more readily searchable corpora.

However, these remain questions for future research – here we can reconstruct the historical changes that must have taken place, but unlike the reconstructions from Section 5, these reconstructions do not help us to explain the resultant patterns.

Abbreviations

1	First person	DETR	Detransitive
2	Second person	DIR	Directional
3	Third person	DIST	Distal Deixis
1+2	First person dual inclusive	DUR	Durative
ABIL	Abilitative	EMPH	Emphasis
ACC	Accusative	ERG	Ergative
ADV	Adverbial	ESS	Essive
AGT	Agent	FACT	Factitive
ALL	Allative	FOC	Focus
ANIM	Animate	FUT	Future
ASP	Aspect	GEN	Genitive
ATTR	Attributive	HSY	Hearsay
AUX	Auxiliary	I	Intransitive
CIRC	Circumstantial	IMPER	Imperative
CNJ	Conjunction	IMPRF	Imperfective
COL	Collective	INF	Infinitive
COMIT	Commitative	INSTR	Instrumental
COMPL	Completive	INTR	Intransitive
CONCL	Conclusive	INVIS	Invisible
COP	Copula	IRLS	Irrealis
DAT	Dative	LOC	Locative
DEF	Definite	NEG	Negative
DESID	Desiderative	NF	Nonfinite

NFUT	Nonfuture		RLS	Realis
NOM	Nominative		RP	Relational Prefix
NZR	Nominalizer		S_A	S_A verb class prefix
OBL	Oblique		SBRD	Subordinate
PAST	Past		SG	Singular
PL	Plural		SM	Subject Marker
POSSD	Possessed		SS	Same Subject
POSTP	Postposition		T	Transitive
PR	Relational Prefix		TAM	Tense-Aspect-Modality
PROSP	Prospective		TEMP	Temporary
PV	Preverbal Particle		TOP	Topic
REL	Relativizer		TRN	Transitive

References

Aikhenvald, Alexandra, R. M. W. Dixon, & Masayuki Onishi, eds. 2001. *Non-Canonical Marking of Subjects and Objects*. Amsterdam: John Benjamins.

Anderson, Stephen R. 1977. On Mechanisms by which Languages Become Ergative. *Mechanisms of Syntactic Change*, ed. by Charles N. Li, 317–363. Austin: University of Texas Press.

Barðdal, Jóhanna & Spike Gildea. 2015. Diachronic Construction Grammar: Epistemological Context, Basic Assumptions and Historical Implications. *Diachronic Construction Grammar*, ed. by Jóhanna Barðdal, Elena Smirnova, Lotte Sommerer & Spike Gildea, 1–50. Amsterdam: John Benjamins.

Blevins, Juliette 2004. *Evolutionary Phonology: The Emergence of Sound Patterns.* Cambridge: Cambridge University Press.

Bybee, Joan, Revere Perkins & William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. Chicago: University of Chicago Press.

Camargo, Nayara S. 2015. Tapayuna (Jê): aspectos morfossintáticos, históricos e sociolinguísticos. Ph.D. dissertation, Universidade Estadual de Campinas.

Castro Alves, Flávia de. 2004. O Timbira falado pelos Canela Apãniekrá: uma contribuição aos estudos da morfossintaxe de uma língua Jê. Ph.D. dissertation, Universidade Estadual de Campinas.

Castro Alves, Flávia de. 2008. O papel das nominalizações na evolução do alinhamento ergativo nas línguas Jê: Dimensões funcionais e estruturais. *Amérindia* 32: 11–25.

Castro Alves, Flávia de. 2010. Evolution of Case-Marking in Timbira. *International Journal of American Linguistics* 76: 439–475.

Castro Alves, Flávia de. 2015. The Grammar of Auxiliaries in Canela: Description and Reconstruction. Paper presented at the Workshop in Diachronic Morphosyntax in South American Languages, Collegium de Lyon & Laboratoire Dynamique Du Langage, Lyon, France, 28–30 May.

Castro Alves, Flávia de & M. A. Reis Silva. 2007. Marcação de caso e aspecto progressivo em Canela e Mebengokre. Paper presented at the V Encontro de Línguas e Culturas Macro-Jê, Universidade de São Paulo, 3–5 May.

Comrie, Bernard. 1989. *Language Universals and Linguistic Typology: Syntax and Morphology* Oxford: Blackwell.

Croft, William. 2003. *Typology and Universals*. Cambridge: Cambridge University Press.

Dixon, R. M. W. 1979. Ergativity. *Language* 55: 59–138.

Dixon, R. M. W. 1994. *Ergativity*. Cambridge: Cambridge University Press.

Donohue, Mark & Søren Wichmann, eds. 2008. *The Typology of Semantic Alignment.* Cambridge: Cambridge University Press.

Dourado, Luciana. 2001. Aspectos morfosintáticos da língua Panará (Jê). Ph.D. dissertation, Universidade Estadual de Campinas.

Dryer, Matthew S. 2007. Clause Types. *Language Typology and Syntactic Description*, vol 1, ed by T. Shopen, 224–275. Cambridge: Cambridge University Press.

Estival, Dominique & John Myhill. 1988. Formal and Functional Aspects of the Development from Passive to Ergative Systems. *Passive and Voice*, ed. by M. Shibatani, 441–491. Amsterdam: John Benjamins.

Gildea, Spike. 1993a. The Rigid Postverbal Subject in Panare: A Historical Explanation. *International Journal of American Linguistics* 59: 44–63.

Gildea, Spike. 1993b. The Development of Tense Markers from Demonstrative Pronouns in Panare (Cariban). *Studies in Language* 17: 53–73.

Gildea, Spike. 1998. *On Reconstructing Grammar: Comparative Cariban Morphosyntax*. Oxford: Oxford University Press.

Gildea, Spike. 2000. On the Genesis of the Verb Phrase in Cariban Languages: Diversity Through Reanalysis. *Reconstructing Grammar: Comparative Linguistics and Grammaticalization Theory*, ed. by Spike Gildea, 65–106. Amsterdam: John Benjamins.

Gildea, Spike. 2003. The Venezuelan Branch of the Cariban Language Family. *Amérindia* 28: 7–32.

Gildea, Spike. 2004. Are There Universal Cognitive Motivations for Ergativity? *L'ergativité en Amazonie*, vol. 2, ed. by Francesc Queixalós, 1–37. Brasília: CNRS, IRD and the Laboratório de Línguas Indígenas, UnB.

Gildea, Spike. 2012. Linguistic Studies in the Cariban Family. *Handbook of South American Languages*, ed. by Lyle Campbell & Veronica Grondona, 441–494. Berlin: Mouton de Gruyter.

Gildea, Spike & Flávia de Castro Alves. 2010. Nominative-Absolutive: Counter-Universal Split Ergativity in Jê and Cariban. *Ergativity in Amazonia*, ed. by Spike Gildea & Francesc Queixalós, 159–199. Amsterdam: John Benjamins.

Gildea, Spike, Berend Hoff & Sérgio Meira. 2010. The story of *ô in the Cariban family. *Fieldwork and Linguistic Analysis in Indigenous Languages of the Americas*, ed. by Andrea L. Berez, Daisy Rosenblum & Jean Mulder. *Language Documentation & Conservation* Special Publication No. 2, 91–123.

Gilij, Filippo S. 1780–1784. *Saggio di Storia Americana*, vols. 1–3. Rome. [1965. *Ensayo de Historia Americana*, vols. 1–3. Trans. Antonio Tovar. Caracas: Italgráfica].

Givón, Talmy. 1979. *On Understanding Grammar*. New York: Academic Press.

Givón, Talmy. 1980. The Drift away from Ergativity: Diachronic Potentials in Sherpa. *Folia Linguistica* 4: 41–60.

Givón, Talmy. 2001. *Syntax: An Introduction*, vols. 1–2. Amsterdam: John Benjamins.

Givón, Talmy. 2009. *The Genesis of Syntactic Complexity*. Amsterdam: John Benjamins.

Goldberg, Adele E. 2006. *Constructions at Work: The Nature of Generalization in Language.* Oxford: Oxford University Press.

Guillaume, Antoine. 2015. *Reconstructing the Morphology of the Pronominal Systems of Takanan Languages*. Paper Presented at the Workshop in Diachronic Morphosyntax in South American Languages, Collegium de Lyon & Laboratoire Dynamique Du Langage, Lyon, France, 28–30 May.

Harris, Alice C. & Lyle Campbell. 1995. *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: Cambridge University Press.

Haspelmath, Martin. 2011. On S, A, P, T, and R as Comparative Concepts for Alignment Typology. *Linguistic Typology* 15: 535–567.

Hawkins, Robert E. 1998. Waiwai. *Handbook of Amazonian Languages*, vol. 4, ed. by Desmond C. Derbyshire & Geoffrey K. Pullum, 3–202. Berlin: Mouton de Gruyter.

Heine, Bernd. 1994. Grammaticalization as an Explanatory Parameter. *Perspectives on Grammaticalization*, ed. by W. Pagliuca, 255–287. Amsterdam: John Benjamins.

Heine, Bernd & Tania Kuteva. 2002. *World Lexicon of Grammaticalization.* Cambridge: Cambridge University Press.

Hewitt, Brian G. 1987. *The Typology of Subordination in Georgian and Abkhaz.* Berlin: Mouton de Gruyter.

Hoff, J. Berend. 1968. *The Carib Language*. The Hague: Martinus Nijhoff.

Koehn, Edward & Sally Koehn. 1986. Apalai. *Handbook of Amazonian Languages*, ed. Desmond C. Derbyshire & Geoffrey K. Pullum, vol. 1, 33–127. Berlin: Mouton de Gruyter.

Koptjevskaja-Tamm, Maria. 1993. *Nominalizations*. London & New York: Routledge.

Malchukov, Andrej L. 2010. "Quirky" Case: Rare Phenomena in Case-Marking and their Implications for a Theory of Typological Distributions. *Rethinking Universals: How*

Rarities Affect Linguistic Theory, ed. by J. Wohlgemuth & M. Cysouw, 139–169. Berlin: Mouton.

Mattéi Muller, Marie-Claude. 1994. *Diccionario Ilustrado Panare-Español, indice Español-Panare*. Caracas: Comisión Quinto-Centenário.

Mattéi Muller, Marie-Claude. 2007. Voir et Savoir en Panaré (langue caribe du Venezuela). *L'énonciation médiatisée II,* ed by Z. Guentcheva & J. Landaburu, 153–170. Louvain-Paris: Peeters.

Meira, Sérgio. 1999. *A Grammar of Tiriyó*. Ph.D. dissertation, Rice University.

Meira, Sérgio. 2000. The Accidental Intransitive Split in the Cariban Family. *Reconstructing Grammar: Comparative Linguistics and Grammaticalization*, ed. by Spike Gildea, 201–230. Amsterdam: John Benjamins.

Meira, Sérgio & Bruna Franchetto. 2005. The Southern Cariban Languages and the Cariban Family. *International Journal of American Linguistics* 71: 127–192.

Nonato, Rafael. 2014. *Clause Chaining, Switch Reference and Coordination*. Ph.D. dissertation, Massachusetts Institute of Technology.

Oliveira, Christiane C. 2003. Lexical Categories and Descriptives in Apinajé. *International Journal of American Linguistics* 69: 243–274.

Oliveira, Christiane C. 2005. *The Language of the Apinajé People of Central Brazil*. Ph.D. dissertation, University of Oregon.

Payne, Doris L. 1994. OVSu Versus VSuO in Panare (Cariban): Do Syntax and Discourse Match? *Text* 14: 581–610.

Payne, Thomas E. 1990. Ergativity and Transitivity in Panare (Carib). *Amazonian Linguistics: Studies in Lowland South American Languages*, ed. by Doris L. Payne 429–453. Austin: University of Texas Press.

Payne, Thomas E. 1997. *Describing Morphosyntax: A Guide for Field Linguists*. Cambridge: Cambridge University Press.

Payne, Thomas E. & Doris Payne. 1999. *Panare: A Cariban Language of Central Venezuela*. 194pp. ms.

Payne, Thomas E. & Doris Payne. 2013. *A Typological Grammar of Panare: A Cariban Language of Central Venezuela*. Leiden: Brill Press.

Popjes, Jack & Jo Popjes. 1986. Canela-Krahô. *Handbook of Amazonian Languages*, Vol. 1, ed. by Desmond C. Derbyshire & Geoffrey K. Pullum, 128–99. Berlin: Mouton de Gruyter.

Queixalós, Francesc & Spike Gildea. 2010. Manifestations of Ergativity in Amazonia. *Ergativity in Amazonia*, ed. by S. Gildea & F. Queixalós, 1–25. Amsterdam: John Benjamins.

Reis Silva, Maria A. 2001. *Pronomes, ordem e ergatividade em Mẽbêngôkre (Kayapó)*. MA thesis, Universidade Estadual de Campinas.

Ribeiro, Eduardo & Hein van der Voort. 2010. Nimuendajú Was Right: The Inclusion of the Jabutí Language Family in the Macro-Jê Stock. *International Journal of American Linguistics* 76: 517–570.

Rodrigues, Aryon D. 2009. A Case of Affinity among Tupí, Karíb, and Macro-Jê. *Revista Brasileira de Linguística Antropológica* 1(1): 137–162.

Salanova, Andrés P. 2007. Nominalizations and Aspect. Ph.D. dissertation, Massachussets Institute of Technology.

Salanova, Andrés P. 2008. Uma análise unificada das construções ergativas em Mẽbêngôkre. *Amérindia* 32: 109–134.

Santos, Ludoviko C. 1997. Descrição de aspectos morfossintáticos da língua Suyá/Kisêdjê (Jê). Ph.D. Dissertation, Universidade Federal de Santa Catarina.

Santos, Ludoviko C. 1999. Aspectos do Sistema de Marcação de Caso da Língua Suyá. *Signum: Estudos da Linguagem* 2: 231–242.

Stout, Mickey & Ruth Thomson. 1974. Modalidade em kayapó. *Série Linguística* 3: 69–97.

Tavares, Petronila S. 2005. A Grammar of Wayana. Ph.D. dissertation, Rice University.

Traugott, Elizabeth C. & Graeme Trousdale. 2013. *Constructionalization and Constructional Change*. Oxford: Oxford University Press.

Urban, Greg. 1985. Ergativity and Accusativity in Shokleng (Gê). *International Journal of American Linguistics* 51: 164–187.

Van de Velde, Freek, Hendrik de Smet, & Lobke Ghesquière. 2013. On Multiple Source Constructions in Language Change. *Studies in Language* 37: 473–489.

Wiesemann, Ursula. 1972. *Die Phonologische und Grammatische Struktur der Kaingáng-Sprache*. Paris: Mouton.

Wiesemann, Ursula. 2002. *Kaingáng-Português: Dicionário Bilíngue*. Curitiba: Editora Evangélica Esperança.

Willis, David. 2011. Reconstructing Last Week's Weather: Syntactic Reconstruction and Brythonic Free Relatives. *Journal of Linguistics* 47(2): 407–446.

Conducting Syntactic Reconstruction of Languages with No Written Records

Kikusawa Ritsuko

Abstract

This article focuses on the methodology for syntactic reconstruction in languages without a written record from the past. The idea is to follow the principles of the Comparative Method, the scientific procedure to compare and reconstruct sounds and lexical items in various proto-languages. The method originally developed out of the comparison and reconstruction of classic languages in Indo-European languages, but has been successfully applied to Austronesian languages, where information about old forms of languages is hardly available from literature. The claim in this article is that there are ways to conduct syntactic reconstruction with languages without a written record. It is shown that, by using correct comparanda and by combining structural analyses with results of sound and lexical reconstruction, clause structures of such languages can be compared and reconstructed, and the developmental paths from one system to another can be traced.

1 Introduction

Syntactic reconstruction was once considered a hazardous if not impossible endeavour, however, the interest these days is more on how to carry out syntactic reconstruction rather than whether it is possible or not.[1] Watkins' (1964) discussion of the reconstruction of "Indo-European sentence structure," has resulted in considerable work being published (Barðdal 2014, Fried 2008, Harris

---

1 This research was conducted during my stay at the University of Ghent in 2015 in association with the EVALISA (The Evolution of Case, Alignment and Argument Structure in Indo-European) Project funded by the European Research Council. I would like to thank Jóhanna Barðdal, the principal investigator of the project and the host, and all the team members and staff members of the university for fruitful academic interactions and their support while I was there. I would also like to thank Lawrence A. Reid, two anonymous reviewers and the editors for their comments on an earlier version of this article.

2008, Gildea 1998 and others), and aspects related to the methodology have been gradually elaborated (Barðdal & Eythórsson in this volume). However, the application of the method to languages without written records is still limited and yet to be established. My aim in this article is, by taking clause structures of Austronesian languages as an example, to demonstrate ways to conduct syntactic comparison of data exclusively from modern languages for the purpose of diachronic reconstruction.

In the comparison and reconstruction of data from languages without a written record, some basic principles and methodology are commonly shared with those languages with philological data. This article follows the basic principles proposed in the research on languages from other families. First, comparanda (comparable units, cf. Ferraresi & Goldbach 2008) must be of surface structure, where changes are directly observable. Second, because each grammatical change is gradual and discrete from other grammatical changes (Fried 2008: 48, Roberts 2007: 6, De Smet 2015), when examining the historical development of a linguistic structure, the linguistic features that form part of the examined structure are decomposed and analysed separately. Thus, changes in pronouns, marking on lexical noun phrases, verb morphology, word order and others are all examined separately (more discussion in Kikusawa 2017). Keeping these principles, the practical procedure applied here is as follows: i) describe the basic clause structures (abstracting relevant patterns) for each language examined, and ii) identify cognacy between the languages, among the structures described; iii) clarify the differences, discuss the changes that brought about the differences, then identify the direction of change. This is also in line with what has been proposed in previous studies on syntactic reconstruction.

In examination of each of these stages, however, special approaches are required so that data from languages with no philological materials can be dealt with and correctly analysed. For example, the description of clause structures requires reanalyses of information available in the description of each language. This is because, typically, the description of each language follows and uses terms according to the type of system each language synchronically exhibits, which is discrete from inheritance. Also, the framework applied in the descriptions differs from language to language. Therefore, consistency is required for a cross-linguistic analysis. Another example is that when historical documentation is not available, the direction of change needs to be identified based on various scientifically based inferences. In the examination of changes that took place in argument structures in Indo-European languages, documented clauses are compared and the cognacy of the compared clauses

is secured by the cognancy of the forms comprising the clauses compared (cf. Barðdal & Eythórsson 2012).

This article is about how these problems are overcome in a specific case of syntactic reconstruction in Austronesian languages, and about how the applied methodology could be generalised. In discussing practical aspects of syntactic reconstruction, Barðdal states that syntactic reconstruction must "abide [by] certain procedural requirements, of which the first one is to identify the cognates, the next is to set up correspondence sets, and the third is to model the reconstructed material with adequate formal tools" (Barðdal 2014: 367). Here, cognate structures are identified by relating the structures with forms reconstructed applying the Comparative Method, the standard method for comparing and reconstructing lexical and morphological items. By identifying the correct morphological component as an anchor for tracing the inherited positions in each structure, it is shown that cognate structures, even when they look completely different today, can be identified without any historical record. Once correspondence sets are determined, various changes can be identified. It is shown that the results of such an endeavour indeed enable us to explain historically various morphological and syntactic phenomena in the modern languages of the family.

The rest of this article is organised as follows. In Section 2, an overview of syntactic reconstruction in the context of Austronesian historical linguistics will be provided. In Section 3, clause structures of five distantly related Austronesian languages are compared. The "basic clause structures" of each language are first schematically represented for the purpose of syntactic comparison and reconstruction, and then by examining the positions where genitive pronouns occur, cognacy among the clause structures is identified. Based on the identified cognacy, it is shown that change in the case alignment systems can be reconstructed and developmental paths described. In Section 4, the results of the examination presented in Section 3 are extended to identify changes that brought about other syntactic phenomena in this language family. Two cases are presented, namely, word-order change and the development of part of the verb morphology. Section 5 provides a summary and concluding remarks.

2   Syntactic Reconstruction and Austronesian Languages

Austronesian languages, consisting of some 1,200 languages spoken in the Pacific and surrounding areas, have insufficient historical text documentation and thus historical comparison and reconstruction is only conducted by

comparing data from modern languages.[2] The Austronesian language family is often referred to as an exemplary case of application of the Comparative Method. As a result, while the languages are known to be typologically diverse, the general subgroup membership of languages is fairly clear. This makes Austronesian languages a good subject for syntactic comparison and reconstruction, not only for clarifying the developmental paths of their syntactic features, but also for developing the methodology of syntactic reconstruction. In addition, many languages in this family are morphologically complex, and their syntax commonly has a morphological correlate, a condition which is considered to favour the reconstructibility of syntactic structures (Harris 2008: 91). However, in traditional Austronesian historical linguistics, on the one hand, "grammatical reconstruction" has typically referred to lexical comparison and the reconstruction of grammatical or functional forms (e.g. Pawley 1970); on the other hand, attempts at direct comparison and reconstruction of linguistic components with a structural nature (or patterns) has been limited.

In the comparison and reconstruction applied to Austronesian languages demonstrated in this chapter, following the principles outlined in Section 1, patterns found in languages are directly compared and reconstructed, and in doing so, the traditional Comparative Method is consciously utilised. For example, reflexes of the earlier genitive pronominal set (hypothesised as having also marked ergative agents of transitive constructions, see 3.1) are regarded as most suitable for tracing clause pattern change and thus are used to identify cognate structures in this study. The reconstruction of pronominal forms, however, is conducted separately from clause structure comparison. But by combining the results of the two, it becomes possible to identify the direction of change in the clause structures, and to clarify the merger and split of syntactic functions associated with positions in a clause in each stage of the development.

One may consider that verb morphology may better serve the same purpose. However, despite the elaborate morphological systems in western Austronesian languages, their supposed remnants are limited in Oceanic languages (cf. Starosta, Pawley & Reid 1982), while the reflexes of the earlier

---

2   Text materials of old forms of the languages include those of Old Cham (4th century AD), Old Malay (7th century AD), and Old Javanese and Balinese (9th Century AD) (Adelaar 2005, Blust 1995). However, because of the phonological uncertainty of the scripts and limited textual content, they do not make good source materials for comparative syntax. Gilles-Maurice de Schryver (pers. comm.) points out that descriptions of languages from colonial periods (16th century on, by Spanish, Dutch and other European visitors) and also translations of the Bible should be considered as useful resources, a perspective missing in my previous research.

genitive pronouns are traceable in most branches of the family. This procedure does not limit the results to the understanding of change in pronominal arguments and case related changes, but further enables us to compare and reconstruct other grammatical features, including verb morphology. To illustrate this, two case studies are presented in this chapter, namely, a word order change, and a change in the distribution of transitive/applicative verb suffixes. My aim is to demonstrate that, once cognacy is identified, specific changes and their directionality can also be identified by applying the principles of the Comparative Method that are used for lexical comparison and reconstruction.

In this article, five selected languages, namely Ibaloy, Pendau, Taba, Fijian, and Tongan, are compared. These are all daughter languages of Proto-Malayo-Polynesian (Figure 3.1), however, they are only distantly related and are spoken in areas that are geographically not adjacent to one another. The advantages of conducting such a macro-comparison is that, as discussed by Kikusawa (2018), it makes it easier to identify direct inheritance. Historical examination of closely related languages (micro-comparison) is often complicated by the mixture of direct and indirect inheritance (i.e., borrowing from closely related languages), as well as sporadic local innovations (forming areal features), where earlier features are obscured by layers of change that have taken place subsequent to the split of the languages.

It should be remembered that the comparison and reconstruction of lexical items and sound systems, which is usually conducted today by applying "bottom-up" methodology, was initially done by macro-comparison, which set the basis for detailed bottom-up micro-comparison (cf. Blust 1990: 137–138). Needless to say, follow-up modification of any proposed hypothesis is necessary, based on new data and the results of micro-comparison. What is presented here is macro-comparison as an initial attempt of investigating syntactic change.

3   Clause Structures and Their Cognacy

In this section, the methodology for comparing and reconstructing clause structures in Austronesian languages will be demonstrated, and the cognacy of clauses from five Austronesian languages will be shown. A working hypothesis is presented in 3.1, with some background linguistic information related to the analyses of the languages. The clause structures of five selected Austronesian languages are schematically described in 3.2. With analyses conducted specifically for comparison of clause patterns of languages with different typological

Austronesian

Eastern Formosan

Malayo-Polynesian
(Extra-Formosan)

(Western Malayo-Polynesian)*   Nuclear Malayo-Polynesian

Ibaloy
Tagalog

Sunda-Sulawesi (Central-Eastern)

Mamuju
Pendau
Indonesian

Central   Eastern
Malayo-Polynesian   Malayo-Polynesian

South-Halmahera/   Oceanic
West New Guinea

Taba

Admiralty Western   Central-Eastern
           Oceanic   Oceanic

Southeast   Eastern (Remote) Oceanic
Solomonic

Central Pacific

Fijian   Polynesian

Tongan

West Futuna-Aniwan   East Futunan, East Uvean

Samoan

Eastern Polynesian

Māori

* Western Malayo-Polynesian is not considered an established subgroup. It is a label referring to a group of languages that do not share the defining innovations of the Nuclear Malayo-Polynesian group.

FIGURE 3.1 Languages referred to in this article (italic font and boxed) and their proposed
           genetic relationships

systems, it will be shown that the clause patterns are abstracted in a systematic way. In 3.3, how this leads to identifying cognate structures and also reconstruction of the developmental paths will be demonstrated. Understanding clause structure change sometimes results in understanding change in the functions of relevant grammatical forms. It will be shown in 3.4, how the findings presented in 3.3 help to re-evaluate previous lexical reconstructions of pronominal systems.

3.1   *Working Hypothesis*

A working hypothesis to be tested in this study is that the parent language of the five languages compared in this article, namely, Proto-Malayo-Polynesian (PMP, sometimes referred to as Proto-Extra-Formosan), was an ergative language. The Agent (A) of a bivalent (syntactically transitive) clause was expressed by a genitive pronoun, while the Subject (S) referring to the actor/ undergoer of intransitive clauses, and the Patient (P) of transitive clauses were expressed by a nominative pronoun. The abbreviations used to indicate the syntactic roles of the arguments of transitive and intransitive clause follow Comrie (1989) and Dixon & Aikhenvald (2000). A full list of abbreviations is given at the end of this chapter.

It should be noted that determining the case of an Austronesian language in the description of the language is usually based on typological criteria. The form or marking on S is by definition nominative. If A receives a different marking from that on S, it is typically labelled as ergative, while if it is P that receives a different marking, it is typically labelled as accusative. The marking on A and the form of the associated pronouns is often shared in Austronesian languages with that of the possessor of a noun in noun phrases and is consequently labelled as genitive (rather than ergative). It should be noted that, unlike Indo-European case labels, such typologically defined terms do not necessarily reflect etymological relationships, and the functional change of each case needs to be traced, based on formal correspondences.

The PMP ergative system is schematically shown in Figure 3.2. The pronoun that expressed A is referred to as genitive, for it also occurred on noun phrases expressing a possessor. Clause structures reflecting this system as well as example clauses are presented in (1) and (2) respectively. Example (2a) is an intransitive clause where the S is expressed as a nominative, while (2b) is a transitive clause where the A appears as genitive, while the P is marked as nominative.
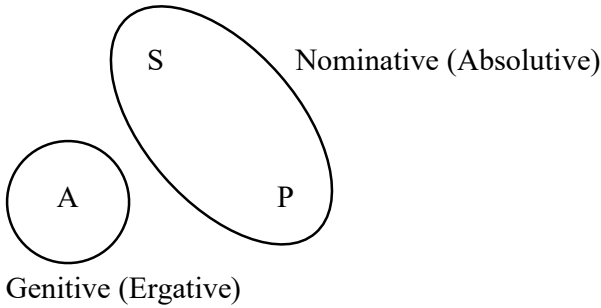
FIGURE 3.2 The PMP case-marking system based on Starosta, Pawley & Reid. (1982) and
            Reid (2009, 2016)

(1)   Assumed PMP clause structures (1) Intransitive and transitive

        INTRANSITIVE   S
                           NOM
                           actor/undergoer

        TRANSITIVE   A   P
                           GEN   NOM
                           actor   undergoer

(2)   Ivatan examples illustrating clause structures shown in (1) (Reid 1966:
        143, data modified following L. A. Reid pers. comm.)

        a. INTRANSITIVE
           mangay   [*ʔo   tao*]ₛ.
           go   NOM   man
           'The man is going.'

        b. TRANSITIVE
           ʔamoʔmohen   [*no   tao*]ₐ   [*ʔo   motdeh*]ₚ.
           frighten   GEN   man   NOM   child
           'The man is frightening the child.'

In PMP, in addition to the two clause structures presented in (1), it is assumed
that there was a dyadic intransitive – Dixon's extended (E) intransitive –
structure. This is a structure which can be described as semantically transi-
tive and syntactically intransitive. Although semantically two participants are
involved, the verb morphology is typically the same as that of the monadic
intransitive clause. In addition, the NP expressed by the nominative case is

identical to the sole NP of a monadic intransitive clause.[3] The clause structure of a dyadic intransitive is shown in (3). An example is given in (4) where the S is expressed in nominative while the E is expressed in oblique.

(3)   Assumed Proto-Malayo-Polynesian clause structure (2) Dyadic intransitive
       DYADIC INTRANSITIVE   S   E
                                  NOM   OBL/LOC
                                  actor   undergoer

(4)   Ivatan (Reid 1966: 22–23, data modified following L. A. Reid pers. comm.)
       DYADIC INTRANSITIVE
       mang-amoʔmo   [*ʔo*   *tao*]$_S$ [*so*   *motdeh*]$_E$.
       frighten   NOM   man   OBL   child
       'The man is frightening a child.'

The E of a dyadic intransitive and the P of a transitive construction are both considered to have carried the semantic macro-role "undergoer" that expressed patient, location, goal, instrument, beneficiary, and other functions. This system is similar to that in many Philippine and Formosan languages today. The full set of assumed PMP clause structures is shown in (5). For more examples illustrating the system, see 3.2.1, particularly (7–9).

(5)   Proto-Malayo-Polynesian clause types based on Starosta, Pawley & Reid
       (1982).
       i.   INTRANSITIVE   S
                                      actor/undergoer
       ii.   DYADIC INTRANSITIVE   S   E
                                      actor   undergoer
       iii.   TRANSITIVE   A   P
                                      actor   undergoer

---

3   Dyadic intransitive clauses may appear to be similar to antipassives, however, they are different constructions. In this article, the former is analysed as one of the basic sentence structures, while the latter is a derived construction and therefore excluded from the discussion. In both structures, the actor (S) is expressed with nominative forms, and the undergoer (E) is expressed with oblique/genitive forms. However, the two are different in that: i) dyadic intransitive clauses may contain the same morphological forms as those occurring in monadic intransitives, such as *mang-* in the verbs in (2a) and (4), while antipassive clauses may contain verb morphology shared with a corresponding transitive clause; ii) the semantic property of the arguments does not always match between a dyadic intransitive clause and a corresponding transitive clause (e.g. 'the child' and 'a child' in (2b) and (4)), while it does between a transitive clause and its derived antipassive clause. In the two structures, for example in Kalanguya, spoken in Northern Luzon, the difference is clearly reflected in verb morphology (Santiago 2016).

The hypothesis that PMP had an ergative system with structures i. through iii. is to be tested against data from modern languages. As the first step for doing so, clause structures in some daughter languages and the function of the possible reflexes (or the remnants) of the earlier genitive pronoun set are identified. The position where such pronouns are found is referred to in this study as the "X-position". In what follows, the basic structures of five Austronesian languages, namely, Ibaloy, Pendau, Taba, Fijian and Tongan, are described and the X-position of each language is examined.

To identify the X-position in each language, the forms of pronouns expressing A are first examined against the reconstructed pronominal forms presented in Table 3.1. In addition, the pronominal forms expressing A are compared

TABLE 3.1   Reconstructed Proto-Malayo-Polynesian clitic pronouns and their variants (based on Blust 1977, Blust & Trussel on-going, Reid 2016) [a,b]

	1MIN	1+2MIN	2MIN	3MIN	1AUG	1+2AUG	2AUG	3AUG
GENITIVE[c]	*=ku	*=ta	*=mu	*=ya	*=mi	*=tamu ~ *=tamuyu ~ *=tayu	*=muyu ~ *=yu	(*da)
NOMINATIVE[d]	*=aku~ *=ak	*=ta	*=kaʔu~ *=kaw ~ *=ka	*Ø	*=kami	*=takamu	*=kamu ~ *=kamuyu ~ *=kayu	(*ida)

a   The terms "minimal" and "augmented" are used for first and second person pronouns to better capture the paradigmatic systematicity of first person inclusive dual (1+2MIN), first person inclusive plural (1+2AUG), and first person exclusive plural (1AUG) pronouns, see Cysouw (2003).

b   Forms for the 1st and 2nd persons are from Reid (2016). Multiple forms are reconstructed when Reid considers that reflexes found in modern languages imply that there were such variants in PMP (Reid 2016). Forms for the 3rd person are based on Blust (1977), Blust & Trussel on-going. The 3AUG forms are independent pronouns and not clitics and are thus shown in parentheses. The forms listed in the original literature are: *i=ya/ni-ya '3SG' (Blust 1977) and *ida 'they, them' (Blust & Trussel on-going).

c   A genitive form could occur as an enclitic to a verb, to function as the actor of a transitive clause, or to a noun, to function as a possessor. It could also appear encliticised to one of the genitive specifiers, PMP *ni ~ *ʔi, and become the complement of an oblique/locative/dative specifier.

d   A nominative form could occur as an enclitic to a verb, to function as the actor or undergoer of an intransitive clause. It could also appear encliticised to the personal noun specifier (PMP *si), or to the topic specifier (PMP *ʔi), and function as an independent pronoun expressing the grammatical subject of a transitive clause, a predicate, or a fronted topic, etc.

with the current genitive forms marking the possessor of the noun. The current function of the reflex/remnant of a pronominal set may not be exclusively marking A. In such cases, the distribution of the pronominal set under examination needs to be described and the developmental path has to be discussed. Reflex forms do not always compose a full set, however, it is often found that a few clear reflexes of the earlier genitive forms express a related function. In this article, I limit the discussion here to clause structures with pronouns, and those with lexical arguments are referred to only when relevant.

Many modern Austronesian languages, including the five languages that are described in this section, show either a reflex set of the reconstructed PAn genitive/ergative set, or a remnant of it. These forms occur as enclitics on the verb, marking either A showing an ergative pattern, or marking A of one of the two transitive clauses, or a semantic actor (A and S) showing an accusative pattern, as is shown in 3.2. The significance of these X-positions for morphosyntactic reconstruction and the direction of morphosyntactic change is further summarised and discussed in 3.3.

3.2 *Clause Structures and the X-Position in Five Austronesian Languages*
3.2.1 Ibaloy
Ibaloy is a language spoken in Northern Luzon in the Philippines. It shows an ergative system (Ruffolo 2004), where the A of a transitive clause is expressed with a genitive pronoun, identifying it as the X-position. Clause structures with pronominal arguments are presented in (6) with the X-position constituent bold, followed by examples in (7–9) demonstrating each clause structure. Analyses and glossing follow the source descriptions of each language, unless otherwise specified.

(6) Ibaloy argument structures with pronominal NPs (without an auxiliary verb)
    i.   Vi$[=\text{NOM}]_\text{S}$
    ii.   Vi$[=\text{NOM}]_\text{S}$ $[son\ \text{IND}]_\text{E}$
    iii. Vt$[\textbf{=GEN}]_\text{A}[=\text{NOM}]_\text{P}$

(7) Ibaloy examples of Clause (i) with pronominal arguments, without an
     auxiliary verb
     a. *ondawakda*
        ʔon-law[=ak]$_S$=la
        ACTV.IPF-go=1.NOM=toward
        'I went away' (Ruffolo 2004: 412)

     b. *naogip*   *ira*
        na-ʔogip   [ʔida]$_S$
        POTPATV.PFT-sleep   3+.NOM
        'They are asleep.' (Ruffolo 2004: 411)

Examples (7a–b) are monadic intransitive clauses. The core argument S is
expressed by a nominative pronoun. A clause with no expressed nominative
form is understood as having a 3rd person singular 'he/she/it' pronoun (Ruffolo
2004: 175). This is indicated by the symbol "=∅" in example sentences.

(8) Ibaloy examples of Clause (ii) with pronominal arguments, without an
     auxiliary verb
     on'aseba   *son*   *si'kato*
     ʔon-ʔasəwa[=∅]$_S$ [son[4] siʔgato]$_E$
     ACTV.IPF-marry=**3MIN/NOM**   OBL   3MIN.IND
     '**She** will get married to him.' (Ruffolo 2004: 150)

The example in (8) is a dyadic intransitive clause where the S is expressed by
a nominative pronoun like in monadic intransitive clauses. The E, expressing
the undergoer, is marked with the oblique marker *son*. An independent (case-
neutral) pronoun *siʔgato* follows this form. In both examples (7) and (8), mo-
nadic and dyadic intransitive clauses respectively, the verb carries the prefix
ʔon- on the verb, thus sharing the same morphology. This, along with the fact
that the S is marked by nominative in both structures, is one of the main rea-
sons why this structure is analysed as dyadic intransitive rather than syntacti-
cally transitive.

---

4 Author's interpretation. The original description by Ruffolo further breaks down the form
   son into *so=n* 'OBL=GEN/PERS'.

(9) Ibaloy examples of Clause (iii) with pronominal arguments, without an auxiliary verb

  a. *ensemektoka*
    ʔən-səmək[=**to**]ₐ[=ka]ₚ
    PotPatV.en-love=**3MIN.GEN**=2MIN.NOM
    '**He** loves you' (Ruffolo 2004: 175)

  b. *intongkaloanto*   *ira*   *ni*   *apag*[5]
    ʔin-toŋgal-an[=**to**]ₐ=j   [ʔida]ₚ [ni ʔapag]ₑ
    BNFV.PFT-buy-BNFV=**3MIN.GEN**   3+.NOM GEN meat
    '**He** bought them some meat' (Ruffolo 2004: 141)

Examples (9a–b) are transitive clauses. The third person augmented nominative form is typically not a clitic, cf. *ʔida* in (9b).

Ibaloy genitive pronouns are listed in Table 3.2. Ibaloy clitic pronouns are Wackernagel clitics and occur in the second position of a clause (Ruffolo 2004: 175). Clause structures with auxiliary verbs, which occur in clause initial position in basic clauses, are shown in (10) to illustrate this, where enclitic pronouns are encliticised to the auxiliary verb and precede the main verb.

TABLE 3.2   Ibaloy genitive pronouns (based on Ruffolo 2004: 175)

	1MIN	1+2MIN	2MIN	3MIN	1AUG	1+2AUG	2AUG	3AUG
Genitive (on N & V)	=k, =ko	=ta	=m, =mo	=to	=mi	=tajo	=jo	=da
Nominative (clitic)	=ak	=kita	=ka	∅	=kami	=kito, =kitajo	=kayo	(=)ʔida[a]

a   The form *ʔida* may appear either as a clitic or an independent form.

<hr>

5   L. A. Reid (pers. comm.) notes that the final *=j* on the 2nd interlinear line of this example is actually the initial component of the 3+.NOM pronoun that Ruffolo analyses as an independent pronoun in the first line of the example. Reid suggests that this pronoun is also an enclitic based on phonological reasons. On the other hand, Ruffolo describes some syntactic behaviours, particularly word order, that are exclusively associated with the 3AUG form and not with other bound pronouns (2004: 175–180). It appears that the characteristics associated with this form are the result of being in a transition stage between being an independent form and a clitic (relevant discussion appears in 4.1).

(10) Ibaloy argument structures with pronominal NPs (with Aux)
    i.   VAUX[=NOM]$_S$ Vi
    ii.   VAUX[=NOM]$_S$ Vi [*son* IND]$_E$[6]
    iii.   VAUX[=**GEN**]$_A$[=NOM]$_P$ Vt

Examples in (11) are sentences illustrating the structures with auxiliary verbs presented in (10). It can be seen that both =*ʔida* '3AUG.GEN', =*to* '3MIN.GEN' and =*ka* '2MIN.NOM' are cliticised to the auxiliary verb *ʔəg* 'negative' that occurs in clause initial position. It appears that the existence of the two structures shown in (6) and (10), contributed to the development of the types of argument structure that are found in some modern Austronesian languages today, that are referred to as ergative, inverse, accusative, etc., as presented in later sections.

(11) Ibaloy examples with pronominal arguments (with auxiliary verb)
    a. *eg'ira*   *ondaw*   *chima*   *pa'dok*
       ʔəg[=**ʔida**]$_S$ ʔon-law   dima   paʔlok
       NEG=3+.NOM   ACTV.IPF-go LOC.DIST   creek
       'They will not go to that creek.' (Ruffolo 2004: 178)

    b. *egtoka*   *kegtinan*
       ʔəg=[**to**]$_A$=[ka]$_P$ gətin-an
       NEG=3.GEN=1.NOM   step-LOCV.IPF
       'He will not step on you.' (Ruffolo 2004: 179)

Finally, the nominative NP may be expressed by a clitic pronoun as in (6), or a corresponding independent pronoun, in which case the clause structure can be described as in (12). Examples are provided in (13), where *siʔgato* '3MIN.NOM' expresses the S of intransitive and ditransitive clauses (13a–b) and the P of a transitive clause (13c).

---

6 Ruffolo analyses the oblique marking form *son* preceding independent pronouns as *so=n* 'oblique=genitive' where =*n* is the genitive form preceding a personal noun or pronoun. However, L. A. Reid (pers. comm.) points out that there is no clear evidence that this form developed from a sequence *so=nen* (the latter being the marker that elsewhere precedes genitive personal nouns in Ibaloy). Nevertheless, the form *ni* is found as a genitive specifier in other languages preceding personal nouns and pronouns, and it is also reconstructed for PMP. Reid suggests, therefore, that =*n* may be a remnant of that earlier form in Ibaloy.

(12) Ibaloy argument structures with independent pronouns (without Aux)
  i. Vi   [NOM/IND]$_S$
  ii. Vi   [NOM/IND]$_S$ [*son* IND]$_E$
  iii. Vt[=GEN]$_A$ [NOM/IND]$_P$

(13) Ibaloy examples demonstrating nominative independent pronouns
  a. *yet*   *mandotopay*   *si'kato*
   jət   man-loto=paj   [**siʔgato**]$_S$
   and then ACTV.IPF-cook=still 3MIN.IND
   'then she will still cook' (Ruffolo 2004: 174)

  b. *emandoto*   *si'kato* *ni* *timol*
   ʔəman-loto [**siʔgato**]$_S$ [ni timol]$_E$
   ACTV.CNTV-cook 3MIN.IND GEN pig.food
   'she is cooking some pig food' (Ruffolo 2004: 145)

  c. *amta* *ni* *daki* *si'kato*
   ʔamta [ni laki]$_A$ [**siʔgato**]$_P$
   know GEN man 3MIN.IND
   'the man knows her/him' (Ruffolo 2004: 419)

In short, i) Ibaloy pronouns may occur in different positions in relation to the main verb; ii) regarding nominative pronouns, either clitic or non-clitic forms may occur. The forms of the pronouns are morphologically different between genitive and nominative and they can thus be said to be morphologically case-marked. The A, when expressed by a pronoun, is consistently expressed by a genitive clitic pronoun and not an independent pronoun as in (14).

(14) Ibaloy example with a genitive clitic pronoun expressing the A
  Saknitantoy   *onas*
  saknit-an[=**to**]$_A$[=j   ʔonas]$_P$
  peel-LOCV.IPF=3MIN.GEN=NOM sugar.cane
  'He will peel the sugar cane.' (Ruffolo 2004: 146)

3.2.2   Pendau

Pendau, a language spoken in South Sulawesi and which probably belongs to the Tomini-Tolitoli group, has been analysed as showing what is referred to as an "inverse" system by Quick (2007). Clause structures with pronominal arguments in Pendau are shown in (15). This language has two dyadic clause

structures, which are referred to by Quick as "active voice" (av) (15b) and "inverse voice" (iv) (15c). Between these, the A of an inverse clause is expressed by a genitive pronoun and all other pronominal arguments are expressed by what are labelled as "absolute" pronouns. The term "absolute" as used here follows that in Quick (2007). The position following the main verb in inverse voice (bold) is identified as the X-position in Pendau.

(15) Pendau argument structures with pronominal NPs (an inverse system) (Quick 2007: 123)
    i. [ABS]$_S$ Vi
    ii. [ABS]$_A$ Vav [ABS]$_P$
    iii. [ABS]$_P$ Viv[(=)GEN]$_A$

Examples are presented in (16) below. In (16a), the sole argument S is expressed by the 3SG absolute pronoun *io*. In (16b), an actor voice clause, two absolute pronouns occur, one preceding and the other following the verb. The functions of the two arguments are determined by their relative position to the verb, namely, the one preceding the verb *io* '3SG.ABS', expresses A, while the one following the verb *ʔaʔu* '1SG.ABS', expresses P. Example (16c) is also a dyadic clause, however, it differs from (16b) in two respects. First, one of the forms of the arguments expressing the A is a genitive clitic pronoun, instead of an absolute pronoun. Second, the relative position of the A and P are reversed, with the argument expressing the P now preceding the verb, while that expressing the A follows the verb. Thus, in (16c), for example, *io* '3SG.ABS' preceding the verb expresses the P of the clause, while *=ʔu* '1SG.GEN' encliticised to the verb expresses the A, and the clause means 'he (A) left me (P)'. In (16d), on the other hand, *ʔaʔu* '1SG.ABS' preceding the verb expresses the P while *=nyo* '3SG.GEN' encliticised to the verb expresses the A and the clause means 'I (A) left him (P)'.

(16) Pendau examples with pronominal arguments[7]
    a. ... *Paey*   *io*   *nopoʔoro,* ...
        paey   [*io*]$_S$ n-popo-ʔoro
        and.then   3SG.ABS   R-SF.POS-stand
        'And then, he stood up ...' (Quick 2007: 140)

---

7   In Quick (1994) the absolute and genitive cases are labelled as proximate and obviate respectively. The glossing here follows that in Quick (2007).

b. *Io*   *nengebiling*   *'a'u*
   [Io]$_A$ neng-ebiling   [ʔaʔu]$_P$
   3SG.ABS   AV.R-leave   1SG.ABS
   'He left me.' (Quick 1994: 467)

c. *Io*   *niebilingo'u*
   [Io]$_P$ ni-ebiling[=**ʔu**]$_A$
   3SG.ABS   IV.R-leave=1SG.GEN
   'I left him.' (Quick 1994: 467)

d. *'a'u*   *niebiling**onyo***
   [ʔaʔu]$_P$ ni-ebiling[=**nyo**]$_A$
   1SG.ABS   IV.R-leave=3SG.GEN
   'He left me.' (Quick 1994: 467)

The two sets of pronouns that express arguments, namely "genitive" and "absolute" are listed in Table 3.3.

Genitive pronouns are either enclitic (1SG, 2SG, 3SG and 1INPL) or non-bound (1EXPL, 2PL, 3PL). Those occurring on the verb expressing the A argument show a slight difference from those occurring on nouns, in that two forms (1SG and 2SG) may occur as a prefix rather than an enclitic, giving a variant to clause structure iii, as in (17).

TABLE 3.3   Pendau pronouns

		1SG	2SG	3SG	1INPL	1EXPL	2PL	3PL
Genitive	on N (marking possessor)	=ʔu	=mu	=nyo	=to	mami	miu	nijimo
	on V (marking A)ᵃ	=ʔu ʔu-, noʔu-	=mu mu-	=nyo	=to	mami	miu	nijimo
Absolute		aʔu/haʔu	ooᵇ	io	ito	ami	emu	jimo

a   In an inverse clause where the A of the clause is 1SG or 2SG, either a prefixed pronominal form (*ʔu-/noʔu-* or *mu-*) or an enclitic pronominal form (=*ʔu* or =*mu*) occurs. According to Quick (2007: 374), the prefixed pronominal forms are portmanteau pronouns, carrying information as to the person and number of A, as well as that of tense (realis/irrealis distinction).

b   Word initial glottal stop is not indicated in this table following the orthography in Quick (2007).

(17)   Pendau argument structures with a prefixed genitive/ergative pronoun
       iii. *var.* $[\text{ABS}]_\text{P}$ $[\text{GEN-}]_\text{A}$ Viv

In addition to the word order shown in (15), absolute NPs may occur follow-
ing all the other constituents of the clause. This is shown in (18). Based on a
comparison with Ibaloy, I claim that those shown in (18) are the earlier word
order in Pendau and those shown in (15) are more innovative forms. See 4.1 for
discussion.

(18)   Pendau argument structures with pronominal NPs (2) (based on Quick
       2007, in particular, pp. 365–366)
       i-2.   Vi $[\text{ABS}]_\text{S}$
       ii-2. Vav $[\text{ABS}]_\text{P}$ $[\text{ABS}]_\text{A}$
       iii-2. Viv$[(=)\text{GEN}]_\text{A}$ $[\text{ABS}]_\text{P}$

3.2.3   Taba

Taba is spoken in southern Halmahera in North Maluku province of Indonesia
(Bowden 2001). This language has a set of cross-referencing forms expressing
the person and number of part of S (actor S but not undergoer S) and A. The
P of a transitive clause does not have any cross-referencing on the verb, and is
simply expressed by an independent pronoun or a lexical noun phrase. In ad-
dition, Taba has a dyadic intransitive structure, which is referred to by Bowden
as "semi transitive". The clause structures in (19) are summaries, based on the
clause types listed by Bowden (2001: 102). Among the seven clause structures
that Bowden describes, structures i to iv occur with "underived" root verbs,
while structures v to vii are derived constructions with verbs with applicative
suffixes. In this study, we focus on the four clause structures that are underived.

    As can be seen in (19), Bowden lists two intransitive clause structures (i and
ii), a "semi-transitive" structure (iii) and a transitive structure (iv). In an "un-
dergoer intransitive" clause (i), the S of the clause is typically an undergoer of
the event. A cross-referencing form indicating the person and number of the
undergoer may optionally appear on the verb in this structure. In an "actor
intransitive" clause (ii), the S is typically the actor of the event. Unlike in un-
dergoer intransitive clauses, a cross-referencing form indicating the person
and number of the actor obligatorily appears on the verb. A "semi transitive"
clause (iii) has the same structure as the actor intransitive clause in that the S
is an actor and a cross-referencing form indicating the person and number of
the actor appears on the verb. In addition, a locative complement phrase ex-
pressing nE follows the verb in this structure. The locative complement phrase

may be optionally followed by a locative post position *li*. A transitive clause
(iv) consists of an independent noun phrase expressing the A, which is cross-
referenced on the verb and another independent noun phrase expressing the
P following the verb.

(19) Taba cross-referencing system

UNDERGOER INTRANSITIVE			Vi	$[\text{IND}]_S$		
ACTOR INTRANSITIVE	$[\text{IND}]_S$	$[\underline{\text{CR}=}]_S$	Vi			
SEMI TRANSITIVE	$[\text{IND}]_S$	$[\underline{\text{CR}=}]_S$	Vst	$[\text{IND}]_E$ (*li*)		
TRANSITIVE	$[\text{IND}]_A$	$[\underline{\text{CR}=}]_A$	Vt	$[\text{IND}]_P$		
NON-ACTOR BIVALENT	$[\text{IND}]_{P1}$		Vb	(P) $[\text{IND}]_{P2}$		
DIRECT DITRANSITIVE	$[\text{IND}]_A$	$[_{\text{CR}=}]_A$	Vdt	$[\text{IND}]_{P1}$	$[\text{NP}]_{P2}$	
REMOTE DITRANSITIVE	$[\text{IND}]_A$	$[_{\text{CR}=}]_A$	Vdt	$[\text{IND}]_{P1}$	(P) NP_{P2}	

Examples illustrating the first four structures are presented in (20), where the
numbers correspond to the structures listed in (19). Example (20)i is an ex-
ample of the undergoer intransitive clause structure. The 3SG independent
pronoun *i* occurs expressing the undergoer S. Example (20)ii is an example
of the actor intransitive clause structure. The 1SG independent pronoun *yak*
occurs expressing the actor S, and the 1SG cross-referencing form *k=* occurs
on the verb indicating the person and number of the actor. Example (20)iii is
an example of a semi transitive clause. A lexical noun phrase *Yanti* '(personal
name)' expresses the actor S, and a 3SG cross-referencing form *n=* occurs on
the verb, indicating the person and number of the actor. A locative comple-
ment phrase *um li* 'house locative' follows the verb. Finally, example (20)iv is
an example of a transitive clause. The 1EX.PL pronoun *am* expresses A and the
3PL independent pronoun *si* (indicated here as being cliticised to the verb) ex-
presses P. The 1EX.PL cross-referencing form *a=* occurs on the verb indicating
the person and number of A.

(20) Taba examples
     a. *Mapot   i.*
        be.heavy   3SG
        'It's heavy.' (Bowden 2001: 102)

     b. *Yak   kwom.*
        yak   k=wom
        1SG   1SG.CR=come
        'I've come.' (Bowden 2001: 187)

c. *Yanti   ncung   um   (li)*[8]
  yanti   n=sung   um   (li)
  Yanti   3SG.CR=enter house   (LOC)
  'Yanti entered the house.' (Bowden 2001: 102)

d. *Am   aamsi   do.*
  am   a=am=si   do
  1EX.PL   1EX.PL=see=3PL   REAL
  'We already saw them.' (Bowden 2001: 35)

One of the characteristics of this language, in contrast to Ibaloy and Pendau, is that there is only a single set of pronouns, namely independent pronouns (IND) in the language. However, in addition, there is a set of "cross-referencing" forms (CR=) that occur on verbs indicating the person and number of A and part of S.

What is interesting here is that the forms of at least some of these cross-referencing formatives share some forms with what Bowden refers to as the "possessive ligature". This implies that the cross-referencing forms developed from an earlier genitive pronominal set. The relevant forms of Taba pronouns are listed in Table 3.4. Explanations follow.

TABLE 3.4   Taba pronouns (cited from Bowden 2001: 271)

	1SG	2SG	3SG	1IN.PL	1EX.PL	2PL	3PL
Independent forms	*yak*	*au*	*i*	*tit*	*am*	*meu*	*si*
CR formatives	*k=*	*m=*	*n=*	*t=*	*a=*	*h=*	*l=*
possessive ligatures	*ni-k*	*ni-m*	*ni-∅*	*ni-t*	*am*	*meu*	*ni-di* ~*di*

---

8 I have been unable to find an example where both of the arguments are expressed by a pronoun for this construction. An example of a locative complement phrase expressed with a pronoun (*yak li*) can be seen below.

Malusa   nim   wlo   maduga   yak   li.
m=ha-lusa   nim   wlo   m=ha-duga   yak   li
2SG=CAUS-say   2SG.POSS liver 2SG=CAUS-only 1SG   LOC.
'You said your heart was only for me.' (Bowden 2001: 323)

A "possessive ligature" expresses the person and number of the possess- or in a possessive construction, connecting the noun phrase expressing the possessor and the possessee. Examples of possessive expressions in Taba are presented in (21). Example (21a) indicates 'my foot', with the possessor expressed by the independent pronoun *yak* '1SG', followed by a possessive ligature *nik* '1SG.POSS', then the possessee *wwe* 'foot'. Example (21b) indi- cates 'Mado's child', with the possessor expressed by a noun *Mado* '(personal name)', followed by a possessive ligature *ni* '3SG.POSS', then the possessee *mtu* 'child'. A possessive ligature may occur without a noun phrase overtly express- ing the possessor. Examples in (22) are the same as those in (21), but with- out the noun phrases *yak* '1SG' (a) and *Mado* '(personal name)' expressing the possessor.

(21) Taba possessive expressions (Bowden 2001: 173)
    a. *Yak   nik   wwe*
       1SG   1SG.POSS foot
       'my foot'

    b. *Mado ni   mtu.*
       Mado 3SG.POSS child
       'Mado's child'

(22) Taba possessive expressions without an overtly expressed possessor (Bowden 2001: 173)
    a. *nik   wwe*
       1SG.POSS foot
       'my foot'

    b. *ni   mtu*
       3SG.POSS child
       'his/her child'

As can be seen in Table 3.5, 1SG, 2SG, 3SG, and 1EX.PL cross-referencing forma- tives and possessive ligatures carry shared consonants *k, m, n,* and *t* respectively. Their etymological relationship becomes more obvious when compared with the reconstructed PMP forms given in Table 3.1. The forms listed in Table 3.5 are the same as the genitive pronouns in Table 3.1, however, both basic genitive forms and the genitive forms preceded by the genitive specifier *ni (see fn. 2)

TABLE 3.5 Reconstructed Proto-Malayo-Polynesian genitive clitic pronouns (cf. Table 3.1)

PMP	1MIN	2MIN	3MIN	1+2MIN	1+2AUG	1AUG	2AUG	3AUG
Genitive	*=ku	*=mu	*=ya	*=ta	*=tamu ~ *=tamuyu ~ *=tayu	*=mi	*=muyu ~ *=yu	(*ida)
*ni + Genitive	*ni=ku	*ni=mu	*ni=ya	*ni=ta	*ni=tamu ~ *ni=tamuyu ~ *=tayu	*ni=mi	*ni=muyu ~ *ni=yu	*ni ida
Taba	1SG	2SG	3SG	1IN.PL	-	1PL	2PL	3PL

are listed. The order of person and number has been modified to match that in Table 3.4 for the sake of easier comparison. Note that the PMP (bare) genitive forms are the ones that are considered to have been cliticised to the verb expressing A. It is commonly known that the reflexes of the earlier genitive pronouns indicating S and/or A appear in both/either enclitics and/or proclitic in Austronesian languages (Himmelmann 1996, Kikusawa 2003c). Based on the above information, it is not unreasonable to assume that the cross-referencing position is the X-position in Taba.

The current cross-referencing system in Taba as described above might be described as showing an accusative type, in the sense that nominative covers S and A, with a split in the marking of S, as has been mentioned above. The cross-referencing marker is obligatory for the A and S of "actor intransitive verbs," while optional for the S of "undergoer intransitive verbs" (Bowden 2001: 147–148, 223). On the assumption that cross-referencing in Taba developed from earlier genitive pronouns which marked A, it may be inferred that genitive pronouns have changed from marking both a syntactic case (genitive) and a semantic role (A) to one in which syntactic case marking has been lost and only the semantic role, actor, remains and was extended to cover what originally was marked by another syntactic case (S_{actor}+A). This process is schematically presented in Figure 3.3.

In Taba, we also note that the cross-referencing form is optional in undergoer intransitive clauses. This implies a further stage, where the function of the cross-referencing form is extending from a semantic role to a syntactic one, namely S, in that it includes both actor and undergoer marking.
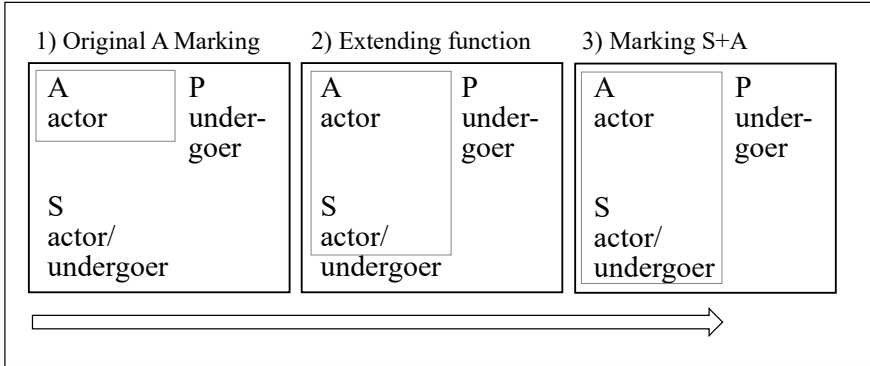
FIGURE 3.3   Possible paths of functional change from A to S+A

3.2.4   Fijian

Fijian languages are spoken in the Republic of Fiji Islands in the South Pacific, and their pronominal systems show a clear accusative pattern. The languages all have a set of "subject pronouns," or cross-referencing forms expressing the person and number of the actor (S+A). In addition, many Fijian languages also have another set of pronouns expressing the P of transitive clauses. Fijian clause structures with pronominal arguments are shown in (23). The components in parentheses are optional and can co-occur with a subject clitic pronoun for emphatic effect. Examples illustrating these structures appear in (24).

(23) Nadrau Fijian clause structures with pronominal arguments
      i. $[_{CR_1}=]_S$ Vi $([o \text{ IND}]_S)$
      iii. $[_{CR_1}=]_A$ Vt$[=_{CR_2}]_P$ $([o \text{ IND}]_A)$

(24) Nadrau Fijian
      a. Intransitive
         $[Au=]_S$ sā   mata   moce.
         1SG   ASP   want   sleep
         'I want to sleep. / I feel sleepy now.' (Kikusawa 2001: 55)

      b. Transitive
         $[Au=]_A$ sā   zivi$[=xexo]_P$ qaca $([oyau]_A)$.
         1SG   ASP   see=2SG   finish   (I)
         'I have already seen you.' (Kikusawa 2001: 86)

A possible trace of the earlier genitive forms in these languages is the 1SG subject pronoun *qu=* or *=qu* [ŋgu] that widely occurs in Western Fijian languages (Pawley & Sayaba 1971), and this is the form that is identical with the 1SG genitive form in possessive noun phrases found throughout the Fijian languages.

Kikusawa (2002) considers this to be a remnant of an earlier genitive pronoun. An example with the form *qu=* is presented in (25).

(25) Nabukelevu Fijian

qu=	*laka*	*a*	*niavi*
I.PAST	go	3SG.PAST	yesterday

'I went yesterday.' (Pawley & Sayaba 1971: 419)

In some of the western Fijian languages, such as Lomawai and Malomalo in Nadrogā, the forms for the 2SG and 3SG subject pronouns are also either fully or nearly identical with the corresponding genitive forms. Malomalo pronouns, where singular subject pronouns are identical to genitive pronouns, are presented in Table 3.6.

TABLE 3.6   Malomalo Fijian pronominal forms (P. Geraghty, unpublished fieldnotes and pers. comm.)

	1SG	2SG	3SG	1IN.DL	1EX.DL	2DL	3DL	1IN.PL	1EX.PL	2PL	3PL
INDEPENDENT FORMS	*yau*	*iko*	*kua*	*ketaru*	*kemaru*	*kemuru*	*kuru*	*ketatou*	*kematou*	*kemutou*	*kora*
SUBJECT PRONOUNS (default)	*qu*	*mu, o*	*a*	*taru*	*maru*	*muru*	*aru*	*tu*	*matu*	*mutu*	*ara*
SUBJECT PRONOUNS (present-future, < default + *i*)	*qɩ*	*ɩ*	*e*	*tarɩ*	*marɩ*	*mɩrɩ*	*erɩ*	*ʋ*	*maɟɩ*	*mɟɩ*	*era*
POSSESSIVE SUFFIXES I	*-qu*	*-mu*	*-(y)a*	*-taru*	*-maru*	*-muru*	*-dru*	*tatou*	*-matou*	*-mutou*	*-dra*
POSSESSIVE SUFFIXES II	*-qu*	*-mu*	*-(y)a*	*-taru*	*-maru*	*-muru*	*-dru*	*-tu*	*-matu*	*-mutu*	*-dra*
PREPOSED POSSESSIVE FORMS (parts of wholes, including body parts) (ex. 4 below)	*qu-*	*mu-*	*e-*	*taru-*	*maru-*	*muru-*	*eru-*	*tu-*	*matu-*	*mutu-*	*era-*

Notes: Possessive suffixes I are used when the suffixed form occurs as a noun, as in exs. 1 and 2 below, while Possessive suffixes II are used when the suffixed form modifies a noun, as in ex. 3. An example of preposed possessive form is also given in ex. 4 below.

ex1. *na le-tatou* 'ours'   ex3. *le-tu were* 'our house'
ex2. *luve-tatou* 'our child'   ex4. *qu-lima* 'my hand' (body parts, names, parts of wholes such as plants, fish, etc.)

TABLE 3.7   Some singular subject pronoun sets found in Fijian languages (based on
            P. Geraghty, unpublished 100 word lists)

	Languages	1SG	2SG	3SG
Set 1	Standard Fijian and some eastern Fijian languages	*au~u*	*o*	*e*
Set 2	Some eastern Fijian and Kadavu languages	*au*	*ko~ʔo*	*e~i*
Set 3	Some western Fijian languages	*qu~kau*	*ko~kō*	*ka~a*
Set 4	Lomawai, Malomalo (in Nadrogā)	*qu*	*mu~mū*	*a*

Table 3.7 is intended to show that there is a wide variety of pronominal forms found in Fijian languages with historical implications, and this is by no means an exhaustive list. Alternating forms indicate variants, and they do not necessarily occur in a single language.

TABLE 3.8   Reconstructed Proto-Oceanic clitic pronouns (Lynch, Ross, & Crowley 2002: 68)

	1SG	2SG	3SG	
Set I	*au=*	*ko=*	*i=*	< nominative
Set II	*ku=*	*=mu*	*=(y)a, ña=*	< genitive
Set III	*[y]a=*	*o-*	*e-*	
Object enclitics	*=au*	*=ko*	*=a*	< nominative

Subject pronoun forms in some Fijian languages are shown in Table 3.7 and the forms of reconstructed Proto-Oceanic singular clitic pronouns are presented in Table 3.8. By comparing these two, it can be seen that Fijian Set 1 and 2 are reminiscent of the Proto-Oceanic Set I, which Lynch, Ross & Crowley (2002: 83) claim may have developed from the earlier nominative set. Clearly, Fijian Set 4 is an obvious reflex set of the Proto-Oceanic Set II, which Lynch, Ross & Crowley claim may have developed from the earlier genitive set. More discussion regarding these forms appears in 3.3. Fijian Set 3 is presented as an example set consisting of forms with a mixed origin.

Tables 3.7 and 3.8 show that Fijian subject pronouns show a clear accusative pattern, with "subject pronouns" indicating the person and number of S and A, and another set of pronouns expressing P. Second, the subject pronoun sets in Fijian languages have various patterns in terms of their origin. Although limited to the singular forms, in some languages, the subject pronoun set reflects an earlier genitive set, while in some, it reflects an earlier nominative set. In some languages, the forms are mixed and appear to show transition. Thus, in Fijian, the subject pronoun position should be treated as a remnant of the X-position.

3.2.5 Tongan

Tongan is spoken in the Kingdom of Tonga in the South Pacific and belongs to the Polynesian language group. Tongan personal pronouns occur in two different patterns, i) a common set of clitic pronouns marking both S and A, with an independent pronoun marking P of a transitive clause, and thus occurring in an accusative case-alignment pattern; ii) independent pronouns occurring in the same ergative pattern as lexical NPs. The accusative clitic pronoun system is commonly shared with other Oceanic languages, such as the one described as "subject pronouns" in Fijian languages, while the ergatively marked independent pronoun system is unique to the Polynesian language group. According to Otsuka (2017), the use of independent pronouns in lexical NP slots is, like in Ibaloy, "marked and has an effect of emphasis." The Tongan pronominal systems described above alternate with their corresponding NPs (pronominal or non-pronominal), and thus, the Tongan system is different from that of Fijian where an NP or an independent pronoun may co-occur with a subject clitic pronoun for the purpose of emphasis. Tongan clause structures with clitic pronouns and independent pronouns are summarised in (26) and (27).

(26) Tongan pronominal system (clitic and independent pronoun for E)
     i.   V$_{AUX}$ [CLTC]$_S$ Vi
     ii.   V$_{AUX}$ [CLTC]$_S$ Vid   [*ki-ate* IND]$_E$
     iii.   V$_{AUX}$ [CLTC]$_A$ Vt   [IND]$_P$

(27) Tongan pronominal system (independent pronouns)
     i.   V$_{AUX}$ Vi   [*ʔa* IND]$_S$
     ii.   V$_{AUX}$ Vid   [*ʔa* IND]$_S$ [*ki-ate* IND]$_E$
     iii.   V$_{AUX}$ Vt   [*ʔe* IND]$_A$ [*ʔa* IND]$_P$

These patterns are exemplified in (28–30).

(28) Tongan examples of structures (i) and (ii) with clitic pronouns (Otsuka
      2017: 993)
      a. *Naʔa   **ku**   kata.*
         PAST   1SG   laugh
         'I laughed.'

      b. *Naʔa   **ku**   ʔaʔahi ki he   fanga tamaiki kotoa ʔi falemahaki.*
         PAST   1SG   visit   to DEF PL   children all   in hospital
         'I visited the children in the hospital.' (Chung 1978: 192)

(29) Tongan examples of structure (ii) with clitic pronouns (Otsuka 2017: 993)
      c. *Naʔa   **ku**   maʔu   ʔa   e   ika.*
         PAST   1SG   get   ABS   SPEC   fish
         'I caught a fish.'

      d. *Naʔe   taaʔi   **au**   ʔe   Sione.*
         PAST   hit   1SG   ERG   John
         'John hit me.'

(30) Tongan examples of structures (i–iii) with independent pronouns
      (Otsuka 2017: 992)
      a. *Naʔe   ʔomai   ʔe   Sione ʔa   e   tohi   **ki-ate**   **au**.*
         PAST   give.me   ERG   John   ABS SPEC   book   to-PRON   1SG
         'John gave a book to me.'

      b. *Naʔe   taaʔi   ʔe   Sione   **ʔa**   koe.*
         PAST   hit   ERG   John   ABS   2SG
         'John hit you.'

      c. *Naʔe   tala   mai   **ʔe**   **ia**   ʔoku   tonu.*
         PAST   tell   DIR   ERG   3SG   PRS   correct
         'He told me (that) it was correct.'

Clitic pronouns in Polynesian languages are typically treated as a separate set
from other pronominal sets. However, when the forms of clitic pronouns and
the genitive forms occurring on nouns are compared, a significant overlap be-
tween them is noted as shown in Table 3.9. It appears that there is some kind of

TABLE 3.9   Tongan personal pronouns

	CLITIC	GENITIVE[a]	INDEPENDENT
1SG	*ou, u*		
	ku	*ku*	
			au
2SG	*ke*		
		o, u	
			koe
3SG	*ne*	*ne*	
		no, na	
			ia
1DL.IN	*ta*	*ta*	
		taua	*(ki)taua*
1DL.EX	*ma*	*ma*	
		maua	*(ki)maua*
2DL	*mo*	*mo*	
		moua	*(ki)moua*
3DL	*na*	*na*	
		naua	*(ki)naua*
1PL.IN	*tau*	*tau*	
		tautolu	*(ki)tautolu*
1PL.EX	*mau*	*mau*	
		mautolu	*(ki)mautolu*
2PL	*mou*	*mou*	
		moutolu	*(ki)moutolu*
3PL	*nau*	*nau*	
		nautolu	*(ki)nautolu*

a   The longer possessive forms are used for emphasis.

historical connection between the clitic pronoun set and the genitive set. Thus, it is worth examining whether earlier genitive pronouns are a possible source for Tongan clitic pronouns.

3.3   *A Summary of Pronoun Position Comparison*
In 3.2, abstracted clause structures and the X-position (a reflex of the earlier genitive position) have been described for five languages. The findings are

summarised in Tables 3.10 and 3.11. In Table 3.10, the formal characteristics of the reflex set of the earlier genitive and its functions are listed. In Table 3.11, pronominal sets that are used to express S, A and P are summarised and the columns that indicate a pronominal set which is related to the earlier genitive in some way is filled with grey. It can be seen in Table 3.11 that the genitive set is exclusively related to the marking of A in Ibaloy and Pendau. In the other three languages, it is evident that there is some continuation of the earlier genitive set, marked with grey, occurring in each language although the function and syntactic distribution is completely different. It is obvious from the two tables that the earlier genitive set merged with the earlier nominative set and that the new pronominal set covers both S and A, which were earlier expressed by nominative and genitive respectively. The X-position is the position where forms in such sets occur in actual clauses, and it is one of the components that can be used to identify cognacy and to clarify the developmental paths of clause structures.

TABLE 3.10   Possible remnants of the earlier genitive pronoun marking ergative

Language	Formal characteristics of the reflex set	Function
Ibaloy	clitic	A
Pendau	clitic/independent	A of Viv
Taba	part of cross-referencing form	S (actor) + A
Fijian	part of clitic (~cross-referencing) set	S + A
Tongan	part of clitic	S + A

TABLE 3.11   A comparison of pronominal sets marking S, A and P

	Ibaloy	Pendau	Taba	Fijian	Tongan	
S (undergoer)	=NOM	ABS		NOM=	=CLTC	ʻa IND
S (actor)	=NOM	ABS	CR=	NOM=	=CLTC	ʻa IND
A	=GEN	=GEN	CR=	NOM=	=CLTC	ʻe IND
P	=NOM or IND	ABS	IND	=ACC (~IND)	IND	ʻa IND

3.4 *Discovery of the Merger of Pronominal Sets in Oceanic Languages*

In Austronesian languages, the pronominal systems in Taiwan and the Philippines show a morphologically marked ergative system like the one in Ibaloy, while those in many Oceanic languages show an accusative pattern, as has been presented in Section 3.3. Kikusawa (2002, 2003b, 2015), based on a comparison of the forms of the pronouns composing relevant sets, proposes a hypothesis that the direction of the change must have been from ergative to accusative. For supporting evidence, as well as references to proposals that the shift was from accusative to ergative rather than the reverse, see Kikusawa 2002, 2003b, and 2017. The directionality of the change is identified with reference to change in pronominal patterns, namely merger, which is known to be strongly unidirectional.

As an example of the change in the order that is proposed here I compare the argument structures with an Auxiliary verb in Ibaloy (repeated in (31)), and the clitic pronoun system in Tongan (repeated in (32)).

(31) Ibaloy argument structures with pronominal NPs (with Aux) (= (10))
  i. $V_{AUX}[=NOM]_S$ Vi
  ii. $V_{AUX}[=NOM]_S$ Vid $[son\ IND]_E$
  iii. $V_{AUX}[=GEN]_A$ Vt $[NOM/IND]_P$

(32) Tongan pronominal system (clitic pronouns) (= (26))
  i. $V_{AUX}\ [CLTC]_S$ Vi
  ii. $V_{AUX}\ [CLTC]_S$ Vid $[ki\text{-}ate\ IND]_E$
  iii. $V_{AUX}\ [CLTC]_A$ Vt $[IND]_P$

By comparing these two systems it can be seen that one of the major differences between them is the set of pronouns that express S or A following the Auxiliary verb. In Ibaloy, the nominative clitic pronoun expresses S and the genitive clitic pronoun expresses A (thus showing an ergative pattern), while in Tongan, a single set of pronouns ("clitic pronouns") express both S and A, or the actor. It has been argued that the latter developed from the former as a result of the merger of the two clitic pronoun sets that occurred in post-Auxiliary positions in the earlier system. The claimed direction and mechanism of the change is indicated in (33), and the assumed precondition is that the position of pronominal forms expressing S and A was fixed as the post-Auxiliary (pre-main verb) position. Once this happened, the forms occurring in the post-Auxiliary position (shown in (33) in a box with dotted lines) must have been recognised as belonging to a single set covering both S and A (those in a box with straight lines). It is hypothesised that this eventually resulted in

the merger of what originally were genitive and nominative sets, resulting in a system that is analysed as having changed to an accusative pattern.

(33) Corresponding arguments and claimed direction of the development
     i.   VAUX [=NOM]$_S$ Vi
     ii.   VAUX [=NOM]$_S$ Vi   [*son* IND]$_E$
     iii$_2$. VAUX [=GEN]$_A$ Vt   [NOM/IND]$_P$

     i.   VAUX [CLTC]$_S$ Vi
     ii.   VAUX [CLTC]$_S$ Vi   [*ki-ate* IND]$_E$
     iii. VAUX [CLTC]$_A$ Vt   [IND]$_P$

The proposed change is supported by the fact that the forms that occur in the boxed positions in the new system are etymologically a mixture. As has been mentioned earlier, Lynch, Ross & Crowley (2002: 83) reconstruct "subject clitic pronouns" and "object clitic pronouns" for Proto-Oceanic, however, they are aware that their reconstructions are not problem-free. They state that "[a]lthough subject proclitics (or prefixes) occur in many well distributed Oceanic languages and we can infer their presence in P[roto-]Oc[eanic], their forms vary considerably and a number of competing reconstructions can be made" (*ibid*, 68). They therefore reorganise the reconstructed forms as in Table 3.12 and point out that "Sets I and II respectively reflect the P[roto-]M[alayo-]P[olynesian] nominative and genitive clitics." Here, I provide the data presented in Table 3.13 to show how diverse the forms are that are found in the reflexes in modern languages, and how their etymology can be identified by assuming multiple sources. A simple examination of the forms of the clitic pronouns reveals that that they have actually come from at least two different sources, namely, earlier nominative and genitive. However, Lynch, Ross & Crowley did not have any explanation as to why this situation was brought about historically.

TABLE 3.12 Reconstructed Proto-Oceanic clitic pronouns (Lynch, Ross, & Crowley 2002: 68) (=Table 3.8)

	1SG	2SG	3SG	
Set I	*au*=	*ko*=	*i*=	< nominative
Set II	*ku*=	=*mu*	=(*y*)*a*, *ña*=	< genitive
Set III	[*y*]*a*=	*o*-	*e*-	
Object enclitics	=*au*	=*ko*	=*a*	< nominative

TABLE 3.13 1SG Clitic pronoun forms found in Central-Pacific languages organised according to their etymology

	< 1SG.GEN	< 1SG.NOM	< 1SG.IND	Source
Rotuman	ŋou			(Churchward 1998)
W. Fijian	ŋgu	–	–	(Pawley & Sayaba 1971)
E. Fijian	–	au	–	(Pawley & Sayaba 1971)
Tongan	ku	ou~u		(Pawley 1970)
Samoan	ʔou	ou	oʔu	(Pawley 1970)
Tokelauan	kō			(Huntsman, Hooper, & Simona 1986)
East Futunan	kau			(Pawley 1970)
East Uvean		u~au		(Pawley 1970)
Tuvaluan	kau		aku	(Besnier 2000)
West Futuna-Aniwan	ŋk~nk~ŋ~n~h		ah	(Dougherty 1983)

With the proposed hypothesis, the existence of a variety of forms in the "clitic pronoun" set expressing the actor in modern Oceanic languages can be readily explained by assuming a merger of the two earlier pronominal sets. The reverse direction is not possible, for it would have to be assumed that every language where the A is marked by a form that has its origin in a genitive pronoun independently re-aligned the system by selecting the same form out of a mix of earlier genitive and nominative forms. In such a case like the one presented in this section, directionality is strongly supported by the merger itself. Merger is a change known for its unidirectionality. A reversal change requires an item-by-item relearning, which could not take place in separate languages for the same system to be produced as a result.

4 Cognacy, Structural Changes and Directionality

In this section, the X-positions identified in Section 3 are used to determine cognate structures. Once cognate structures are identified, the principles of the Comparative Method are applied, then the scenario of the historical development among the cognate structures is clarified. This process also involves an examination of the diachronic relationship between those with the X-position and those without. The resulting hypothesis should be an "optimal theory of

the differentiation of" the daughter systems developing from a proto-system, that is, that the proto-system must be plausible as a human language and the assumed changes must be natural (Clark 1976: 24–28). To postulate such a hypothesis, distributional evidence is first scrutinised, and then the results are examined as to their plausibility.

There are two identifiable cases of change that can be determined by applying the methods that are presented in this section. The first case is a word order change of pronominal NPs, the directionality of which is ascertained based on the possible motivation of the change (4.1). The second case is the development of applicative systems (4.2). In the latter, the directionality of the change is first hypothesised based on the distributional pattern of relevant grammatical morphemes. Results of the examination show a transitional stage between the old system and the system where the innovative applicative system exists, and thus support the proposed working hypothesis. These hypotheses are summaries of discussion from previous publications (Kikusawa 2002, 2003, 2015).

4.1   *Word Order Change of Pronominal NPs*

Ibaloy and Pendau are relatively closely related, nevertheless, the former shows a clear ergative system while the latter shows an inverse system. In this section, I argue that one of the major changes by which the two systems developed is a word order change of pronominal NPs, where the earlier nominative NP that occurred in post-verbal position acquired a preverbal clause initial position. The developmental paths and the process of the change are identified in this section based on a comparison of cognate structures.

The focus of this chapter is strictly on the method for the applying the Comparative Method to syntactic features. That the verb-initial word order is more conservative and the nominative NP came to occupy preverbal position has been discussed in Kikusawa (2003) and Aldridge (2010). The discussion in the latter is based on theoretical underpinnings, rather than comparative, and the fact that two researchers with different perspectives separately come up with the same conclusion supports and strengthens the two hypotheses.

The schematic structures of Ibaloy (without Auxiliary, (6)) and Pendau (15) are repeated in (34), this time with pairs of cognate structures indicated with boxes. Among these, X-positions occur in dyadic structures in iii., and these are identified as cognate clause structures. Structures in i. are the sole single argument structures in each language and are assumed to have developed from the intransitive structure in the shared proto-language. Structures in ii. are dyadic intransitive structures, and based on negative evidence (the lack of the X-position), they are also provisionally treated as possible cognate structures.

(34) Ibaloy and Pendau argument structures with pronominal NPs and their cognacy

Ibaloy (without Aux)   Pendau

i.   Vi[=NOM]$_S$ i. [ABS]$_S$ Vi

ii.   Vi[=NOM]$_S$ [*son* IND]$_E$ ii. [ABS]$_{Aav}$ Vav   [ABS]$_{Pav}$

iii. Vt[=GEN]$_A$[=NOM]$_P$ iii.[ABS]$_{Piv}$ Viv[=GEN]$_{Aiv}$

Based on a comparison of the occurrence distribution and function of NPs in the structures in iii., the nominative NP in Ibaloy can be analysed as corresponding to the absolute NP in Pendau. The genitive NP in Ibaloy by default corresponds to the genitive NP in Pendau, both being the X-position. However, the positions of supposedly corresponding nominative NPs in the two languages do not match. For example, in Ibaloy, both genitive and nominative pronouns expressing A and P are encliticised to the verb, while in Pendau, the absolute pronoun expressing P precedes the verb, while the genitive pronoun expressing A is encliticised to the verb just as in Ibaloy. Since it is known that the two languages developed from a single system, namely PMP, it can be hypothesised that the word order changed in either or both of the languages.

To identify their developmental paths, first I examine the word order restrictions of each of the two languages. Starting with Ibaloy, as mentioned in 3.2.1, the pronoun expressing the nominative NP of the clause may be either a clitic or an independent pronoun. The relative word order of the pronouns to the main verb is commonly shared by both languages in that they follow the verb in the order of genitive, then nominative. Clause structures with independent pronouns are repeated in (35).

(35) Ibaloy argument structures with independent pronouns (without Aux, =(12))

i.   Vi   [NOM/IND]$_S$

ii.   Vi   [NOM/IND]$_S$ [*son* IND]$_E$

iii. Vt[=GEN]$_A$ [NOM/IND]$_P$

In addition, in Ibaloy, either the nominative or genitive component may be topicalised. In such a case, an independent NP may appear in the clause initial position (Ruffolo 2004: 469). The NP appearing in the clause initial position may be a pronoun or a lexical NP, and whichever the case is, it co-occurs with the corresponding clitic nominative pronoun. Example (36) is an example where 1SG independent pronoun *siʔgak* is topicalised and occurs before the verb indicating "as for me" and co-occurs with the 1SG genitive clitic

pronoun =*ko*. The clause structure with topicalised independent pronouns are shown in (37). It is assumed here that the topicalised NP could also be an independent pronoun when it corresponds to the nominative component.

(36) Ibaloy example with a topicalised pronominal NP
nem   si'kak,   kowankoy   aychi!
nəm   **siʔgak**   kowan[=ko]$_A$[=j   ʔajdi]$_P$
but   1.IND   say=1.GEN=NOM   no
'but as for me, I said no!' (Ruffolo 2004: 474)

(37) Ibaloy argument structures with topicalised pronominal NPs
    i.   PRON.IND   Vi[=NOM]$_S$
    ii.   PRON.IND   Vi[=NOM]$_S$ [*son* IND]$_E$
    iii. PRON.IND   Vt[=GEN]$_A$[=NOM]$_P$

Thus, Ibaloy exhibits three structures for the transitive clause, including the one with a topicalised NP, as shown in (38).

(38) Variations of Ibaloy structure iii.
    a. Vt[=GEN]$_A$[=NOM]$_P$
    b. Vt[=GEN]$_A$ [NOM.IND]$_P$
    c. PRON.IND   Vt[=GEN]$_A$[=NOM]$_P$

In Pendau, on the other hand, Quick (2007: 123), in his description of basic clause structures, introduces the notion "pivot" to refer to the component that occurs in clause initial position, the function of which overlaps with those that are typically associated with subjecthood, such as quantifier float, relativisation and equi-subject deletion (2007: 127–132). However, according to Quick (2007: 365–366), the position of the absolute NPs is not completely fixed, and the component occurring in the pivot position may occur following all the other constituents of the clause. Such structures are repeated in (39).

(39) Pendau argument structures with pronominal NPs (2) occurring in the phrase final position (=(18))
    i.   Vi   [ABS]$_S$
    ii.   Vav [ABS]$_P$ [ABS]$_A$
    iii. Viv[(=)GEN]$_A$ [ABS]$_P$

(40) Variations of Pendau structure iii.
    a. [ABS]$_{Piv}$ Viv[=GEN]$_{Aiv}$
    b.   Viv[(=)GEN]$_A$ [ABS]$_P$

To compare and reconstruct the Ibaloy and Pendau systems, I first focus on the variations of structure iii. in the two languages, namely those listed in (38) and (40), which we know are cognate structures. Then, the results are examined as to whether they are consistent with the structures in i and ii. In doing so, principles of the Comparative Method summarised by Clark (1976: 24–27) are applied. That is, in reconstruction, positing a hypothesis with fewer and more natural changes based on distributional evidence and plausibility is considered to be more appropriate than otherwise.

Between Ibaloy and Pendau structures, there is one structure that appears to be commonly shared between the two languages, namely (38-ii) and (40-ii). These clause structures are analysed differently by the analyst of each language (transitive for Ibaloy, and inverse for Pendau), however, they are both i) a two argument structure, ii) verb initial, iii) with a genitive clitic expressing the grammatical function A and, iv) with an independent pronoun expressing P following the A. Assuming that the "majority rule" is valid for the reconstruction of syntactic structures, this two-argument verb structure is reconstructed as a proto-structure as in (41). I refer to it as Proto-Ibaloy-Pendau (PIP). The rest of the structures are listed in (42). The question here is how many of the structures in (42) can be explained by the hypothesis in (41).

(41) Possible Proto-Ibaloy-Pendau two argument structure
      $V[=\text{GEN}]_A \;\; [\text{IND}]_P$

(42) Structures that are different from the reconstructed one
      i.   Ibaloy $Vt[=\text{GEN}]_A[=\text{NOM}]_P$ $=(34c)$
      ii.   Ibaloy   PRON.IND $Vt[=\text{GEN}]_A[=\text{NOM}]_P$ $=(37c)$
      iii. Pendau $[\text{ABS}]_{Piv}$ $Viv[(=)\text{GEN}]_{Aiv}$ $=(34c)$

The difference between the reconstructed structure in (41) and (42-i) is whether the NP expressing the P is a clitic or not. It is commonly known that pronouns can be grammaticalised to become clitics and eventually agreement markers (e.g., De Vogelaer 2008: 223–225). It appears that there is a good possibility that Ibaloy, along with other Philippine languages today, shows a stage in the change where a nominative pronoun is becoming a clitic and this claim contradicts current reconstructions of nominative pronouns as clitics. According to Reid & Liao (2001: 21–24), the status of the pronoun expressing P varies among Philippine languages, being either a clitic or an independent pronoun or both, and this appears to support the claim that the clitic status was acquired after the dispersal of these languages. PMP Nominative pronouns have been reconstructed as clitics as shown in Table 3.1. It is necessary to examine both the form and distribution of these pronouns in western Austronesian languages,

combining morphological reconstruction with clause structure reconstruction to clarify the actual developmental path.

Structure (42-ii) is a derived structure of (42-i), with an independent NP expressing the topicalised element of the clause. Topicalisation of an argument by fronting an NP is widely found in Austronesian languages, although the conditions vary. Further discussion of this point appears in Section 4.1.

The difference between the reconstructed structure and (42-iii) is the position of the NP expressing P. It should be noted here that the two sets of pronouns in Pendau, which Quick refers to as absolute and genitive, are formally non-bound and bound pronouns respectively. The former carries functions carried by independent pronouns in other languages, occurring in both argument positions of equative clauses and copula clauses, the object of prepositional phrases, and other argument positions except for the argument expressing A of an inverse structure (2007: 126). When we compare (41) and (42-iii) bearing this in mind, the major difference between the two is the position of the "independent" NP expressing P. Here I show that by assuming a topicalised construction in the proto-system, we can explain not only the development of (42-iii) but also that of the other clauses. By topicalised construction is meant here a structure where an independent NP expressing the nominative element is fronted. In (43–44), these assumed proto-structures are presented. An independent NP expressing S or P in the clause appears following the verb in (43), while in (44), it occurs preceding the verb.

(43) Proto-Ibaloy-Pendau basic argument structures

i.	Vi	$[\text{NOM.IND}]_S$	
ii.	Vi	$[\text{NOM.IND}]_S$	$[son\ \text{IND}]_E$
iii.	Vt$[=\text{GEN}]_A$	$[\text{NOM.IND}]_P$	

(44) Proto-Ibaloy-Pendau topicalised constructions

i.	$[\text{NOM.IND}]_S$	Vi	
ii.	$[\text{NOM.IND}]_S$	Vi	$[son\ \text{IND}]_E$
iii.	$[\text{NOM.IND}]_P$	Vt$[=\text{GEN}]_A$	

The structures in (43) are shared by both languages. The current Ibaloy system can be explained as resulting from nominative pronouns developing into clitics from this system.[9] The Pendau system on the other hand appears to have

---

9   In structure ii, in Ibaloy, it is the nominative S that is phonetically attached to the verb, while in Pendau, it is the P (equivalent of E in Ibaloy) that is closely attached to the verb (Quick 2007: 366). This appears to support the idea as well that the clitic status of nominative in

developed as a result of the structures in (44) becoming basic (unmarked) structures. This hypothesis would be naturally motivated if the PMP (thus PIP) system was not only morphologically ergative but also syntactically ergative. It has been reported that in some Philippine-type languages spoken today, nominative NPs are the only core NPs that can be fronted for topicalising (or, "extracted", cf. Payne 1982, Aldridge 2004).[10] Thus, the word order change between the two systems is readily explained by the allowed fronting of nominative arguments in system (i) becoming a fixed position in system (ii) (Kikusawa 2003, To appear).

The assumed sequence of the development of basic clause structures in Ibaloy and Pendau is presented in (45–47). In each set, the first two lines show the reconstructed PIP structures, followed by the Ibaloy and Pendau reflex clauses.

(45) A sequence of development of the intransitive clause (< *Intransitive)

PIP BASIC		Vi	$[\text{NOM}]_S$ =(43a)
PIP TOPICAL.	$[\text{NOM.IND}]_S$	Vi	=(44a)

> Ibaloy BASIC		Vi$[=\text{NOM}]_S$	=(35a)
> Ibaloy TOPICAL. PRON.IND		Vi$[=\text{NOM}]_S$	=(37a)

> Pendau BASIC		Vav	$[\text{ABS}]_S$ =(39a)
> Pendau BASIC	$[\text{ABS}]_S$	Vav	=(15a)

(46) A sequence of development of the first dyadic clause (< *Dyadic Intransitive)

PIP BASIC		Vi $[\text{NOM.IND}]_S$	$[son\ \text{IND}]_E$	=(43b)
PIP TOPICAL.	$[\text{NOM.IND}]_S$	Vi	$[son\ \text{IND}]_E$	=(44b)

> Ibaloy BASIC		Vi $[\text{NOM.IND}]_S$	$[son\ \text{IND}]_E$	=(35b)
> Ibaloy TOPICAL. PRON.IND		Vi$[=\text{NOM}]_S$	$[son\ \text{IND}]_E$	=(37b)

> Pendau BASIC		Vav $[\text{ABS}]_P$	$[\text{ABS}]_A$	=(39b)
> Pendau BASIC	$[\text{ABS}]_A$	Vav	$[\text{ABS}]_P$	=(15b)

Ibaloy and also the syntactic attachment of P in Pendau are both the result of innovations after the languages split. These facts imply that the free position of the absolute A NP in Pendau is an innovation subsequent to the change where P became syntactically closely attached to the verb.

10   In Ibaloy, however, the genitive NP (expressing A) can also be extracted. How the Ibaloy case relates to this historical development requires further examination.

(47) A sequence of development of the second dyadic clause (< *Transitive)

PIP BASIC		Vt$[=$GEN$]_A$	$[$NOM.IND$]_P$	$=$(43c)
PIP TOPICAL	$[$NOM.IND$]_P$	Vt$[=$GEN$]_A$		$=$(44c)
> Ibaloy BASIC		Vt$[=$GEN$]_A$	$[$NOM.IND$]_P$	$=$(35c)
> Ibaloy TOPICAL. PRON.IND		Vt$[=$GEN$]_A$	$[=$NOM$]_P$	$=$(37c)
> Pendau BASIC		Viv$[(=)$GEN$]_A$	$[$ABS$]_P$	$=$(39c)
> Pendau BASIC	$[$ABS$]_P$	Viv$[(=)$GEN$]_A$		$=$(15c)

The cognacy of the clause structures under examination and the presented hypothesis as to how the Ibaloy and Pendau developed is supported by the occurrence pattern of a verbal affix (the so-called *maN-) (Kikusawa 2017, To appear).

The following is a summary of the characteristics of PIP:

i)   It was a verb initial language

ii)   The nominative pronoun expressing S and P was a free form and not a clitic

iii) In transitive clauses, the genitive pronoun expressing A was encliticised to the verb

Space does not allow detailed discussion, however, it should be mentioned here that parallel correspondences are found in clause structures with lexical NPs in Ibaloy and Pendau, and the same hypothesis can be applied to explain their differences and possible developmental paths. This supports the proposed hypothesis that the Pendau system developed by fronting an NP rather than by the clitic pronoun being stranded by "aux-axing" in preverbal position as proposed by Starosta, Pawley & Reid (1982).

Clause structures with lexical NPs in the two languages are shown in (48–49) and examples illustrating them are given in (50–51).

(48) Abstracted argument structures with lexical NPs in Ibaloy (an ergative system)

i.   Vi $[si/ʔi$ NP$]_S$

ii.   Vi $[si/ʔi$ NP$]_S$ $[son/ni_2$ NP$]_E$

iii. Vt $[nən/ni_1$ NP$]_A$ $[si/ʔi$ NP$]_P$

The forms *si* and *ʔi* in (48) mark the difference between personal and common nouns that express the nominative NP of the structure, likewise *nen* and *ni* mark the difference between personal and common nouns that express the genitive NP of the structure.

(49) Pendau argument structures with Lexical NPs (an inverse system)

    i.   Vi $[si/\varnothing\,\text{NP}]_S$

    ii. $[si/\varnothing\,\text{NP}]_{Aav}$ Vav $[si/\varnothing\,\text{NP}]_{Pav}$

    iii. $[si/\varnothing\,\text{NP}]_{Piv}$ Viv $[ni/nu\,\text{NP}]_{Aiv}$

The alternation between *si* and $\varnothing$ in (49) indicates the marking before proper and common NPs respectively.

(50) Examples in Ibaloy (I-ii) and Pendau (P-ii) (1)

    I-ii.   *engoney*   *i*   **aki**   *ni*   *otot*

               ʔəN-ʔonəj   [ʔi   ʔaki]$_S$ [ni$_2$ ʔotot]$_E$

               ACTV.PFT-see   NOM   monkey   OBL   mouse

               '**The monkey** saw a mouse'   (Ruffolo 2004: 238)

    P-ii.   [*Si*   *kai*]$_A$ *neng-ita-i*   [*si*   *be'e*]$_P$

               ABS.PNM   grandfather   AV.R-see-LOC   ABS.PNM   grandmother

               '**The grandfather** saw the grandmother.' (Quick 1994: 466)

(51) Examples of Ibaloy (I-iii) and Pendau (P-iii) (2)

    I-iii.   *naon'an*   *ni dedaki*   **sota**

               na-ʔonəj-an   [ni$_1$ RDP-laki]$_A$ [sota

               POTLOCV.PFT-see-LOCV   GEN PL-man   NOM.REC

               bibiid   ***Batan***

               RDP-biʔi=d   batan]$_P$

               PL-woman=LOC   Batan

               'The men happen[ed] to see **the women of Batan**' (Ruffolo 2004: 306)

    P-iii.   [*Si*   *be'e*]$_P$ *ni-ita-i*   [*ni*   *kai*]$_A$

               ABS.PNM   grandmother   IV.R-see-LOC   GEN.PNM   grandfather

               'The grandfather saw **the grandmother**.' (Quick 1994: 466)

Examples in Ibaloy with fronted nominative NPs (52) and a relativised nominative NP (53) are provided. There are structures in Ibaloy where the nominative NP precedes the verb. For example, a nominative NP may be fronted, or "clefted" in Ruffolo's terms (2004: 379), and it is only a nominative NP that may be fronted in such a construction. Examples are presented in (52). In (52b), the third person independent noun that occurs in the clause initial position corresponds to the nominative NP (actor) of the dyadic intransitive verb, and the

word *ŋanto* 'what' in the clause initial position corresponds to the nominative
NP (Patient) of a transitive verb in (52c).

(52) Ibaloy examples with fronted nominative NP
    a. ***si'katoy***   *dimaw*   *chi*   *Bagiw*
       si?gato=j   \<im\>law   di   bagiw
       3.IND=NOM   \<ACTV.PFT\>go   LOC   Bagiw
       'Who went to Baguio?' (Ruffolo 2004: 380)

    b. ***ngantoy***   *dingkato?*
       ŋanto=j   \<in\>laga=to
       what=NOM   \<PatV.PFT\>do=3.GEN
       'What did he do?' (Ruffolo 2004: 380)

A nominative NP may also precede the verb in relativised clauses, and in this
case also, it is only the nominative NP that may be relativised. An example of a
relativised transitive clause is presented in (53).

(53) Ibaloy transitive clause with relativised nominative NP
    a. *bara*   *kono*   *i*   ***titit*** *ya* *chakaichemang*
       wada   kono   ?i   titit   ja   daka=?i-dəmaŋ
       exist   hearsay NOM bird LK 3+.GEN.ASP=THMV.CNTV-see
      ya   *emeboteng*
       ja   ?əmə-botəŋ
       LK   STAPatV.CNTV-drunk
       'It is said that they keep seeing drunken birds' (Ruffolo 2004: 407) (LIT.
       There are, it is said, **birds** that they keep seeing [∅] and (that) are
       drunk[∅].)

Considering the fact that Ibaloy has a clear condition as to when a nominative
NP can precede or follow the verb, while in Pendau, there is no clear condition
specified by which the position of the corresponding argument, namely, the
absolute NP, is determined, it seems reasonable to assume that the direction of
the change was from the Ibaloy system to the Pendau one. At least the precon-
dition of the change appears to be clear; that the nominative NP acquired the
position preceding the verb as its default position. In contrast, if we assume
that the Pendau system was the earlier one, an explanation is needed as to why
the NPs preceding the verb came to occur in a post-verbal position in Ibaloy.

    Based on the above observation, the direction of the change is shown sche-
matically in (54). NPs that are nominative are bold. The change is applicable to
both pronominal and lexical NPs.

(54) Illustration of word order change from the Ibaloy to Pendau system

Structure	PMP		pre-Pendau		Pendau
i.	*V S	→	S V	→	S V
ii.	*V S E	→	S V E	→	A V P
iii.	*V A P	→	P V A	→	P V A

In what follows, reconstructed PMP clause structures and their reflex structures will be referred to as Structures i., ii., and iii., according to their cognacy.

4.2 *Development of Applicative Verb Suffixes*

In this section, the development of applicative verb suffixes in some Sunda-Sulawesi languages is examined.[11] The discussion starts with the recognition of a discrepancy between the correspondence of clause structures and the distribution of certain verb suffixes in Ibaloy and their apparently corresponding suffixes in Pendau. The suffixes in these two languages could both be referred to as "applicative suffixes," however, the ones in Ibaloy occur only in Structure iii., while supposedly corresponding suffixes in Pendau occur in both Structures ii. and iii. This fact appears to cast a question on the accuracy of the cognate identification carried out in Section 2 above.

In this section, I show that the proposed clause correspondence helps to clarify the differences between the two languages, and to identify the direction of change. It is shown that the Ibaloy suffixes can be identified as reflecting an earlier system, and the Pendau suffixes extended their distribution from Structure iii. to Structure ii. The directionality of the change is determined based on the fact that some "applicative suffixes" in Pendau show the same syntactic function as those in Ibaloy, occurring only in Structure c. The Pendau system, where two systems are combined, appears to show a transition period between the Ibaloy system and those found in other Sunda-Sulawesi languages, as discussed below. It is possible that the existence of the form *-i* with the wider distribution is a result of contact with Indonesian and other languages spoken in the area. For a detailed discussion, see Kikusawa, 2012 and To appear.

The distribution of relevant verb forms in Ibaloy and Pendau are compared in (55). Semantically, the transitive suffixes in Ibaloy, including *-i* (in complementary distribution with *-an*), and the applicative suffixes *-i* and *-aʔ* in Pendau appear to correspond to each other respectively. However, the Ibaloy affixes mark transitive constructions and occur only in Structure c. as shown in (56), while in Pendau, some affixes occur only in Structure iii. (Quick 2007:

---

11 The equivalent of the Ibaloy forms in closely related Austronesian languages has been referred to as "applicative" in some literature (Ross & Teng 2005, Daguman 2004, Aldridge 2004, Kaufman 2017).

304), as in the Ibaloy examples in (58–59), while the majority occur in both Structures ii. and iii. (Quick 2007: 288, see 60). For convenience, I will hereafter refer to the first type of affixes as applicative$_1$ and the second type of affixes as applicative$_2$. Applicative related phenomena in Pendau are discussed in detail in Quick (2007: 288–312).

(55) Distribution of "applicative" verb affixes in Ibaloy and Pendau and the semantic roles of applied argument

	IBALOY		PENDAU	
	TRANSITIVE AFFIX		APPLICATIVE$_1$	APPLICATIVE$_2$
Structure i.				
Structure ii.				
Structure iii.	-ən ~ -a	'patient'[a]	-aʔ$_1$ 'instrumental'	-aʔ$_2$ 'benefactive,
	-an ~ -i	'locative'[b]	-i$_1$ 'locative'	instrument'
	i-	'thematic'		-i$_2$ 'goal, locative'
	i- -an	'benefactive'		

a   The suffix -a occurs in continuative and progressive aspects (Ruffolo 2004: 254).
b   The suffix -i occurs in continuative, progressive and imperative aspects (Ruffolo 2004: 266) and also in various circumfixes in these aspects (Ruffolo 2004: 293, 297, 299, 304).

The examples in (56) illustrate Structures ii. and iii. in Ibaloy with the verb ʔonəj 'see'. Note that in ii., the verb has a reflex of PMP *maN- and the actor aki 'the monkey' is expressed as a nominative NP and the undergoer otot 'a mouse' as genitive. In structure iii., the verb has the locative affix -i and the actor is expressed with genitive forms, while the undergoer of the event is expressed with nominative forms.

(56) Ibaloy example illustrating Structures ii.
     engoney   i   aki   ni   otot
     ʔəN-ʔonəj   [ʔi   ʔaki]$_S$ [ni   ʔotot]$_E$
     ACTV.PFT-see   NOM   monkey   GEN   mouse
     'The monkey saw a mouse.' (Ruffolo 2004: 238)

(57) Ibaloy example illustrating Structures iii.
     on'im   kari   iman!
     ʔonəj-i[=m]$_A$ kadi   [ʔiman]$_P$
     see-LOCV/IMP=2.GEN   request NOM.DIST.PRON
     'Look at that one!' (Ruffolo 2004: 164)

The examples in (58–59) illustrate parallel examples to those presented above in Pendau. The verb *guntung* 'to light' in (58a) carries a reflex of PMP *maN- and the actor is expressed as an absolute pronoun. The form *-i* does not occur on the verb in this structure. In (58b) on the other hand, the verb carries the suffix *-i* and the actor appears in genitive (*=nyo* '3SG.GEN') and the undergoer *palan* 'light' appears in the absolute case preceding the verb. The pair shows a parallel system to the one shown in (56) for Ibaloy. Likewise, in (59a), the verb *mene?* 'to go up' carries *N-*, a reflex of PMP *maN-, the actor is expressed as an absolute pronoun *?a?u* '1SG' and the undergoer *niu* 'coconut' follows the verb. The verb does not carry the suffix *-i* in this structure. In (59b), on the other hand, the verb carries the suffix *-i*, and the actor is expressed by the genitive pronoun *=nyo* '3SG.GEN'. The undergoer *taipang* 'mango tree' in absolute case follows the verb in this example.

(58) Pendau examples with suffix *-i* occurring only in Structure iii. (1)

    a. *A'u   moguntung   palan.*
       ?a?u   m-pong-guntung   palan
       1SG.ABS   IR-SF.PT-light   light
       'I will light the lamp.' (Quick 2007: 305)

    b. *Palan   roguntuninyo.*
       palan   ro-guntung-**i**=nyo
       light   IV.IR-light-DIR=3SG.GEN
       'He/she will light the lamp.' (Quick 2007: 304)

(59) Pendau examples with suffix *-i* occurring only in Structure iii. (2)

    a. *A'u   nemene'   niu.*
       ?a?u   n-pe-mene?   niu
       1SG.ABS   RE-SF.DY-go.up   coconut
       'I climbed the coconut tree.' (Quick 2007: 331)

    b. *Nipene'inyo   taipang   uo.*
       ni-pene?-i=nyo   taipang   ?uo
       IV.RE-go.up-DIR=3SG.GEN   mango.tree yonder
       'He climbed up that mango tree.' (Quick 2007: 304)

It should be noted that the examples in (58–59) are not typically occurring patterns of applicative suffixes in Pendau. The majority of applicative suffixes *-i* and *-a?* in Pendau occur in both Structures ii. and iii., as in (60) below. In (60a), the verb *?omung* 'to carry' has a reflex of PMP *maN-* and also the suffix *-i*. The

actor is expressed in an absolute pronoun *ʔaʔu* and the undergoer *bau* 'fish' follows the verb. In (60b), the verb carries the suffix *-i* and the actor is expressed by a genitive pronoun *=ʔu*, and the undergoer *bau* 'fish' appears preceding the verb.

(60) Pendau examples with suffix *-i* occurring in both Structures ii. and iii.
  a. *A'u*   *mongkomuni*   *bau*   *riunjung.*
     ʔaʔu   m-pong-ʔomung-**i**   bau   ri=junjung
     1SG.ABS   IR-SF.PT-carry-DIR   fish   LOC=house
     'I will hold the fish at my house.' (Quick 2007: 303)

  b. *Bau*   *uo*   *ni'omuni'u*   *riunjung.*
     bau   ʔuo   ni-ʔomung-**i**=ʔu   ri=junjung
     fish   yonder   IV.R-carry-DIR=1SG.GEN   LOC=house
     'I will hold the fish at my house.' (Quick 2007: 303)

Referring to the distribution of the two applicative suffixes, Quick summarises as follows, "most applicative forms occur in both active and inverse constructions. However, a few applicative forms only occur in the inverse construction" (2007: 288). Among these, the function and the distribution of the latter (referred to as "applicative₁" in (55)) match those of the transitive affixes in Ibaloy. Thus, the characteristics of applicative₁ are shared by the two languages, while applicative₂ (55) is found only in Pendau. Possible directions of change, therefore, could be either that i) Ibaloy reflects the earlier system, and the distribution of the affixes extended in Pendau, or ii) Pendau reflects the earlier system, and the distribution of the affixes became restricted to Structure iii. in Ibaloy. The claim here is that various types of evidence suggest that the former is the actual direction of change.

Part of the claim that the general direction of the change is from the restricted occurrence, as in Ibaloy, to the extended one, as in Pendau, comes from the distribution of the two systems in Austronesian languages. The Ibaloy system is shared in languages in Taiwan and the Philippines and possibly in some non-Sumba-Sulawesi languages spoken in Indonesia (Figure 3.1). Languages outside of Sumba-Sulawesi, such as Taba, Fijian and Tongan, can be said to share the Ibaloy system in the sense that the affixed verbs occur only in a single clause structure. This is summarised in Table 3.14. Thus, the distribution of the two systems in the context of their proposed genetic relationship seems to imply that the Ibaloy system is more conservative, and the Pendau system is more innovative.

TABLE 3.14 The distribution of the applicative verb ending in non-Sunda-Sulawesi languages

	IBALOY	PENDAU	TABA	FIJIAN	TONGAN
Structure i.					
Structure ii.		-i, -a?			
Structure iii.	-i, -a, -an	-i, -a?	-o	-i	(-i, -Caki)
	and others		-ak	-aki	

Note: Parentheses indicate restricted occurrence.

This hypothesis is supported by the syntactic characteristics associated with the form -i in Pendau. Quick (2007: 302) claims that the suffix -i in Pendau has at least four known functions, and "[s]ome of these are clearly applicative, some seem marginally applicative and others appear to be idiosyncratic occurrences." He notes that -i has a "low degree of productivity," and many of the verbs with this form do not have a "locative" function. These appear to imply that the -i form is more lexicalised or somewhat fossilised and is probably older than the more productive suffix -a?, which more clearly shows the nature of applicative$_2$. It should be noted that the occurrence of verbal affixes is lexically determined in Ibaloy and the other languages mentioned in Table 3.14.

The assumed direction of change from an Ibaloy distribution where the affixes are restricted to Structure iii., to one where the affixes have spread to Structure ii. seems to be further supported by the distribution of the applicative suffix in Mamuju (Kaufman 2017). Mamuju is another language spoken in Sulawesi. It shows a similar system to that in Pendau, in that it has an A marking with origins in PMP genitive pronouns and in that it has two applicative suffixes, -i and -ang. However, according to Kaufman (2017), the forms -i and -ang in Mamuju are "incompatible" with structure ii. It is further mentioned that structures ii. are, however, allowed to combine freely with the forms -i and -ang when the agent is extracted, in, for example, relative clause constructions. This is the same condition under which definite patients are possible for actor voice predicates in Philippine-type languages (Kaufman 2017, see also Adams & Manaster-Ramer 1988). Thus, the distribution of the forms -i and -ang in Mamuju can be said to show a stage in between the Ibaloy and Pendau systems, as summarised in Table 3.15.

The condition Kaufman describes for the occurrence of an applicative suffix in structure ii. should help identify the motivation and developmental paths of the applicative suffix extending its distribution from Structure iii. to ii.

TABLE 3.15   The distribution of the applicative verb ending in Sunda-Sulawesi languages

	IBALOY	MAMUJU	PENDAU	INDONESIAN/MALAY
Structure i.				
Structure ii.		(*-i, -ang*)		*-i, -aʔ*	*-i, -kan*
Structure iii.	*-i, -a, -an* and others	*-i, -ang*	*-i, -aʔ*	

Note: Parentheses indicate restricted occurrence.

Indonesian and Malay have been added to the table. These languages share the applicative$_2$ system with Pendau. It is necessary, moreover, to examine the possibility that it was Indonesian, a lingua franca in the area, that provided the source of the applicative$_2$ in Pendau.

One of the findings of this distribution is a change in the co-occurrence pattern with the prefix *maN-* and the applicative suffix. The distribution of *maN-* was restricted to structure ii. in PMP (later spreading to structure i. in some languages), and never co-occurred with transitive ("applicative") suffixes. However, after the applicative verb suffixes extended their distribution to cover Structures ii. and iii., reflexes of PMP *maN- and one of the two PMP transitive suffixes, *-i or *-an, now co-occur in structure ii. The loss of the earlier functional difference between Structures ii. and iii. may be associated with this distributional change of the verb forms, however, this requires further investigation of other linguistic features associated with them.

5   Summary and Conclusion

The aim of this article has been to focus on the methodology as to how syntactic reconstruction is conducted with languages without a written record from the past. Any research on the historical development of such languages has to be based on the comparison and reconstruction of data almost solely from modern languages. It is claimed that Austronesian languages make a good candidate for this endeavour, for the genetic relationship among the languages is relatively well established, while the languages show diverse typological characteristics.

Five Austronesian languages were analysed and compared for the purpose of syntactic reconstruction. The selected languages are Ibaloy (ergative), Pendau (inverse), Taba (split between S and A), Fijian (accusative), and Tongan

(accusative pronominal and ergative non-pronominal systems). As the languages show typologically diverse systems, the question was how these different systems developed from an earlier system.

To answer this question, clause structures were abstracted. Structural patterns were described based on the combination of the verb and its argument noun phrases, and then they were classified according to the patterns of the occurrence of the grammatical case of the argument noun phrases. Monadic intransitive, dyadic intransitive, and transitive clauses were considered as basic clause structures and were described based on the argument structure. This descriptive method made it possible to compare structural patterns across typologically divergent languages.

In identifying the cognacy of the abstracted structures, it was shown to be useful to determine the position of each clause where the reflex of earlier genitive pronouns or the remnant of them occurs (labelled as the X-position). It was shown that the reflex could appear in a full or clitic pronominal set, or in reduced (grammaticalised) forms such as verb agreement or simply a consonant occurring on the verb. The existence of this position in each structure is considered a mark of the clause having developed from an earlier transitive clause, since genitive pronouns occurred as the A of transitive clauses.

Although the earlier genitive set is considered to have marked the A of transitive sentences, and thus the structure has been labelled as ergative, the positions associated with the reflexes turned out not to be found only in transitive clauses. Monadic and dyadic intransitive clauses also may have a reflex of the earlier genitive pronouns, including in a clearly accusative language, such as Fijian. This implies that the ergative marking function that was carried by the earlier genitive pronouns changed in some languages as a result of syntactic change. The distribution of the X-positions and the occurrence of the reflexes of the earlier genitive set were analysed in conjunction with one another. The functional change of the pronouns was shown to have resulted from the merger of two earlier pronominal sets, namely, nominative and genitive. It was argued that the motivation for this change was the change from a morphologically marked case-marking system to a word-order oriented system. Based on lexical reconstruction, the pronominal system of Proto-Oceanic had been reconstructed as showing three sets. However, there had been no explanation as to why there were three sets and how each developed in the pronominal system of Oceanic languages. The results of the syntactic reconstructions presented in this chapter clarified the development of the pronominal sets in Oceanic languages. Thus, it was shown that clarifying changes in clause structures also contributes to a better understanding of the development of morphological forms and systems.

The results of the presented reconstruction have the potential for being the bases for the comparison and reconstruction of other syntactic phenomena. To demonstrate this point, two cases of comparison and reconstruction were presented, namely, examination of word-order changes in Ibaloy and Pendau, and one on the extension of the function of certain verbal morphemes in some languages in Indonesia.

Syntactic reconstruction is indeed possible with languages, such as Austronesian, that typically have no old written records. It is possible by the comparison and reconstruction of the surface clause structures, integrated with lexical comparison and reconstruction, applying the traditional Comparative Method. Syntactic reconstruction supplements information that cannot be obtained through lexical comparison and reconstruction. In particular, knowing how syntactic features have changed is inevitable for tracing changes in the function of grammatical forms, since such changes are triggered by or result from syntactic change. In Austronesian historical linguistics, there is much that awaits such examination. One such example is the reconstruction of the verb morphology of Proto-Austronesian, proposed by Ross (2015). This is based on a rigorous examination of form and function correspondences of relevant reflexes and the results are presented in long paradigmatic lists. It is likely that the size of the list is partially due to the existence of what could be referred to as functional doublets. But sorting out syntactic change and the functional changes resulting from them, it is possible that these paradigms will be reduced with information about their developmental histories, just like the reconstructed pronominal system in Proto-Oceanic has been clarified.

As a new area of research, syntactic reconstruction has much to offer, not only for bringing in new knowledge about syntactic change but also extending the limits of the Comparative Method.

Abbreviations

-	boundary between an affix and its root
<xxx>	indicates that xxx is an infix
[xxx]$_{Aav}$	(Pendau) A of Vav
[xxx]$_{Aiv}$	(Pendau) A of Viv
[xxx]$_{Pav}$	(Pendau) P of Vav
[xxx]$_{Piv}$	(Pendau) P of Viv
[**xxxx**]$_{A}$	argument expressing A
[**xxxx**]$_{E}$	argument expressing E
[xxxx]$_{P}$	argument expressing P
[**xxxx**]$_{S}$	argument expressing S

+	(Ibaloy, PMP) augmented pronoun number
=	boundary between a clitic and its host
1	first person
1+2	first person inclusive
2	second person
3	third person
A	Agent of transitive verbs
ABS	(Pendau) absolute case
ABS	(Tongan) absolutive case
ActV	(Ibaloy) Actor Verb
ASP	aspect marker
AUG	augmented number
Aux	Auxiliary
AUX	Auxiliary
AV	(Pendau) active voice
BnfV	(Ibaloy) Beneficiary(-oriented) verb
CLTC	(Tongan) clitic
CNTV	(Ibaloy) continuative aspect
CP	clitic pronoun
CR	(Taba) cross-referencing form
DIR	(Tongan) directional
DIST	(Ibaloy) distal demonstrative
DL	dual
E	Extended argument of intransitive verbs
ERG	ergative case
EX	exclusive
GEN	genitive case
IMP	(Ibaloy) imperative
IN	inclusive
IND	independent pronoun
IPF	(Ibaloy) imperfective aspect
LOC	Locative
LOCV	(Ibaloy) Locative(-oriented) verb
MIN	minimum number
N	Noun
NEG	negative
NOM	Nominative
NP	Noun Phrase
OBL	(Ibaloy) oblique case
P	Patient of transitive verbs
P	preposition

PAST	past tense
PATV	(Ibaloy) Patient-oriented Verb
pb	proto-structure ii.
PFT	(Ibaloy) perfective aspect
PL	plural
PMP	Proto-Malayo-Polynesian
PNM	proper noun marker
POS	(Pendau) postural (verb class VI)
PotLocV	(Ibaloy) potentive Locative(-oriented) verb
PotPatV	(Ibaloy) potentive Patient(-oriented) verb
PRON	(Ibaloy) pronoun
PRS	present
R	(Pendau) realis
RDP	reduplicated part
REAL	(Taba) realis
REC	recognitional demonstrative; reciprocal marker
S	Subject (actor/undergoer) of intransitive verbs
SF	augmenting stem prefix former
SG	singular
SPEC	(Tongan)
StaV	(Ibaloy) Stative verb
THMV	(Ibaloy) Theme(-oriented) verb
V	Verb
VAUX	Auxiliary verb
Vav	(Pendau) active voice verb
Vdt	ditransitive verb
Vi	intransitive Verb
Vid	dyadic intransitive verb
Viv	(Pendau) inverse voice verb
Vst	(Taba) semi-transitive verb
Vt	transitive verb

References

Adams, Karen L. & Alexis Manaster-Ramer. 1988. Some Questions of Topic/Focus Choice in Tagalog. *Oceanic Linguistics* 27: 79–101.

Aldridge, Edith. 2004. *Ergativity and Word Order in Austronesian Languages*. Ph.D. dissertation, Cornell University.

Aldridge, Edith. 2010. Directionality in Word Order Change in Austronesian Languages. *Continuity and Change in Grammar*, ed. by Anne Breitbarth, Christopher Lucas, Sheila Watts & David Willis, 169–180. Amsterdam: John Benjamins.

Barðdal, Johanna. 2014. Syntax and Syntactic Reconstruction. *The Routledge Handbook of Historical Linguistics*, ed. by Claire Bowern & Bethwyn Evans, 343–373. London: Routledge.

Barðdal, Johanna & Thórhallur Eythórsson. 2012. Reconstructing Syntax: Construction Grammar and the Comparative Method. *Sign-Based Construction Grammar*, ed. by H. C. Boas & I. A. Sag, 257–308. Stanford, California: CSLI Publications.

Blust, Robert. 1990. Summary Report: Linguistic Change and Reconstruction Methodology in the Austronesian Language Family. *Linguistic Change and Reconstruction Methodology*, ed. by Philip Baldi, 133–154. Berlin: Walter de Gruyter.

Blust, Robert & Stephen Trussel. On-going. *The Austronesian Comparative Dictionary*. On-line at http://www.trussel2.com/acd/.

Bowden, John. 2001. *Taba: A Description of a South Halmahera Language*. Canberra: Pacific Linguistics.

Comrie, Bernard. 1989. *Language Universals and Linguistic Typology.* Chicago: University of Chicago Press.

Cysouw, Michael. 2003. *The Paradigmatic Structure of Person Marking.* Oxford: Oxford University Press.

De Smet, Hendrik. 2015. How Gradual Change Progresses. GLIMS [Ghent Research Team on Linguistic Meaning and Structure] Lecture Series. University of Ghent, Belgium.

De Vogelaer, Gunther. 2008. (De)grammaticalisation as a Source for New Constructions: The Case of Subject Doubling in Dutch. *Constructions and Language Change*, ed. by Alexander Bergs & Gabriele Diewald, 231–257. Berlin: Mouton De Gruyter.

Dixon, R. M. W. & Alexandra Y. Aikhenvald. 2000. *Changing Valency: Case Studies in Transitivity*. Cambridge: Cambridge University Press.

Ferraresi, Gisella & Maria Goldbach. 2008. Syntactic Reconstruction: Methods and New Insights. *Principles of Syntactic Reconstruction*, ed, by Gisella Ferraresi & Maria Goldbach, 1–26. Amsterdam & Philadelphia: Benjamins.

Fried, Mirjam. 2008. Constructions and Constructs: Mapping a Shift Between Predication and Attribution. *Constructions and Language Change*, ed. by Alexander Bergs & Gabriele Diewald, 47–80. Berlin: Mouton De Gruyter.

Gildea, Spike. 1998. *On Reconstructing Grammar: Comparative Cariban Morphosyntax*. Oxford: Oxford University Press.

Harris, Alice C. 2008. Reconstruction in Syntax: Reconstruction of Patterns. *Principles of Syntactic Reconstruction*, ed. by Gisella Ferraresi & Maria Goldbach, 73–96. Amsterdam & Philadelphia: John Benjamins.

Himmelmann, Nikolaus P. 1996. Person Marking and Grammatical Relations in Sulawesi. *Papers in Austronesian Linguistics* No. 3, ed. by Hein Steinhauer, 115–136. Canberra: Pacific Linguistics (*PL A*-84).

Kaufman, Daniel. 2017. Lexical Category and Alignment in Austronesian. *The Oxford Handbook of Ergativity*, ed. by Jessica Coon, Diane Massam & Lisa deMena Travis, 589–630. Oxford: Oxford University Press.

Kikusawa, Ritsuko. 2001. A Sketch of the Syntactic Structures of Nadrau Fijian. *Journal of Asian and African Studies* 62: 35–90.

Kikusawa, Ritsuko. 2002. *Proto Central Pacific Ergativity: Its Reconstruction and Development in the Fijian, Rotuman and Polynesian Languages*. Canberra: Pacific Linguistics.

Kikusawa, Ritsuko. 2003a. The Development of Some Indonesian Pronominal Systems. *Historical linguistics 2001: Selected Papers from the 15th International Conference on Historical Linguistics, Melbourne, 13–17 August 2001,* ed. by B. J. Blake & K. Burridge (with J. Taylor), 237–269. Amsterdam, Benjamins.

Kikusawa, Ritsuko. 2003b. A New View of the Proto-Oceanic Pronominal System. *Oceanic Linguistics* 42(1): 161–186.

Kikusawa, Ritsuko. 2003c. The Two Genitive Clitic Pronoun Positions in Some Extra-Formosan Languages: Problems in Determining Their Historical Development. *Papers from the Workshop on Brunei-Borneo Studies*, 1–20. Bandar Seri Begawan: University of Brunei.

Kikusawa, Ritsuko. 2015. Typological Generalisations and Diachronic Analyses: Actancy Systems in Austronesian Languages. *Historical Linguistics in Japan* 4: 3–35.

Kikusawa, Ritsuko. 2017. Ergativity and Syntactic Change in Austronesian Languages. *The Oxford Handbook of Ergativity*, ed. by Jessica Coon, Diane Massam & Lisa de-Mena Travis, 553–588. Oxford: Oxford University Press.

Kikusawa, Ritsuko. 2018. What the Tree Model Represents: Language Change, Time Depth, and Visual Representation. *Let's Talk about Trees: Genetic Relationships of Languages and Their Phylogenic Representation*, ed. by Ritsuko Kikusawa & Lawrence A. Reid, 171–193. Osaka: National Museum of Ethnology.

Kikusawa, Ritsuko. To appear. A Diachronic Typology of Applicative Verbs in Western Austronesian Languages: Toward a Comparison and Reconstruction (Synopsis). *Diachronic Typology Of Voice And Valency-Changing Categories*, ed. by L. Kulikov & S. Kittila.

Otsuka, Yuko. 2017. Ergative-Absolutive Patterns in Tongan: An Overview. *The Oxford Handbook of Ergativity*, ed. by Jessica Coon, Diane Massam & Lisa deMena Travis, 989–1006. Oxford: Oxford University Press.

Pawley, Andrew K. & Timoci Sayaba. 1971. Fijian Dialect Divisions: Eastern and Western Fijian. *The Journal of the Polynesian Society* 80(4): 405–436.

Payne, Thomas E. 1982. Role and Reference Related Subject Properties and Ergativity in Yup'ip Eskimo and Tagalog. *Studies in Language* 6(1): 75–106.

Quick, Phil. 1994. Active and Inverse Voice Selection Criteria in Pendau, a Western Austronesian language. *Proceedings of the Seventh International Conference on Austronesian Linguistics, Leiden 22–27 August 1994*, ed. by Cecilia Odé & W. A. L. Stokhof, 461–481. Amsterdam/Atlanta, GA: Rodopi.

Quick, Phil. 2007. *A Grammar of the Pendau Language*. Canberra: Pacific Linguistics.

Reid, Lawrence A. 2009. The Reconstruction of a Dual Pronoun to Proto Malayo-Polynesian. *Discovering History Through Language: Papers in Honour of Malcolm Ross*, ed. by Bethwyn Evans, 461–477. Canberra: Pacific Linguistics.

Reid, Lawrence A. 2016. Accounting for Variability in Malayo-Polynesian Pronouns: Paradigmatic Instability or Drift? *Journal of Historical Linguistics* 12: 130–164.

Roberts, Ian G. 2007. *Diachronic Syntax*. Oxford: Oxford University Press.

Ross, Malcolm D. 2015. Reconstructing Proto Austronesian Verb Classes. *Language and Linguistics* 16(3): 279–345.

Ruffolo, Roberta. 2004. Topics in the Morpho-Syntax of Ibaloy, Northern Philippines. Ph.D. dissertation, Australian National University.

Santiago, Paul Julian. 2016. *A Grammar of Kalanguya*. PhD dissertation, University of Osaka.

Starosta, Stanley, Andrew K. Pawley & Lawrence A. Reid. 1982. The Evolution of Focus in Austronesian. *Papers from the Third International Conference on Austronesian Linguistics, Vol. 2: Tracking the Travelers*, ed. by Amran Halim, Lois Carrington, & Stephen A. Wurm, 145–170. Canberra, Pacific Linguistics.

Watkins, Calvert. 1964. Preliminaries to the Reconstruction of Indo-European Sentence Structure. *Proceedings of the Ninth International Congress of Linguists*, ed. by Horace G. Lunt, 1035–1045. The Hague: Mouton.

External Possessor Constructions in Indo-European

Silvia Luraghi

Abstract

Two external possessor constructions occur in ancient Indo-European languages: the dative external possessor construction, and the double case construction. They both indicate adnominal possession by means of syntactically independent NPs, and basically refer to inalienable possession. In this article, I analyze the two constructions, describe their meaning and their syntactic properties, and review the comparative evidence for each of them. Neither construction is uniformly attested throughout the Indo-European language family. In addition, the dative external possessor construction seems to be quite unstable over time. Based on the data presented, I conclude that the former can be reconstructed as an original Proto-Indo-European construction, while the latter must be regarded as a language specific construction, with different properties in the languages in which it occurs.

1 Introduction

External possessor constructions have been described in the typological literature as constructions in which two NPs indicate a possessor and a possessum, but are not in a relation of syntactic dependency. Typically, external possessor constructions involve two syntactically independent NPs, each of which can function as an argument of the predicate with which they co-occur, as argued in Payne & Barshi (1999: 3), who provide the following definition: "constructions in which a semantic possessor-possessum relation is expressed by coding the possessor … as a core grammatical relation of the verb and in a constituent separate from that which contains the possessum."

Two types of external possessor construction occur in ancient Indo-European languages: one in which the possessor is in the dative, and one involving double case. Homeric Greek offers examples for both constructions:

(1)   *Hérēi   d'   ouk ékhade   sthêtos   khólon*
      Hera:DAT PTC NEG contain:AOR.3SG breast:NOM wrath:ACC
      'Hera's breast could not withhold wrath.' (Hom. *Il.* 4.24 – Greek)

(2)   *tón   r'   Oduseùs …   bále   dourì   kórsēn*
      3SG.ACC PTC Odysseus:NOM hit:AOR.3SG spear:DAT brow:ACC
      'Odysseus hit his brow with a spear.' (Hom. *Il.* 4.501–502 – Greek)

In both sentences we find a nominal expression that indicates a person plus
the name of a body part that belongs to this person. In (1), the dative NP *Hérēi*
'Hera.DAT' refers to the possessor of the body part *sthêtos* 'breast.NOM', which
is the subject of the sentence. Similarly, in (2) the accusative demonstrative *tón*
'this.ACC' refers to the possessor of the body part *kórsēn* 'brow.ACC', which is
the direct object of the sentence.

   In this article, I review the data concerning these two constructions in order
to test their reconstructability as a common Proto-Indo-European heritage.
I argue that the former should be reconstructed as a Proto-Indo-European
construction, in spite of not being as uniformly attested as is usually said in the
literature, while the latter most likely cannot.

   The article is organized as follows. In Section 2 I describe the dative external
possessor construction. After a brief introduction in which I show what types
of possessa it usually involves, I review the comparative data from Ancient
Indo-European languages (Section 2.1). I then discuss the meaning of the
construction, and the semantic role of the external possessor (Section 2.2).
In Section 2.3, I contrast the dative external possessor construction with the
construction of verbs of depriving. Section 2.4 describes diachronic changes
in various Indo-European languages, in which the dative external possessor
construction has either been replaced by other constructions, or has arisen as
a replacement. In Section 2.5 I discuss the reconstructability of the construc-
tion. Section 3 is devoted to double case constructions. I start by describing the
double accusative in Homeric Greek (3.1). In Section 3.2 I discuss the mean-
ing of the Greek construction. Then I survey the evidence from other ancient
Indo-European languages, except for Hittite and Armenian (3.3). Section 3.4 is
devoted to the Hittite and Armenian data concerning double case construc-
tions. In Section 3.5 I discuss the reconstructability of the double case con-
struction based on comparative evidence. Section 4 contains the conclusion.

2   Dative External Possessor

External possessor constructions involving the dative are a frequently oc-
curring feature of Indo-European languages, including many modern ones.
According to Haspelmath (1999), external possessor constructions of this type
are an areal feature of the modern languages of Europe. Concerning the an-
cient Indo-European languages, this construction has been long recognized
and it is exhaustively described in the handbooks. For example Havers (1911)
offers an exhaustive description covering all ancient Indo-European languages
except Anatolian and Tocharian, which were not known at the time when the
book was written. He shows the extent to which different types of possessors
could feature in this construction, and the development of the construction
over time. He rightly acknowledges the difference between this type of dative,
for which he uses the name of *dativus sympatheticus,* and the beneficiary da-
tive, even though it is not completely clear that all the types of *dativus sympa-
theticus* he lists are to be kept apart from beneficiary uses of the dative case.
Indeed, the trademark of the *dativus sympatheticus* according to Havers is that
it is equivalent to a genitive or to a possessive pronoun/adjective, and in fact
genitive NPs and possessives also occur in the same contexts. This is true for
his types (i)–(v) (examples 3–8), while type (vi) (examples 9–10) is slightly dif-
ferent, as equivalent expressions feature various types of spatial NPs or PPs,
rather than possessive expressions. Havers' term *dativus sympatheticus* is often
translated as 'dative of affection', or 'dative of interest'. In this article, I will
mostly use *dativus sympatheticus* to refer to it, in order to avoid confusion with
other possible functions of the dative.

  Below are the six types of *dativus sympatheticus* according to Havers, with
examples from ancient Indo-European languages.

i)   The event affects the human body or one of its parts:
   (3)   *minatur   mihi   oculos   exurere*
         threaten:PRS.3SG 1SG.DAT eye:ACC.PL burn_out:INF.PRS
         'He's threatening to burn my eyes out.' (Pl. *Men.* 843 – Latin)

ii)   The event affects the spirit or the feelings of a human being:
   (4)   *hós   moi   epôrse   ménos*
         REL.NOM   1SG.DAT   arouse:AOR.3SG   strength:ACC
         'Who aroused my strength.' (Hom. *Il.* 20.93 – Greek)

iii)   The event affects personal possessions of a human being, including kin-
       ship relations:
       (5)   *ašte   komou   bratrŭ   oumŭretŭ*
             if   INDF.DAT   brother:NOM   die:PRS.3SG
             'If somebody's brother dies.' (Luke 20.28 – Old Church Slavic)

iv)   Sentences that contain verbs of depriving:
       (6)   *jah saei   ni   habaiþ   jah þatei   habaiþ*
             and REL.NOM NEG have:PRS.3SG and DEM.NOM have.PRS.3SG
             afnimada   imma
             take:PRS.3SG.P   3SG.DAT
             'And he that has not, from him shall be taken even that which he
             has.' (Mark 4.25 – Gothic)

v)   Human relations:
       (7)   *sijai   allaim   skalks*
             be:SBJ.3SG all:DAT.PL   slave:NOM
             '(That one) will be the slave of all.' (Mark 10.44 – Gothic)

       (8)   *patḕr   dé   moí   esti   Polúktōr*
             fater:NOM   PTC   1SG.DAT be:PRS.3SG Polyctor:NOM
             'My father is Polyctor.' (Hom. *Il.* 24.397 – Greek)

vi)   Contexts in which the dative alternates with a locative expression:
       (9)   *oúte tis   oûn moi   noûsos   epḗluthen*
             NEG INDF.NOM PTC 1SG.DAT disease:NOM come_upon:AOR.3SG
             'Nor did any disease come upon me.' (Hom. *Od.* 11.200 – Greek)

       (10) *hōs Aphrodítē   êlthen   Árēi   epíkouros*
             as   Aphrodite:NOM come:AOR.3SG Ares:DAT helper:NOM
             'As Aphrodite came as Ares' helper.' (Hom. *Il.* 21.430–431 – Greek)[1]

Havers' six categories do not all instantiate the same construction. In this ar-
ticle, I will mainly be concerned with categories (i), (ii), (iii) and (v), which

---

1   An anonymous reviewer wonders whether this is not an occurrence in which the dative indi-
    cates a human relation, as in (7) and (8). In fact, the difference is pretty clear. In (10) it is not
    said that Aphrodite is Ares' helper, but that she came as helper: in other words, it is the whole
    phrase *epíkouros érkhesthai* 'come as helper' that takes a beneficiary dative, which could also
    be encoded by a prepositional phrase.

are indeed cases of inalienable possession. The other two categories should be treated differently, for various reasons that I illustrate shortly in this section, and will take up again in Sections 2.3 and 2.4. It is also remarkable that some of the occurrences in category (v) seem better understood as cases of beneficiary dative, as I will show in Section 2.5.

Type (iv) is often treated as containing a dative external possessor (cf. for example Fried 1999, Haspelmath 1999), and it frequently occurs in languages that also feature types (i)–(iii) and (v). However, here I take it as a different construction, as verbs of depriving are ditransitive (see Malchukov et al. 2010). Consequently, I consider the occurrences of a dative third argument with such verbs in the larger framework of language specific ditransitive constructions, and suggest that it be viewed as an extension of the construction of verbs of giving, which, in the same languages, typically take a third argument in the dative. Type (vi) cannot be regarded as instantiating a possessive construction: rather, the dative here has other functions, indicating a participant which is indeed affected by an event, but cannot be understood as a possessor. The examples in Havers (1911) show that the most frequent interpretation of type (vi) datives is beneficiary, as in (10) or, less frequently, experiencer in inchoative experiential situations, as in (9). Notably, in Vedic this type frequently involves nouns rather than pronouns, and it only has a beneficiary interpretation (see below, Section 2.5).

2.1 *Comparative Data*

External possessors in the dative are prototypical possessors, that is, they are animate, most often human.[2] As I have already remarked above, Havers (1911) gives abundant data for all ancient Indo-European languages except Anatolian and Tocharian, data which indicate that the dative external possessor in the categories (i)–(iii) and (v) was especially frequent for first and second person pronouns, while being less frequent or even unattested depending on the language for third person (demonstrative) pronouns[3] and other nominal categories.

However, a word of caution must be said concerning whether the occurrence of a dative in this construction in all languages described by Havers is

---

2 Inanimate possessor can also appear in external possessor constructions in some languages, as long as they are in a part-whole relation with the possessee. See Baldi & Nuti (2010: 352) on Latin.

3 In ancient Indo-European languages the third person pronoun was in origin a demonstrative, and no dedicated third person pronoun similar to first and second person pronouns can be reconstructed for Proto-Indo-European.

real. Indeed, as Havers himself acknowledges, it is hard if not impossible to distinguish morphologically the dative from the genitive of first and second person pronouns in some of these languages. In particular, in Sanskrit the enclitic forms *me* and *te* are both genitive and dative of first and second person singular pronouns. (Likewise, dual and plural clitic personal pronouns do not have distinct forms for the dative and the genitive.) Dative accented forms also exist, which are distinct from genitive forms, i.e. *máhyam* and *túbhyam*. Notably, however, they do not occur as *dativus sympatheticus* (Havers mentions a single occurrence of *tubhyam*, see below, Section 2.5). So, when Havers (1911: 44) says that the two occurrences in (11) and (12) typically illustrate the distinction between first and second person-dative on the one hand and third person-genitive on the other, one must be aware that synchronically this is not the case.

(11)   *... te ...   yujnami   hári*
        2SG.DAT   yoke:PRS.1SG   stallion:ACC.PL
        'I yoke your stallions.' (RV 3.35.4 – Sanskrit)

(12)   *yunjánt   asya ...   hári ...   ráthe*
        yoke:PRS.3PL   3SG.GEN   stallion:ACC.PL   cart:LOC
        'They yoke his stallions to the cart.' (RV 1.6.2. – Sanskrit)

Havers (1911: 9–10; 14–17) takes great care in demonstrating that the forms *me* and *te* could function as dative in other constructions, for example with verbs of giving. Crucially, however, with such constructions other nominal categories also occur in the dative, so it is plausible that the clitic forms were synchronically felt as dative. But when we come to the *dativus sympatheticus*, the complementary distribution of the genitive and the putative dative forms does not support this interpretation: from the synchronic point of view there is no reason to set up a distinction between a dative and a genitive here (cf. also Haudry 1977: 69–70). So, synchronically, the forms *me* and *te* can be regarded as genitives, an interpretation which is supported by the occurrence of the genitive in this construction with all other types of constituent. The Iranian evidence, too, is hardly compelling. In Old Persian, syncretism of the dative and the genitive was already complete for all types of nominal categories before the time of the written attestations. Avestan, in which the two cases are still separate in the declension of nouns and accented pronouns, provides limited evidence only for Havers' category (v), but crucially not for types (i)–(iii), as, similar to Sanskrit, it features syncretic genitive-dative forms for pronominal clitics.

From the diachronic point of view, *me* and *te* go back to Proto-Indo-European **moi* and **toi*, which are dative forms. So one could still value the Indo-Iranian evidence, as it could preserve an ancient Proto-Indo-European construction, which was no longer synchronically available for speakers. However, one must also consider the evidence from the languages that were not yet known in Havers' time, that is, Hittite and Tocharian. Concerning the latter, not much can be said, again on account of the merger of the dative and the genitive. Hittite, the oldest attested Indo-European language, offers a quite unexpected picture. In Old Hittite original texts, pronominal datives such as those shown above in the other Indo-European languages do not occur. Instead, a set of possessive enclitics occurs, as shown in examples (13) and (14).

(13) *nu=*   *us*   *appa*   *ishi=*   *ssi*
      CONN   3PL.ACC   back   master(C):DAT   3SG.POSS.DAT.SG.C
      pennai
      drive:PRS.3SG
      '(S)he takes them (= the oxen) back to their (sg.) owner.'

                                                                                    (*Laws* § 79 – Hittite)

(14) *ki*   ᴳᴵˢ̌TUKUL-*li=*   *met*   *ki=*   *ma*
      DEM.NOM   craft(N):NOM   POSS.1SG.N   DEM.NOM   CONN
      sahha=   *mit*
      feudal.duty(N):NOM   POSS.1SG.N
      'This is my craft, and this is my feudal duty.' (*Laws* § 40 – Hittite)

According to Gamkrelidze & Ivanov (1995: 250–253), possessive enclitics are limited to inalienable possession in Old Hittite. However, examples such as (14) cast some doubts on this hypothesis. Baldi (2002) further mentions occurrences such as *halugatalla(n)=tin* 'your envoy', and adds that genitive forms of personal pronouns are almost never used in Old Hittite texts, while enclitic possessives are virtually the only items that can express pronominal possession. It thus seems safer not to posit a link between possessive clitics and inalienability. Such clitics also occur as cross-referencing devices with possessive genitives, in a construction that involves both head and dependent marking, as in (15) and (16).[4]

---

4 Following common transcription conventions for cuneiform languages, capitalized roots (e.g. LUGAL) represent Sumerian logograms, which are separated by a hyphen from the second part of the word, which reproduces Hittite lexical material. This hyphen does not indicate a morpheme boundary, and so has no analogue in the gloss line.

(15) LUGAL-*was aras=*   *mis*
  king:GEN friend(C):NOM 1SG.POSS.NOM.C
  'The friend of mine, the king.' (KUB 29.1 i 35 – Hittite)

(16) [*laba*]*rnas* LUGAL-*as* NINDA= *san*   *adue*[*ni*]
  labarna:GEN king:GEN bread(C) 3SG.POSS.ACC.C eat:PRS.1PL
  [*wata*]*r*= *set=*   *a* *akueni*
  water(N) 3SG.POSS.ACC.N PTC drink:PRS.1PL
  'Let us eat the bread of the Labarna, the king, let us drink his water!'
              (StBoT25.140, Rev. 5'–6' – Hittite)

In the case of this latter construction a connection with inalienable possession has also been suggested (Garrett 1998). As remarked by Baldi (2002: 32) it is hard to draw universal boundaries of what can be inalienably possessed, and the discussion becomes circular (see further Chappell & McGregor 1996 for a general discussion of inalienable possession, Kockelman 2009 on linguistic reflexes of inalienable possession, and Heine 1997 on inalienability and the role of body parts). In any case, what is clear even about cross-reference clitics with genitives is that they are not limited to part-whole relations.

  As noted above, dative external possessors are not attested in Old Hittite texts written in Old Script (see Güterbock 1983: 75; Luraghi 1997: 23). Some occurrences can be found in later copies of Old Hittite texts, but this does not mean that they existed in the original, due to the widespread and well attested scribal habit of updating the language (see further the discussion of example (55) in Section 3.4). Indeed, dative external possessors appear starting from the Middle Hittite period and are frequent in New Hittite. Examples are (17) and (18).

(17) ᴸᵁ́KÚRᴹᴱˢ= *mu=* *kan* ᴸᵁ́·ᴹᴱˢ*arsanatallus* ᴰ*IŠTAR* GAŠAN-
  enemy-PL 1SG.OBL PTC evious:ACC.PL Istar lady
  YA ŠU-*i* *dais*
  my hand:DAT put:PRT.3SG
  'Istar my Lady put in my hand my enemies and those who were envious of me.' (*StBoT* 24 i 58–59 – Hittite)

(18) *nasma=* *ssi=* *kan* *garates* *adantes*
  or   3SG.DAT PTC bowels:NOM.PL eat:PTCP.NOM.PL
  'Or (when) its bowels have been eaten.' (KUB 7.1+ i 2 – Hittite)

As the examples show, this construction closely resembles the dative external possessor construction of the other ancient Indo-European languages: but the chronology within Hittite is puzzling, as it seems to be an independent development out of an older stage, in which possession was expressed differently with respect to the other languages.

2.2   *The Meaning of the Dative External Possessor Construction*

As already remarked, dative external possessors are especially frequent with first and second person pronouns. After his remarks on Sanskrit, Havers (1911: 317–320) argues that this must be true of Indo-European in general, as is shown by the evidence from the other languages. In Homeric Greek, for example, Havers analyzed 550 occurrences with the dative and 395 with either the genitive or a possessive adjective, and found that first and second person pronouns featured 410 times in the first group of occurrences, while occurring only 85 times in the second group (1911: 104; similar data come from Germanic, pp. 274–285; 299–300 and 323–324).

Thus, especially for the categories which clearly involve inalienable possession, the pattern of the ancient Indo-European languages in which the dative and the genitive are distinct (with the exception of Anatolian) points toward a split between first and second person pronouns on the one hand, and other nominal categories on the other hand. The dative external possessor was the preferred construction for first and second person pronouns, while other nominal categories most often took genitive possessors or possessive adjectives. Due to the limited number of occurrences, it is hard to gauge the difference between the dative external possessor construction and other constructions. Havers (1911) and Bally (1926), as well as more recent studies such as Fried (1999); König & Haspelmath (1997) and Haspelmath (1999) highlight the importance of affectedness, or 'interest' in more traditional terms, as a feature of the semantic roles taken by external possessors. However, as remarked by Fried (1999), affectedness is a feature of patients, and external possessors are clearly not patients. In particular, they do not undergo a change of state; in addition, they are high on the animacy hierarchy, while prototypical patients are inanimate. What seems distinctive for external possessors is the possibility of perceiving the effects of an event. König & Haspelmath (1997) quote examples (19) and (20) from Roldán (1972).

(19) *Sus*   *ojos*   *se*   *llenaron*   *de*   *lágrimas*
     POSS.3SG.PL   eyes   REFL   filled.up   of   tears

(20) *Los   ojos   se   le   llenáron   de   lagrimas*
  the   eyes   REFL   3SG.DAT   filled.up   of   tears
  'His/her eyes filled up with tears.' (Spanish)

Example (19) is appropriate in a situation in which the possessor is not mentally affected, i.e. not because some type of emotion made him/her start weeping: this would be the appropriate situation for (20). The situational context for (19), instead, could be one in which tears are triggered by some mechanical effect, as in the case of a patient being treated during surgery.

  It is hard to provide evidence from ancient languages, in which interpretations may be speculative. However, examples (21) and (22) from Homeric Greek might perhaps match the Spanish examples above.

(21) *séo   d'   ostéa   púsei   ároura   keiménou*
  2SG.GEN PTC bone:ACC.PL rot:FUT.3SG earth:NOM lie:PTCP.PRS.GEN
  en   Troíēi
  in   Troy:DAT
  'And your bones shall the earth rot as you lie in the land of Troy.'
              (Hom. *Il.* 4.174–175 – Greek)

(22) *grápsen   dé   hoi   ostéon   ákhris   aikhmḕ*
  scratch:AOR.3SG PTC 3SG.DAT bone:ACC spear:GEN point:NOM
  Pouludámantos
  Polydamas:GEN
  'The spear-point of Polydamas cut to his bone.'
              (Hom. *Il.* 17.599–600 – Greek)

In (21), which features a genitive possessor (*séo* 2SG.GEN), reference is made to the bones of a person after his death, and perception by the possessor is ruled out here. In (22), instead, the dative possessor (*hoi* 2SG.DAT) refers to a living warrior, who perceives the wound.

  Being animate, typically human, and being perceivers are features of experiencers, rather than of patients. External possessors are also frequently described as similar to beneficiaries, for example in Haspelmath (1999). Notably, the author highlights the relevance of mental affectedness for external possessors (p. 112). Note however that mental affectedness is not a necessary feature of beneficiaries. Beneficiaries are partially affected by the event in which they take part, in the sense that they must be "capable of making use of the

benefit bestowed upon them" (Kittilä & Zuñiga 2010: 2), but they do not necessarily perceive the situation when it takes place. Consider example (23).

(23) *I baked a cake for Mary.*

In (23) it is not implied that the beneficiary has any consciousness of the prospective benefaction at the time of the event. On the other hand, experiencers are by definition involved in an experience, i.e. they perceive the effects of a situation. Thus, I conclude that the semantic role of dative external possessors is experiencer, rather than beneficiary. This is in accordance with the connection between this construction and inalienable possession. Inalienably possessed entities are conceptualized as parts of the possessor, and an event which affects a part of a human possessor is necessarily perceived by the possessor (see also below, Section 3.2).

2.3   *Verbs of Depriving: Maleficiary Dative*

Let us now turn to the dative with verbs of depriving, which I mentioned briefly in Section 2. Rosén (1959) argues that in Homeric Greek the verb *apháireō* 'take away' takes the dative most frequently when the thing which is stolen from someone is an inalienable possession (four occurrences out of five).[5] However, this does not seem to be the case in other languages discussed by Havers, such as Latin or various Germanic languages, and it is by no means true of Hittite, in which all ditransitive verbs normally take the dative when the third argument is animate at all language stages (Starke 1977, Luraghi 1986), while types (i)–(iii) of *dativus sympatheticus* do not occur in the oldest texts (see Section 2.1). Some examples are:

(24) nu=   ssi=   ssan   ᴳᴵˢ*huesan*   ᴳᴵˢ*hulali*=   ya   arha
      CONN   3SG.DAT   PTC   spindle:ACC   distaff:ACC   and   off
      dahhi   nu=   ssi   ᴳᴵˢBAN ...   *pehhi*
      take:PRS.1SG   CONN   3SG.DAT   bow   give:PRS.1SG
      'I take the spindle and distaff from him and give him a bow.'
                                        (KUB 9.27 + KUB 7.8 i 23–25 – Hittite)

---

5   Verbs of depriving normally took the double accusative in Homer; the cases in which one finds a dative are limited. See Jacquinod (1989: 215, 223–225) and Luraghi & Zanchi (2018: 28–33).

(25) *Haec   prius   illi   detrahenda   auxilia*
DEM.ACC.PL before 3SG.DAT take.away:SUP.ACC.PL auxiliary:ACC.PL
existimabat   quam   ipsum   bello   lacesseret
think:IMPF.3SG   than   same:ACC   war:ABL   provoke:SBJ.IMPF.3SG
'He thought that these auxiliaries ought to be detached from him before
he provoked him to war.' (Caes. *Gal.* 6.5.5 – Latin)

(26) *pàvogé   tám   séniui   tá   óżká*
steal:AOR.3SG   DEM.DAT   old.DAT   DEM.ACC.PL   goat:ACC.PL
'He stole the goats from the old man.'
                    (Schl. 128.7, quoted from Havers 1911: 304 – Lithuanian)

(27) *dyde   him   of   healse   hring   gyldenne*
do:PRT.3SG   3SG.DAT   from   neck   ring   golden
'He took away the golden ring from his neck.' (*Beowulf* 2810 – Old English)

The cross-linguistic distribution of this construction and the type of items in-
volved also differs with respect to types (i)–(iii) and (v) of *dativus sympatheticus*.
As Havers (1911) shows, this construction is infrequent in Homeric Greek and
Vedic Sanskrit, in which types (i)–(iii) and (v) are very widespread, while it
is very common in Latin, Germanic and Balto-Slavic. In these languages, the
construction of verbs of depriving is frequent with all nominal categories, in-
cluding common nouns, and does not display a split between first and second
person pronouns on the one hand, and other nominal categories on the other
hand, even though types (i)–(iii) and (v) show this split to varying extents.

In the modern languages of Europe, many of which also feature this con-
struction, it is by no means clear that it only concerns inalienable possession.
Consider examples (28) and (29) from Italian.

(28) *Gli   hanno   rubato   l'   auto*
3SG.DAT   they.have   stolen   the   car
'Somebody stole his car.' (Italian)

(29) *Gli   hanno   rubato   l'   auto   che   gli   aveva*
3SG.DAT   they.have   stolen   the   car   that   3SG.DAT   he.had
prestato   Giovanni
lent   Giovanni
'Somebody stole from him the car that he had borrowed from Giovanni.'
                                                                (Italian)

While in (28) the default interpretation is that the car belongs to the person indicated by the dative pronoun *gli*, (29) makes it clear that this implication is not obligatory. Note further that verbs of depriving take the same construction as the verb 'give', as shown in (30).

(30) *Gli   ho   dato   un   libro*
     DAT.3SG   I.have   given   a   book
     'I gave him a book.' (Italian)

Remarkably, the occurrence of a dative does not per se indicate which construction occurs in a sentence, even in cases in which two sentences contain the same verb, as shown in examples (31) and (32).

(31) *Gli   ho   aperto   gli   occhi*
     DAT.3SG   I.have   opened   the   eyes
     'I opened his eyes.' (Italian)

(32) *Gli   ho   aperto   la   porta*
     DAT.3SG   I.have   opened   the   door
     'I opened the door for him.' (Italian)

In (31) the dative *gli* indicates a possessor, and has the same function as the external possessor dative in the examples from the ancient Indo-European languages discussed thus far. In (32), the same dative form has a beneficiary function: the different meaning is triggered by the occurrence of an inalienably possessed entity in (31) and one which is not in (32). (Note that this is not limited to dative forms of pronouns: in both examples, one can also replace *gli* with a prepositional phrase, such as *a Giovanni* 'to Giovanni.') Even the occurrence of a noun that refers to an entity which is inalienably possessed in normal conditions is not enough to rule out a beneficiary or maleficiary interpretation, if the context makes it possible, as shown in (33).

(33) *Il   gatto   gli   ha   mangiato   il   fegato   che   aveva*
     the   cat   3SG.DAT   has   eaten   the   liver   that   had
     lasciato   sul   tavolo
     left   on_the   table
     'The cat ate the liver that he left on the table **on him**.' (Italian)

2.4 *The Dative External Possessor Construction and Stability over Time*

Havers (1911) analyzes the development of the *dativus sympatheticus* and argues that, especially in Greek and in Latin, the construction started declining

at a rather early time, and was virtually absent toward the end of antiquity, as shown by the data from the New Testament. The absence of the dative external possessor construction in Late Latin is quite striking, as this construction is well attested in the early stages of all Romance languages (Havers 1911: 232).[6] In fact, it seems that Havers' arguments for increasing limitations of the construction throughout the history of Latin are based on an incorrect evaluation of the style of different authors. According to Baldi & Nuti (2010: 351) "we find examples in every kind of text and from authors of every age. External possession is differently applied by the authors (e.g., more often in Caesar, less so in Cicero …), and it strengthens with time." Concerning Late Latin, the authors cite an occurrence from the *Peregrinatio Egeriae* (6th century CE, p. 352). In addition, while Havers claims that the Latin New Testament does not feature the construction, a cursory exam of first and second person singular pronouns limited to the four Gospels shows that this is not the case. In (34), a dative external possessor matches a genitive pronoun in the Greek text:

(34) | *Quid* | *fecit* | *tibi* | *quomodo* | *aperuit* | *tibi* |
|---|---|---|---|---|---|
| INDF.ACC | do:PF.3SG | 2SG.DAT | how | open:PF.3SG | 2SG.DAT |
| *oculos?* | | | | | |
| eye:ACC.PL | | | | | |
| *Tí* | *epoíēsén* | *soi?* | *pôs* | *énoixén* | *sou* |
| INDF.ACC | do:AOR.3SG | 2SG.DAT | how | open:AOR.3SG | 2SG.GEN |
| *toùs* | *ophthalmoús?* | | | | |
| ART.ACC.PL | eye:ACC.PL | | | | |

'What did he do to you? How did he open your eyes?'

(John 9.26 – Latin and Greek)

It is possible that the rarity of the construction in the Latin New Testament can be caused by the influence of the Greek text: as is well known, Latin translations tried to stick as close as possible to the original (see e.g. Ceresa-Gastaldo 1975).

The history of Greek offers a very interesting picture. According to Havers (1911: 167–169), the dative external possessor construction had already started

---

6   Bauer (2000: 158) thinks that "The loss of the *dativus sympatheticus* is presumably related to the loss of the case system. This assumption is supported by evidence from languages where its use is still attested, such as German or Russian, which have case. Similarly in Modern French the structure is still used in pronominal, but not in nominal contexts: in contrast to nouns, pronouns still feature some case marking." This statement ignores the fact that in other Romance languages, such as Italian or Spanish, prepositional phrases with *a* 'to' function as the *dativus sympatheticus* (e.g. *Ho rotto una gamba a Giovanni* 'I broke John's leg' while a similar sentence in French is hardly acceptable, see also Bally 1926 and Haspelmath 1999).

declining during the classical time, and, as remarked above, was lost by the time of the New Testament. At the wake of the Byzantine age, the dative merged with the genitive in the paradigms of all nominal categories. In spite of this, in Modern Greek it is possible to distinguish between the dative and the genitive of enclitic pronouns, as the former takes the verb as its host, while the latter cliticizes to it head noun, as argued in König & Haspelmath (1997: 555). Consider (35) and (36).

(35) *tu*   *éspases*   *to*   *xéri*
      3SG.DAT   break:PRT.2SG   ART   hand

(36) *éspases*   *to*   *xéri*   *tu*
      break:PRT.2S   ART   hand   3SG.GEN
      'You broke his hand.' (Modern Greek)

In (35), the clitic *tu* is hosted by the verb: this is the position of dative clitics, and this is also the normal way to convey the propositional content of the example. In (36), the same form is attached to the nominal head *xéri* 'hand', and functions as a genitive modifier. Thus, the dative external possessor construction which was lost before the disappearance of the dative case was recreated at a later time.

König & Haspelmath (1997) and Haspelmath (1999) argue that the dative external possessor construction is an areal feature of the modern languages of Europe: in the first place, it is not limited to Indo-European languages, but also occurs in Basque, Maltese, Modern Hebrew, and Hungarian. Notably, at least in the Semitic languages, contact seems to have played a relevant role, as this construction does not occur in Classical Arabic or in Biblical Hebrew. In the second place, the construction does not occur in languages that are spoken at the margins of the European linguistic area, such as English and the Scandinavian languages, Celtic, Finnic languages, Turkish, and East Caucasian languages (Haspelmath 1999: 116). Remarkably, both Old English and Old Norse featured the construction in ancient times, as the other Germanic languages (Havers 1911: 299).

To sum up, the dative external possessor construction seems to be a rather unstable pattern: in the course of the attested history of Greek, for example, one can see its decay and its renewal. Changes are attested in other Indo-European languages: Old Hittite displays possessive pronouns, with the dative external possessor construction emerging during the Middle Hittite time, accompanied by the decay of possessive pronouns. In West Germanic the construction remained relatively stable over time, but it decayed in English

and in North Germanic. In at least two of the non-Indo-European languages of Europe, it emerged, possibly because of language contact. In the Slavic languages, the construction remained stable over time. Remarkably, however, the attested history of these languages starts at a much later time than the attested history of Greek, Latin, and the non-European Indo-European languages (and the same can be said for West Germanic).

In the non-European Indo-European languages, the construction was soon lost, because of early merger of the dative and the genitive. In fact, this development, which started with clitic forms of personal pronouns, makes it difficult to gauge the extent to which the construction existed at all in these languages. In Sanskrit, syncretism affected only first and second person clitic pronouns, but this made it synchronically impossible to detect the construction, as such pronouns featured prominently in it in other languages. (Note further that both dative and genitive clitic forms occurred in P2 in Vedic Sanskrit, making it impossible to distinguish between them, as one can do in Modern Greek.) The construction did not even arise at a later time, as the dative and the genitive merged in Middle-Indo-Aryan. Viti (2004) argues that inalienable possession tended to be expressed via nominal composition in Vedic. In the case of pronominal possessors, Rosén (1989) shows that compounds such as *mad-dehas* 'my body' or *tvat-putra* 'your son' are the most frequent strategy for the encoding of inalienable possession in Classical Sanskrit. Such compounds are formed with the stem of personal pronouns (*mad-* 1SG, *tvad-* 2SG) plus a noun. This stem, which only occurs in compounds, is also the base for derivation of a set of possessive adjectives, whose usage, however, remained limited in Sanskrit. Putative external possessors in the dative are not a feature of Classical Sanskrit: apparently, the morphological strategy of compounding replaced the syntactic strategy in the encoding of inalienable possession.

2.5   *Reconstructability of the Construction*

In Section 2.4 I have shown that the dative external possessor construction has been a quite unstable feature of the Indo-European languages over the centuries. In addition, it seems to be easily borrowed through contact. The interpretation of the data from ancient Indo-European languages is not as unambiguous as it is usually thought to be: among the oldest branches of Indo-European, only Greek and Latin, both from the same area, offer abundant evidence for the construction. Germanic and Balto-Slavic also provide good evidence, but these languages are attested much later, and have been in contact already in pre-literary times. Hittite, the most ancient attested Indo-European language, provides puzzling evidence: the construction does not occur in Old Hittite, but emerges immediately thereafter. The evaluation of the Indo-Aryan data

is more complex than usually thought, but it is of crucial importance in order to rule out the possibility that the dative external possessor construction can really be reconstructed for Proto-Indo-European and is not an unstable feature of the Indo-European languages of Europe.

As we have seen in Section 2.1, first and second person pronouns cannot offer evidence for the existence of the dative external possessor construction in Vedic Sanskrit. Pronominal clitics, which only had a single form for the dative-genitive, did indeed derive from ancient Indo-European datives. However, as they were synchronically not analyzable as such, and since other nominal categories occurred in the genitive in the same construction, they could have been introduced into it on account of their genitive meaning. Evidence for dative external possessors is very scanty: Havers (1911: 30) mentions a passage in which the form *túbhyam* (2SG.DAT) occurs as *dativus sympatheticus* of type (ii). However, as Havers shows, the form *táva* (2SG.GEN) is also attested in a similar passage (cf. *ib.*), and with type (i) he mentions a number of occurrences of the accented genitive form *mama* (1SG.GEN). Is it possible to consider this single occurrence as evidence for the original form of the construction? The interpretation of the very limited number of other dative forms of pronouns and nouns in Havers' types (i)–(iii) is doubtful: as Havers argues, they can mostly be taken as beneficiary datives. The only type of *dativus sympatheticus* in which nominal categories other than first and second person pronouns occur is type (v). Notably, Vedic occurrences listed by Havers in this category do not seem to indicate possession in as clear a manner as those from other languages. Consider example (37).

(37) *taṃ*   *tvā*   *vayaṃ ...*   *śāsmahe ...*   *sakhe*
     3ACC.SG   2.ACC.SG   1NOM.PL   pray:PRS.MID.1.PL   friend:VOC.F
     vaso   *jaritṛbhyaḥ*
     beneficent:VOC.F   praiser:DAT.PL
     'We pray you for that, o friend beneficent to the praisers.'
                                                                            (RV 30.10 – Sanskrit)

In (37) the dative *jaritṛbhyaḥ* 'to the praisers' can easily be understood as a beneficiary dative. This difference becomes clear if we compare (37) with (8) from Homeric Greek: in (8) it is not said that Polyctor is a father for the speaker, but that he is his father. In (7) from Gothic one could also argue for a beneficiary interpretation of the dative. However, the occurrence of a genitive in the Greek original (*pántōn* 'of all') supports a possessor interpretation, as does the whole context: "But so shall it not be among you: but whosoever will be great among

you, shall be your minister. And whosoever of you will be the chief, shall be servant of all." (Mark 10.44).

The beneficiary meaning of the dative, inherited from Proto-Indo-European, remains productive in Indo-Iranian, as far as one can see in the languages in which the dative has not merged with the genitive. However, in Sanskrit dative experiencers are less frequent than in many other ancient Indo-European languages, and apparently their occurrence is semantically restricted. Regarding Vedic, Dahl (2014) notes that dative experiencers are limited to positive situations, "while parallel predicates denoting a negative state of mind ... have a genitive-marked Experiencer argument in Early Vedic." This fact shows that dative experiencers share the features of the beneficiary role.

As I discussed in Section 2.3, the dative in external possessor constructions must be understood as an experiencer. Partial disconnection between the dative and the experiencer role may be the reason why the morphological merger of the dative with the genitive of clitic pronouns brought about the decay of the construction. In the meantime, the association of the genitive case with the possessor role on the one hand, and of the dative case with the beneficiary role on the other hand became stronger (notably, the preferred strategy for encoding experiencers in Vedic Sanskrit was the nominative case, see Dahl 2014). In this connection, it is also remarkable that Havers' type (vi) of *dativus sympatheticus* mostly feature a beneficiary. In particular, occurrences similar to (9) from Homeric Greek, in which the dative is an experiencer involved in a (negative) inchoative situation do not occur in Vedic, in accordance with Dahl's observation. All occurrences have a strong positive facet, and support a beneficiary reading, as in (38).

(38) ā   gantā   nūnaṃ   nas   avasā   yathā purā
     PREV   come:PTCP   now   2PL.ACC help:INS   as   in_the_past
     itthā   kaṇvaya   bibhyuṣe
     so   Kanva:DAT   frighten:PTCP.DAT
     'Come now to us with your help, as in the old days, so now for frightened Kanva's sake.' (RV 1.39.7 – Sanskrit)

On account of the above remarks, the limited evidence offered by Vedic can be considered consistent with the reconstruction of the dative external possessor construction as belonging to Proto-Indo-European in general, and not only for the Indo-European languages of Europe; even so, it must be stressed that the Hittite evidence remains problematic. Concerning Vedic, the decay of the *dativus sympatheticus* can likely be considered one of the manifestations of

ongoing marginalization of the dative, which lost, or never featured, a number of functions typical of the dative in other ancient Indo-European, languages: for example, the dative of agent with gerundives is attested in Vedic but was later was replaced by the instrumental (see Luraghi 2016), the dative of possession does not occur even in the oldest texts. In view of the increasing reduction of the functions of the dative in Sanskrit, the evidence offered by some dative accented forms of personal pronouns noted by Havers must be taken as preserving an ancient construction, thus they are valuable for reconstruction of Proto-Indo-European.

3   Double Case

This construction has been referred to by different names: part and whole construction, partitive apposition (Hahn 1953, 1954), and, limited to Hittite and Armenian, case attraction (Luraghi 1993 and 2008). It is found in Greek, Armenian, and Anatolian; examples from Latin are limited and partly controversial while the remaining languages display no compelling evidence for it. Sporadic occurrences in Indo-Aryan point toward low entrenchment of the construction (Wiedmer 2014), and show notable difference with respect to Homeric Greek, as discussed in Section 3.3. Although the Armenian construction had long been recognized as due to language contact (see below Section 3.4), Hahn (1953, 1954) held double case to be the oldest way in which possession was expressed in Indo-European. More recently, Luraghi (1990, 1993) and Garrett (1990: 79–91) have argued that it was the outcome of a recent development in Anatolian as well. As I will show while reviewing the data, cases possibly featured in double case constructions change significantly depending on the language. As the double accusative in Homeric Greek is the best known and best described instantiation of double case construction, I start by illustrating its features in the next section.

3.1   *The Double Accusative in Homeric Greek*
The double accusative is widespread in Homeric Greek, occurring in cases of inalienable possession limited to body parts, feelings, and, in a small number of passages (four occurrences according to Jacquinod 1989: 25), pieces of a warrior's armor.[7] An example has already been given in (2); further occurrences are (39) and (40).

---

7 Jacquinod also mentions one occurrence with an inanimate possessor, also featuring a part-whole relation; see further Luraghi & Zanchi (2018: 15).

(39) *hḗ*   *se*   *pódas*   *nípsei*
      DEM.NOM.F   2SG.ACC   foot:ACC.PL   wash:FUT.3SG
      'She will wash your feet.' (Hom. *Od.* 19.356 – Greek)

(40) *hṓs min*   *érōs*   *pukinàs*   *phrénas*   *amphekálupsen*
      as   3SG.ACC desire:NOM wise:ACC.PL mind:ACC.PL roll_up:AOR.3SG
      'Then love encompassed his wise heart.' (Hom. *Il.* 14.294 – Greek)

The double accusative, which is very frequent in Homeric Greek, is mostly limited to poetic style in later authors, most likely influenced by Homer. It corresponds to Havers' groups (i) and (ii): the four occurrences in which parts of an armor are referred to can also be included here, as the armor is understood as an integral part of the body of a fighting warrior. In some occurrences, the two constructions are used with the same verbs, as shown in Jaquinod (1989: 16). Thus we find occurrences such as *tòn ... blápse phrénas* '(a god) injured his (acc.) mind (acc.)' (Hom. *Od.* 14.178) with a double accusative, and *blápse ... hoi ... goúnata* '(he) injured his (dat.) knees (acc.)' (Hom. *Il.* 7.271), with a dative external possessor.

    Hahn (1953) further mentions several passages that she claims provide evidence for other double case constructions. However, a closer examination shows that this is not the case: purported instances of double dative feature a dative external possessor and a locative expression in the dative, most often accompanied by a spatial particle, as in (41). Similarly, putative double genitives present a genitive possessor and a genitive second argument, as in (42).

(41) *ménos*   *dé*   *hoi*   *en*   *phresì*   *thêke*
      courage:ACC   PTC   3SG.DAT   in   heart:DAT   put:AOR.3SG
      'And (he) set courage in his heart.' (Hom. *Il.* 21.145 – Greek)

(42) *goúnōn*   *hápsasthai*   *Laertiádeō*   *Odusêos*
      knee:GEN.PL   clasp:INF.AOR   son.of.L.:GEN   O.:GEN
      'To clasp the knees of Odysseus, son of Laertes.' (Hom. *Od.* 22.339 – Greek)

In (41), the dative *hoi* indicates the possessor, while the dative *phresí* in connection with the particle *en* denotes the endpoint of the trajectory indicated by the verb *títhemi* 'put'. This verb normally takes a direction complement in the dative, or with *en* plus dative (see Luraghi 2003: 83–84). In (42) again we find a possessor expression, this time in the genitive, *Laertiádeō Odusêos*. The genitive *goúnōn* is the second argument of the verb *háptō*, which always takes the genitive in Homeric Greek.

3.2   *The Meaning of the Construction*

In this construction, we find transitive verbs that typically take accusative direct objects. The special feature here is that we find two accusative objects, one that indicates the possessor, and one that indicates a special type of possessee, that is, a part of the possessor. From the semantic point of view, we find two patients in this construction. Indeed, inalienable possession as instantiated here implies that both the possessee and the possessor are affected in the same way by the event: if I get injured in a part of my body, I myself get injured, and if desire gets hold of my mind, it gets hold of myself (see the discussion in Jacquinod 1989: 26–28 and the references therein). This is the reason why types of possessee in the double accusative construction are more limited than those that occur with the dative external possessor construction. Dative external possessors are affected when an event affects the entity that they possess inalienably. However, the type of affectedness is not necessarily the same. Kinship relations are a case in point: if one of my close relatives is injured, this certainly affects me, but it does not imply that I myself am injured.

Syntactically, the two accusative NPs in the double accusative construction have different status. The possessor NP displays the properties of a direct object, in that it can be passivized, while the possessee NP cannot, and also occurs in the accusative in passive sentences, as shown in (43) and (44).

(43) *blêto …*   *knémen   dexiterén*
      hit:AOR.3SG.M/P   calf:ACC   right:ACC
      'He was wounded in his right leg.' (Hom. *Il.* 4.518–519 – Greek)

(44) *Atreídes   d'   ákhei   megáloi   bebolménos*
      Atreus'_son:NOM   PTC   pain:DAT   great:DAT   hit:PTCP.PF.M/P.NOM
      êtor
      soul:ACC
      'Atreus' son, hit in his soul by great pain.' (Hom. *Il.* 9.9 – Greek)

The possessee NP is syntactically in apposition to the possessor NP, and in the passive it takes an adverbial status. Such an appositional accusative also features in another construction, commonly known as 'Greek accusative' or 'accusative of respect'. It also often involves body parts. In Homer, it most frequently occurs with the indication of a quality of a person, and the body part to which this quality is especially referred, as in (45).

(45) *pódas   okùs   Akhilleús*
      foot:ACC.PL   swift:NOM   Achilles:NOM
      'Achilles swift-footed' (Homer, *passim* – Greek)

Besides body parts, this type of appositional accusative occurs with nouns denoting moral qualities, bodily activities, shape, and measure or other inherent qualities (see Jacquinod 1989: 42–43). In later Greek, the last group of words became frequent. Examples are (46) and (47).

(46) *phrourâs   eteías   mêkos*
watch:GEN   yearly:GEN   length:ACC
'A watch that has been lasting years.' (Aesch. *Ag.* 2 – Greek)

(47) *diaphérei   he   gunḕ   andròs   tḕn*
differ:PRS.3SG   ART.NOM   woman:NOM   man.GEN   ART.ACC
phúsin
nature:ACC
'Women are different from men by their nature.' (Plato *Rep.* 453b – Greek)

It seems likely that the appositional accusative which originally occurred in double accusative possessive constructions later acquired the status of a special construction, in which it indicated the particular area of a referent to which a general predication applied. This special construction, which is a trademark of Ancient Greek syntax, does not occur in other ancient (or modern) Indo-European languages.

3.3   *Double Case in Other Ancient Indo-European Languages*
As remarked in Section 3, evidence for the double accusative constructions in other ancient Indo-European languages is hardly compelling. Jacquinod reviews a limited number of occurrences mentioned by Delbrück (1888), and concludes that the only convincing occurrence in Vedic is (48), which contains a noun that refers to an abstract property.

(48) *ahaṃ   ni-anyaṃ   sahasā   sahas   karaṃ*
1SG.NOM down-other:ACC strength:INS strength:ACC do:INJ.AOR.1SG
'I overcome another's strength with strength.' (RV X 49.8 – Sanskrit)

Note however that, as Jaquinod (1989: 59) also points out, several other interpretations have been suggested for this passage. Wiedmer (2014) surveys some other sporadic occurrences of double case in Vedic. As shown in (49), concrete nouns can also occur.

(49) *ahám*   *etã́ñ ...*   *dvă-dvā ...*   *índraṃ*   *yé*
     1SG.NOM   DEM.ACC.PL   two_by_two   Indra:ACC   REL.NOM.PL
     vájraṃ   *yudháye*   *ákr̥ṇvata*
     mace:ACC.PL   fight:INF   make:IMPF.3PL
     'I (struck down) by twos those who caused Indra's mace to fight.'
                                                           (Rigveda X 48.06)

In Vedic, this type of double accusative is very infrequent, but, in the mean-
time, it is not limited to body parts as it is in Greek (in fact, body parts do not
occur at all): as shown in (48) and (49), featured nouns may refer to concrete
or abstract entities. In addition to this, Wiedmer also discusses another type of
double accusative, with verbs of depriving, in which however the two accusa-
tives must be regarded as depending on the verb, since, as already discussed in
Sections 2 and 2.3, verbs of depriving are bivalent predicates. Indeed, Wiedmer
(2014: 33) points out that '[t]his type [is] syntactically clearly different from
Ancient Greek.' Among other things, both arguments can be passivized, while
in the Greek double accusative construction described in Sections 3.1 and 3.2
only the possessor noun can be passivized (Section 3.2). Wiedmer (2014: 22)
also mentions sporadic occurrences of double locative. One of the few clear
ones is (50).

(50) *ā́*   *hí*   *ruhátam*   *aśvinā*   *ráthe*   *kóśe*
     to   PTC   ascend:IMP.2DU   Aśvin:VOC.DU   chariot:LOC   cask:LOC
     hiraṇyáye
     golden:LOC
     'Ascend into the golden cask of the chariot, Aśvins!' (Rigveda VIII 22.9)

In (50), the NP *kóśe hiraṇyáye* 'golden cask' indicates a part of the chariot
(*rátha-*), and both are inflected in the locative. Notably, this pattern differs
both from the pattern in (49) and from the pattern found in Homeric Greek,
because the possessor is inanimate. (Wiedmer further cites an isolated occur-
rence from Avestan.)

   The Latin evidence, which has been especially highlighted in Hahn (1954),
boils down to a couple of examples in Plautus, which, however, can also be
taken as left dislocated accusatives, a construction which is quite frequent in
Latin informal discourse (Ernout & Thomas 1959). Consider example (51):

(51) *hunc   senem   osse   fini   dedolabo   assulatim*
      DEM.ACC old.man:ACC bone:ABL end:ABL hew:FUT.1SG in.bits
      viscera
      entrails:ACC.PL
      'This old fellow, I will hew to his very bone, (making) his entrails into
      mincemeat.' (Pl. *Men.* 858–859 – Latin)

In (51) *hunc senem* 'this old fellow' is the possessor of *viscera* 'entrails'.
Syntactically, it seems better to analyze it as a hanging topic, rather than as the
head noun of a phrase with an apposition.

   A few other examples from later poetry, notably from Vergil, are better un-
derstood as stylistic imitations of Greek epic poetry. Hahn (1953) also mentions
an alleged example of double ablative with instrumental meaning in Latin,
shown in example (52):

(52) *dextera   digitis   rationem   computat*
      right:ABL finger:ABL.PL calculation:ACC count:PRS.3SG
      'With (his) right (hand), he counts on (his) fingers.'
                                                    (Pl. *Mil.* 203–205 – Latin)

Hahn's translation reflects her assumption that *dextera digitis* must be
taken as a partitive apposition: 'he counts using the finger of his right hand'.
However, the context does not support this assumption. The opposition *lae-
vam manum … dextera* 'his left hand … his right hand' rather points toward the
correctness of the translation given in (52): *ecce avortit; nixus laevo in femine
habet **laevam manum, dextera** digitis rationem computat; ferit femur dexterum*
'Look, he turns; he is leaning with his left hand on his left thigh, with his right
hand he is counting on his fingers; he hits his right thigh.' (see Luraghi 1993 for
further discussion). Finally, Jacquinod (1989: 60) mentions a small number of
occurrences from Middle High German, which could be taken as double ac-
cusatives but could also be interpreted in other ways, while Old High German
does not contain traces of this construction.[8]

   The adduced evidence shows different types of appositional constructions,
with different properties and containing different types of nouns, rather than
attesting to a pattern consistently reflected in the languages surveyed (i.e. the
appositional constructions surveyed are not cognate).

---

8 Hahn (1969) also indicates the naming construction of the type *rājā Nalo nāma* 'a king, Nala
   (was) is name' (Sanskrit) as reflecting a partitive apposition. For a critical appraisal of this
   book, which was published posthumously, see Beekes (1973).

3.4   *Case Attraction in Armenian and Hittite*

Double case constructions involving both the accusative and other cases are clearly and abundantly attested only in Armenian and in Hittite. Let us consider examples (53)–(56) from Armenian (examples (54–56) are from Vogt 1932).

(53) *i*   *mijoy*   *cocoy*   *kʻumē*
    from   middle:ABL   bossom:ABL   POSS.2SG.ABL
    'From the center of your heart.' (*Psalms* 73.11 – Armenian)

(54) *i*   *knojê*   *tʻagaworē-n*
    from   wife:ABL.SG   king:ABL.SG-ART
    'By the wife of the king.' (Armenian)

(55) *baznowtʻeamb*   *zawrawkʻ-n*   *Hayoc*
    majority:INS   troops:INS-ART   Armenian
    'With most of the Armenian troops (lit.: the majority of the troops).'
                                                  (Armenian)

(56) *varowkʻ*   *lawowtʻeamb*
    life:INS.PL   virtue:INS
    'Through a virtuous way of life (lit.: a life of virtue).' (Armenian)

In (53) and (54) we find two occurrences of double ablative; the first occurrence features a part-whole relation between a referent (the heart) and one of its regions (the center), while in the second we find a kinship relation (the wife of the king). Examples (55) and (56) show two double instrumentals. Here, again, we find a part-whole relation in the first example (the majority of the troops), while the second indicates that a quality (virtue) is attributed to a certain referent (life).

As one can easily see from the examples, there are significant differences between the double case construction in Armenian and the double accusative construction in Homeric Greek, concerning both the semantics of the arguments and their morphological encoding. On the semantic plane, it must be remarked that, in the first place, inanimate possessors in Armenian are quite frequent in this construction. In the second place, types of possessee in the Armenian constructions are not limited to (body) parts, and the construction is not limited to part-whole relations, as shown by example (54). This example also shows that the two NPs involved need not have the same semantic role, as something which is done by the king's wife cannot be said to be done by

the king. This is in contrast to the semantics of the Homeric double accusative construction (see Section 3.2). Moreover, on the plane of morphological encoding, it is important to note that double case constructions in Armenian only involve the ablative and the instrumental: there are no double genitives and, most notably, no double accusative to match the Homeric construction (see Caha 2013: 1023).

Some Hittite examples are (57)–(60).

(57) *takku* LÚ.ULÙ^{LU}-*an* *ELLAM* KAxKAK= *set*
if man:ACC free nose 3SG.POSS.ACC
kuiski *waki*
INDF.NOM bite:PRS.3SG
'If someone bites a free men on his nose.' (*Laws* § 13 (B i 33) – Hittite)

(58) *nu=* *za* *ke* KUR.KUR ^{LÚ}KÚR *INA* MU 10.KAM
CONN PTC DEM.ACC.PL countries enemy in year ten
ammedaz ŠU-*az* *tarahhun*
1SG.ABL hand:ABL conquer:PRT.1SG
'I conquered these enemy lands in ten years by my hand.'
(*AM* 136.45–46 – Hittite)

(59) *nu=* *kan* GAL-*in* *arunan* ^D*Ku(ma)rbiyaza* É-*irza* ...
CONN PTC big:ACC sea:ACC Kumarbi:ABL house:ABL
uwater *n=* *an* *INA* É- *ŠU arha pehuter*
bring:PRT.3PL CONN 3SG.ACC into house his back bring:PRT.3PL
'They brought the big sea out of Kumarbi's house, and carried him to his (own) house.' (*StBoT* 14.11.16–19 – Hittite)

(60) *IŠTU* ^{HUR.SAĜ}*Hahruwa* *tuedaz* *assiyantaza*
from Hahruwa _mountain 2SG.ABL love:PTCP.ABL
'From your beloved mountain Hahruwa.' (KUB 36.90.19 – Hittite)

Example (57) contains a double accusative (the nose of a free man). This occurrence is indeed similar to the Homeric Greek double accusative construction, in that it features a noun referring to a human being and one referring to a body part. However, the Hittite construction also shows differences with respect to the Homeric one, both on the semantic and on the morphological plane. Indeed, examples (58)–(60) show double ablatives, and in fact the ablative case is frequent in this construction. The first example contains an ablative with instrumental function, again featuring a part-whole relation with a

body part noun (my hand). Examples (59) and (60) feature animate possessors with different type of possessee, the house in (59) and an external region (the mountains) in (60). While examples (57) and (58) are evocative of the Homeric construction, (59) and (60) show that in Hittite, too, types of possessee are not as restricted. In particular, while someone's house is likely to be categorized as an instance of inalienable possession, the occurrence of the name of a mountain in (60) even casts doubts on the fact that the construction is really limited to inalienable possession.

While the rise of the Armenian construction has long been recognized as caused by influence from Georgian (see Vogt 1932 and the discussion below), Hittite examples of double case have been held by some scholars as representing the oldest Indo-European pattern, only later replaced by the adnominal genitive (see Hahn 1954: 199, Jacquinod 1989: 62–64). However, following the chronology of the texts it becomes apparent that double case is a later development in Hittite. In the first place, it must be remarked that examples as (57) and (59) could not date back to the OH period, since personal pronouns are never inflected in the ablative in Old Hittite. Indeed, even ablative forms of animate nouns such as ^{D}Ku(ma)rbiyaza in (58) apparently never occur in Old Hittite original texts (see Starke 1977).

Double accusatives such as the one in (57) in principle could be possible in Old Hittite. However, the Old Hittite Corpus written in Old Script hardly offers evidence for such a hypothesis. On the contrary, comparing older with more recent versions of the Hittite Laws it becomes apparent that the double accusative has replaced the older construction with the possessor noun in the genitive co-indexed by a possessive enclitic pronoun (discussed in Section 2.1). In fact, example (61) is an older version of the same passage given above as example (57), which is taken from a post-Old Hittite copy of the Laws:

(61) *takku*   LÚ.ULÙ^{LU}-*as*   *ELLAM-as*   KAxKAK=   *set*
      if   man:GEN   free:GEN   nose(N)   3SG.POSS.ACC.N
      kuiski   *waki*
      INDF.NOM   bite:PRS.3SG
      'If someone bites the nose of a free man.' (*Laws* § 13 (A i 24) – Hittite)

In (61) the possessor noun is inflected in the genitive: this is the only available construction in Old Hittite original texts. The possessee takes the enclitic possessive adjective *=set*, which is a neuter accusative and indicates the direct object function. Note that this clitic also occurs in (57). This does not mean that possessive clitics were a feature of the double case construction: indeed they were not, and the occurrence of *=set* in (57) is a consequence of the copyist's partial updating of the text. Here, the copyist replaced the genitive with an

accusative for the possessor noun, in order to introduce the new construction, but then forgot to leave out the possessive clitic, which was typical of the older construction.

As early as 1932, Vogt pointed out that agreement of head and modifier, generally called case attraction, in Classical Armenian must have developed under the influence of an Old Georgian pattern called suffix copying, as in example (62):

(62) *šecevn-ita   cmid-isa   sameb-isa-jta*
     help-INS   holy-GEN   trinity-GEN-INS
     'With the help of the Holy Trinity.' (from Boeder 1995: 159 – Georgian)

In (62) we find a head noun (help) inflected in the instrumental, and a genitive modifier (Holy Trinity). The modifier, after the genitive ending, also 'copies' the instrumental ending of the head noun. Notably, Georgian has agglutinating morphology, rather than the fusional type of the ancient Indo-European languages. With regard to the two types of construction, Vogt writes:

> In Georgian the case ending of the head noun … is repeated after all modifiers, either adjectives, pronouns, genitives, prepositional phrases or noun cases which already contain the expression of a case relation…. The difference between the two expressions [i.e. suffix copying in Georgian and double case in Armenian] consists in this, that Georgian allows cumulation of case endings whereas in Armenian the ending that marks the constituent forces away the ending that expresses the relation between head and modifier. Since the two languages have different [morphological] means, facts are not the same in detail. However, the general tendency of the two languages displays a striking similarity, making case endings also function as markers of noun phrases (1932: 75).[9]

In Luraghi (1993), (1994), and (2008) I have argued that the double case construction attested in Hittite is a borrowing from a neighboring language,

---

[9] "En géorgien la desinence casuelle du nom déterminé … se répète après tout déterminant, que ce soi un adjectif, un pronom, un génitif, une expression prépositionelle, ou encore un group pronominal comportant déjà l'expression de la relation casuelle…. La différence entre les deux expressions consiste en ceci, que le géorgien permet l'accumulation des désinences casuelles tandis qu'en arménien, la désinence marquant le group chasse la désinence marquant le rapport entre le déterminé et le déterminat. Comme les moyens dont disposent les deux langues sont différentes, le détail des faits n'est pas le meme. Les tendances générales des deux langues n'en montre pas moin une concordance frappante, en laissant les désinences casuelles servire aussi d'indicateurs des groupes nominaux".

Hurrian, which also featured agglutinative morphology and suffix copying (see Wilhelm 1995), as shown in (63):

(63) *sen(a)-iffu-we-ne-va*   *torub(i)-i-va*
      brother-my-GEN-ART.SG-DAT   enemy-his-DAT
      'To my brother's enemy.' (Mit. III 114 – Hurrian)

In practice, in the Armenian and Hittite double case constructions, nouns show a partly adjectival behavior, in that they are assigned case not based on their function, but as targets of agreement. In Hittite, personal pronouns go as far as agreeing in number with the head noun, thus becoming real possessive adjectives. An example is (64), which features a plural form of the second person singular pronoun:

(64) *tuedas*   *assiyantas*   *pedas*
      2SG.DAT.PL   love:PTCP.DAT.PL   place:PTCP.DAT.PL
      'In your favorite places.' (KUB 36.90.16 – Hittite)

According to Vogt, there is evidence, for example in the use of the article (examples 54 and 55) for assuming that the two nouns that occur in the Armenian case attraction construction belong to the same constituent. In other words, the noun denoting the possessor functions both morphologically and syntactically as an attributive adjective. Evidence from pronouns seems to point in the same direction for Hittite. Thus, the double case construction appears to be syntactically different from the Homeric Greek double accusative, in which the two accusative NPs remained independent, as shown by their behavior with regard to passivization. Therefore, the Homeric and the Hittite constructions do not appear to be cognate.

3.5   *Reconstructability of the Construction*

The evidence reviewed in the preceding sections does not support the reconstruction of a double case external possessor construction in Proto-Indo-European. Indeed, the comparative evidence shows that ancient Indo-European languages did not consistently feature this construction, either with the accusative or with other cases, the only exception being the Homeric Greek double accusative.

The Armenian and the Hittite double case constructions were later developments, independent of other Indo-European languages, and were both due to language contact. In both languages, double case constructions arose under the influence of partly similar constructions featured by non-Indo-European

languages with agglutinating morphology. Their syntactic and semantic properties were different from those of the Homeric double accusative construction, and both historical and structural considerations indicate that, although superficially similar to one another, these constructions in the three languages were unrelated.

Thus, the Homeric Greek double accusative construction, despite being attested in one of the most ancient branches of the Indo-European language family, must be considered a language specific construction, rather than a piece of evidence for a Proto-Indo-European reconstruction.

4   Conclusion

In this article, I have analyzed two types of external possessor constructions occurring in ancient Indo-European languages: the dative external possessor construction, and the double case construction. For each of them, I reviewed the comparative data, and described the relevant semantic and syntactic properties. Both constructions have been considered to be of common Proto-Indo-European heritage by several scholars, even though the comparative data does not provide unambiguous evidence for either of them. In particular, the dative external possessor construction, which is today an areal feature of the languages of Europe, was more clearly attested in the European Indo-European languages already in antiquity. In spite of early merger of the dative and the genitive of pronominal clitics in Indo-Iranian languages, limited evidence from accented personal pronouns in Vedic attests to the antiquity of the construction. Concerning the double case construction, sizable and reliable evidence is provided only by three languages: Homeric Greek, Classical Armenian, and Middle and Late Hittite.

Based on the evidence, I suggest that only the former construction must date back to Proto-Indo-European, and that similar instantiations in the Indo-European languages actually provide pieces of evidence for its reconstruction, despite its early disappearance from the non-European Indo-European languages. The latter construction, in its turn, cannot be reconstructed. Double case constructions in Greek, Armenian, and Hittite have different syntactic and semantic properties, and arose from different historical backgrounds, while sporadic occurrences of double case in other languages do not reflect the construction attested in Homeric Greek. Hence, they should better be regarded as different constructions, independent of one another, rather than as reflexes of an original proto-construction.

Abbreviations

a)   *Glosses*

1	1st person	M/P	medio-passive
2	2nd person	MID	middle
3	3rd person	N	neuter
ABL	ablative	NEG	negation
ACC	accusative	NOM	nominative
AOR	aorist	OBL	oblique
ART	article	P	passive
C	common gender	PF	perfect
CONN	connective	PL	plural
DAT	dative	POSS	possessive
DEM	demonstrative	PREV	preverb
DU	dual	PRS	present
F	feminine	PRT	preterite
FUT	future	PTC	partiple
GEN	genitive	PTCP	participle
IMP	imperative	REFL	reflexive
IMPF	imperfect	REL	relative
INDF	indefinite	SBJ	subjunctive
INF	infinitive	SG	singular
INJ	injunctive	SUP	supine
INS	instrumental	VOC	vocative
LOC	locative		

b)   *Sources*

Aesch.	Aeschylus	*Men.*	*Menaechmi*
Ag.	*Agamemnon*	*Mil.*	*Miles Gloriosus*
AM	*Mursilis' Annals*	Mit.	Mitanni letters
Caes.	Caesar	*Od.*	*The Odyssey*
Gal.	*The Gallic War*	Pl.	Plautus
Hom.	Homer	*Rep.*	*The Republic*
Il.	*The Iliad*	RV	*Rigveda*
KUB	Keilschrift Urkunden aus Bogazkoy	*StBoT*	*Studien zu den Bogazkoy Texten*

References

Baldi, Philip. 2002. Inalienable Possession in Hittite and Proto-Indo-European. *Donum Grammaticum: Studies in Latin and Celtic Linguistics in Honour of Hannah Rosén*. ed. by Lea Sawicki & Donna Shalev, 29–37. Louvaine-la-Neuve: Peeters.

Baldi, Philip & Andrea Nuti. 2010. Possession in Latin. *New Perspectives on Historical Latin Syntax*, ed. by Philip Baldi & Pierluigi Cuzzolin, vol. 3, 239–387. Berlin: Mouton de Gruyter.

Bally, Charles. 1926. L'expression des idées de sphère personnelle et de solidarité dans les langues Indo-européennes. *Festschrift Louis Gauchat*, ed. by Franz Fankhauser & Jakob Jud, 68–78. Aarau: Verlag H. R. Sauerländer & Co. [English translation: 1996. The Expression of Concepts of the Personal Domain and Indivisibility in Indo-European Languages. *The Grammar of Inalienability*, ed. by Hilary Chappell & William McGregor, 31–61. Berlin/New York: Mouton de Gruyter.]

Bauer, Brigitte L. M. 2000. *Archaic Syntax in Indo-European: The Spread of Transitivity in Latin and French*. Berlin/New York: Mouton de Gruyter.

Beekes, R. S. P. 1973. Review of E. A. Hahn, *Naming-Constructions in Indo-European Languages* Cleveland: Press of Western Reserve University, 1969. *Mnemosyne* 26(4): 399–401.

Boeder, Winfried. 1995. Suffixaufnahme in Kartvelian. *Double Case*, ed. by Frans Plank, 151–215. Berlin/New York: Mouton de Gruyter.

Caha, Pavel. 2013. Explaining the Structure of Case Paradigms by the Mechanisms of Nanosyntax: The Classical Armenian Nominal Declension. *Natural Language and Linguistic Theory* 31(4): 1015–1066.

Ceresa-Gastaldo, Aldo. 1975. *Il latino nelle antiche versioni bibliche*. Roma: Studium.

Chappell, Hilary & William McGregor. 1996. Introduction: Prolegomena to a Theory of Inalienability. *The Grammar of Inalienability*, ed. by Hilary Chappell & William McGregor, 3–30. Berlin/New York: Mouton de Gruyter.

Dahl, Eystein. 2014. The Morphosyntax of the Experiencer in Early Vedic. *Perspectives on Semantic Roles*, ed. by Silvia Luraghi & Heiko Narrog, 181–204. Amsterdam/Philadelphia: Benjamins.

Delbrück, Berthold. 1888. *Altindische Syntax*. Halle: Verlag der Buchhandlung des Waisenhauses.

Ernout, Alfred & Francois Thomas. 1959. *Syntaxe latine*. Paris: Klincksieck.

Fried, Mirjam. 1999. From Interest to Ownership: A Constructional view of External Possessors. *External Possession*, ed. by Doris L. Payne & Imanuel Barshi, 473–504. Amsterdam/Philadelphia: Benjamins.

Gamkrelidze, Thomas & Vjaceslav V. Ivanov. 1995. *Indo-European and the Indo-Europeans: A Reconstruction and Historical Analysis of a Proto-Language and a Proto-Culture*. Berlin/New York: Mouton de Gruyter.

Garrett, Andrew. 1998. Remarks on the Old Hittite Split Genitive. *Mír Curad: Studies in Honor of Calvert Watkins*, ed. by Jay Jasanoff, H. Craig Melchert & Lisi Oliver, 155–163. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.

Garrett, Andrew. 1990. The Syntax of Anatolian Pronominal Clitics. Ph.D. Diss., Harvard University.

Güterbock, Ans Gustav. 1983. Noch einmal die Formel parnaššea šuwaizzi. *Orientalia Nuova Serie* 52: 73–80.

Hahn, E. Adelaide. 1953. Vestiges of Partitive Apposition in Latin Syntax. *TAPhA* 82: 92–123.

Hahn, E. Adelaide. 1954. Partitive Apposition in Homer and the Greek Accusative. *TAPhA* 85: 197–289.

Hahn, E. Adelaide. 1969. *Naming-Constructions in Some Indo-European Languages*. Cleveland: Case Western Reserve University Press.

Haspelmath, Martin. 1999. External Possession in a European Areal Perspective. ed. by Payne, Doris L. & Barshi, Immanuel, 109–135. *External Possession*. Amsterdam/Philadelphia: Benjamins.

Haudry, Jean. 1977. *L'emploi des cas en védique: Introduction à l'étude des cas en indoeuropéen*. Lyon: Editions l'Hermès.

Havers, Wilhelm. 1911. *Untersuchungen zur Kasussyntax der indogermanischen Sprachen*. Strassburg, Trübner.

Heine, Bernd. 1997. *Possession: Cognitive Sources, Forces and Grammaticalization*. Cambridge: Cambridge University Press.

Jacquinod, Bernard. 1989. *Le double accusatif en grec d'Homère à la fin du V siècle avant J.-C.* Louvaine-la-Neuve: Peeters.

Kittilä, Seppo & Fernando Zúñiga. 2010. Introduction: Benefaction and Malefaction from a Cross-Linguistic Perspective. *Benefactives and Malefactives: Typological Perspectives and Case Studies,* ed. by Seppo Kittilä & Fernando Zúñiga, 1–28. Amsterdam/Philadelphia: Benjamins.

Kockelman, Paul. 2009. Inalienable Possession as Grammatical Category and Discourse Pattern. *Studies in Language* 33(1): 25–68.

König, Ekkehard & Martin Haspelmath. 1997. Les constructions à possesseur externe dans les langues d'Europe. *Actence et valence dans les langues d'Europe*, ed. by J. Feuillet, 525–606. Berlin: Mouton – de Gruyter.

Luraghi, Silvia. 1986. Der semantische und funktionelle Bau des althethitischen Kasussystems. *Zeitschrift fur vergleichende Sprachforschung* 99: 23–42.

Luraghi, Silvia. 1990. The Structure and the Development of Possessive Noun Phrases in Hittite. *Historical Linguistics 1987* ed. by Henning Andersen & Konrad Koerner, 309–324. Amsterdam/Philadelphia: Benjamins.

Luraghi, Silvia. 1993. La modificazione nominale in anatolico. *Archivio Glottologico Italiano* 78(2): 144–166.

Luraghi, Silvia. 1994. Suffix Copying and Related Phenomena: A Prototype Approach. *Linguistics* 32: 1095–1108.

Luraghi, Silvia. 1997. *Hittite*. München/New Castle: Lincom Europa.

Luraghi, Silvia. 2008. Possessive Constructions in Anatolian, Hurrian, Urartean, and Armenian as Evidence for Language Contact. *Anatolian Interfaces*, ed by Billie J. Collins, Mary R. Bachvarova & I. C. Rutherford, 147–155. Oakville, CT: Oxbow Press.

Luraghi, Silvia. 2016. The Dative of Agent in Indo-European Languages. *STUF – Language typology and Universals* (Special Issue on *Non-Central Usages of Datives*) 69(1): 15–47.

Luraghi, Silvia & Chiara Zanchi. 2018. Double Accusative Constructions and Ditransitives in Ancient Greek. *Ditransitive Constructions in a Cross-Linguistic Perspective*, ed. by A. Korn, C. Jahani & A. Malchukov, 13–35. Wiesbaden: Reichert.

Malchukov, Andrej, Martin Haspelmath & Bernard Comrie. 2010. Ditransitive Constructions: a Typological Overview. *Studies in Ditransitive Constructions: A Comparative Handbook*, ed. by Andrej Malchukov, Martin Haspelmath & Bernard Comrie, 1–64. Berlin: Mouton de Gruyter.

Payne, Doris L. & Immanuel Barshi. 1999. External Possession: What, Where, How, and Why. *External Possession*, ed. by Doris L. Payne & Immanuel Barshi, 3–32. Amsterdam: Benjamins.

Roldán, Mercedes. 1972. Concerning Spanish Datives and Possessives. *Language Sciences* 21: 27–32.

Rosén, Haim. 1959. Die Ausdrucksform für 'veräusserlichen' und 'unveräusserlichen' im Frühgriechischen. *Lingua* 8: 264–295.

Rosen, Haim. 1989. A Marginal Note on Sanskrit Case Syntax. *Hanjamana*, ed. by Subhadra Kumar Sen, 33–39. Calcutta: University of Calcutta Press.

Starke, Frank. 1977. *Die Funktionen der dimensionalen Kasus und Adverbien im Althethitischen*. Wiesbaden: Harrassowitz.

Viti, Carlotta. 2004. Funzioni semantiche e pragmatiche nelle strategie di possesso dell'antico indiano. *Archivio Glottologico Italiano* 89: 41–83.

Vogt, Hans. 1932. Les groups nominaux en arménien et géorgien ancien. *Norsk Tidsskrift for Sprogvidenskap* 5: 57–81.

Wiedmer, Paul. 2014. Partitivity and Part-Whole Constructions in Indo-European. Paper presented at the workshop *Partitivity in Romance and Beyond*, Zurich, 11–13 December 2014.

How to Identify Cognates in Syntax? Taking Watkins' Legacy One Step Further

Jóhanna Barðdal and Thórhallur Eythórsson

Abstract

As a reaction to three different proposals on how to reconstruct basic word order for Proto-Indo-European, Watkins and his contemporaries in the 1970s succeeded in aborting any attempt at reconstructing syntax for a long time to come. As a consequence, syntactic reconstruction has generally been abandoned, regarded as a doomed enterprise by historical linguists for several different reasons, one of which is the alleged difficulty in identifying cognates in syntax. Later, Watkins (1995) proposed a research program aimed at reconstructing larger units of grammar, including syntactic structures, by means of identifying morphological flags that are parts of larger syntactic entities. As a response to this, we show how cognate argument structure constructions may be identified, through a) cognate lexical verbs, b) cognate case frames, c) cognate predicate structure and d) cognate case morphology. We then propose to advance Watkins' program, by identifying cognate argument structure constructions with the aid of non-cognate, but synonymous, lexical predicates. As a consequence, it will not only be possible to identify cognate argument structure constructions across a deeper time span, it will also be possible to carry out semantic reconstruction on the basis of lexical-semantic verb classes.

1 Introduction

While phonological, morphological and lexical reconstruction continue to thrive in historical-comparative linguistics, syntactic reconstruction has been balked at for several decades now, ever since the influential article by Watkins (1976) in which three different reconstructions of basic word order in Proto-Indo-European were weighed, measured and found wanting.* Since

---

then, different forces have made an entrance into the field of syntactic reconstruction, most notably Harris and Campbell who systematically argued for the feasibility of syntactic reconstruction (Harris & Campbell 1995, Campbell & Harris 2003), but were remonstrated against by Lightfoot and his followers (Lightfoot 2002a, 2002b; Pires & Thomason 2008). The debate on the legitimacy of syntactic reconstruction continues, although it seems that there is more resonance in the historical linguistic community with syntactic reconstruction now than ever before (cf. Gildea 1998, 2000; Mendoza 1998; Bauer 2000; Kikusawa 2002, 2003; Roberts 2007; Bowern 2008; Willis 2011; Kulikov & Lavidas 2013; Walkden 2014; Viti 2014; Smitherman 2015; Daniels 2015, 2017; Dunn et al. 2017; Danesi, Johnson & Barðdal 2017; Pooth et al. 2019; Johnson et al. 2019; *inter alia*), despite there still being strong forces in the community arguing against it (Pires & Thomason 2008; Mengden 2008; Walkden 2013; Seržant 2015).

Our goal in this article is to take Watkins' legacy, that syntax is reconstructable on the basis of morphological clues, one step further. A central question to be dealt with is how to identify cognates in syntax. We propose that within the area of argument structure constructions, cognate argument structures may be identified as such on the basis of a) cognate lexical material, b) cognate case frames, c) cognate predicate structure and d) cognate case morphology. After demonstrating in practice the viability of the proposed research program, we proceed to show how cognate argument structure constructions may be identified on the basis of non-cognate, but synonymous, lexical material. This last step in the methodology entails improving on Watkins' proposal, such that cognate argument structure constructions are not only identifiable on the basis of morphological or morphosyntactic clues, but also on the basis of semantic clues.

In Section 2 we give an outline of the problem, which has its roots in the marginal status of syntax during the Neogrammarian period, the underdeveloped syntactic theories of the structuralists, as well as a putative fundamental difference between phonology, morphology and the lexicon, on the one hand, and syntax on the other. We briefly discuss four of the five major arguments that have been presented against syntactic reconstruction, before turning to the fifth, the alleged lack of cognates in syntax, in Section 3. There we showcase how cognates in argument structure may be identified, illustrating our point

with a reconstruction of the argument structure constructions of three verbs, 'think', 'thirst' and 'answer', for Proto-Germanic. What these three verbs have in common is that they all deviate from the canonical Nom-Acc transitive proto-type, hence exhibiting a certain degree of arbitrariness in the linking between case marking and grammatical relations.

Moreover, we present a general argument against the arbitrariness requirement in syntax, and claim instead that canonical structures are reconstructable as such on the basis of common regularities in the daughter languages, exactly as for phonology, morphology and the lexicon. Our reconstructions are in part based on attested syntactic changes, that are regular in nature, indeed demonstrating the directionality that has been presumed to be absent in syntax.

2   The Problem

It has been consistently assumed in historical linguistics that syntactic reconstruction is more or less untenable (Watkins 1976; Jeffers 1976; Winter 1984; Lightfoot 1979, 2002a, 2000b, 2006; Harrison 2003; Holland 2003; Pires & Thomason 2008; von Mengden 2008). The reasons for this are five-fold:
– lack of regularity in syntactic change
– lack of arbitrariness in syntax
– lack of simple form–meaning correspondences in syntax
– lack of continuous transmission of syntactic structures during acquisition
– lack of cognate material in syntax
Let us start with the lack of regularity in syntactic change, since regularity has been a key concept for deciding on the potential directionality of syntactic change. This idea of lack of regularity is based on the notion that while sound change is regular and can be captured by sound laws, syntactic change, in contrast, is not; there are no syntactic "laws" that can be used in reconstruction to determine the regularity of change (cf. Miranda 1976; Lightfoot 1979, 2002a; Campbell & Mithun 1980), aiding in the decision of which alternant of a correspondence set provides basis for the reconstruction.

This first criticism of syntactic reconstruction is based on two misconceptions. The first misconception is that the sound laws are regular and apply without exception. The second one is that there is no regularity, and hence no directionality, in syntactic change. The truth of the matter is first that sound laws are only regular by definition (Hoenigswald 1978), while semi-regular and less regular sound changes are as such systematically excluded from the notion of a sound law. Thus, not all phonological changes can be reconstructed on the basis of sound laws. Second, contrary to widespread beliefs, regularity

can indeed be found in syntactic change, although as Willis (2011) points out, this regularity is *local* and not necessarily *universal.* In other words, by carefully examining the data, regularities may be discerned, although this clearly has to be done on the basis of each data set and that data set alone (cf. also Harris & Campbell 1985; Campbell & Harris 2003; Barðdal & Eythórsson 2012a; Dunn et al. 2017).

Turning to the second and the third criticisms listed above, the general conception within the historical linguistic community is that syntax lacks the form–meaning correspondences which are needed to provide a basis for a secure reconstruction, and hence that the arbitrariness requirement cannot be fulfilled in syntax.

Exactly as the first criticism, the second one is also based on two misconceptions: first, there is in fact a great deal of arbitrariness in syntax, contra the received opinion, and second, in any case, the requirement of arbitrariness is simply not needed in syntax. Harrison (2003) argues that the requirement of arbitrariness is first and foremost relevant when the goal of the reconstruction is to establish genetic relatedness. Since syntactic reconstruction is usually carried out after genetic relatedness has been established, the arbitrariness requirement is superfluous in syntax. In addition, as we discuss below, there is a substantial amount of arbitrariness in syntax, so the requirement can in part be fulfilled in any case.

The third criticism, that syntax does not consist of simple form–meaning correspondences, has its roots in the traditionalist/structuralist view of sentence meaning, namely that the meaning of a sentence is derived from the meaning of the lexical items instantiating it (cf. Klein 2010). On such a view, sentences do not consist of form–meaning correspondences, but are instead combinations of words and phrases, according to specific rules.

This third criticism does not apply on a constructional approach to language where larger syntactic units than words are regarded as form–meaning correspondences (Fillmore et al. 1988; Goldberg 1995; Croft 2001; *inter alia*). On a constructional approach, not only words are regarded as signs, in the Saussurian sense, but also larger constructions. In fact, constructions are assumed to range on a scale from lexically filled to schematic constructions, as well as ranging on a scale from the atomic to the bound, and from the bound to the combinatoric (cf. references in Croft & Cruse 2004). On a constructional approach, therefore, syntactic entities also count as form–meaning correspondences, and are as such legitimate objects of the Comparative Method (Eythórsson & Barðdal 2011, 2016; Barðdal & Eythórsson 2012a, 2012b; Barðdal 2013, 2014; Daniels 2015; Danesi, Johnson & Barðdal 2017; Johnson et al. 2019; Vázquez-González & Barðdal 2019).

The fourth criticism, concentrating on the issue of the alleged lack of continuous transmission of syntactic structures during acquisition, was in particular put forward and emphasized by Lightfoot (1979, *inter alia*). The idea here is that words are inherited from one generation to the other, while clauses are not, since the speaker does not inherit the grammar of his/her language, but derives it on the basis of the input.

This fourth criticism of syntactic reconstruction is based on quite a simplistic view of lexical items, as being somehow less abstract than clauses (Barðdal & Eythórsson 2012a). In fact words are complex form–meaning correspondences and are as such abstract entities (cf. Adger 2003; Tomasello 2003; Goldberg 2006: 69ff). This means that there is, in essence, no qualitative difference between the transmission of lexical items and larger structural units. We realize that this goes against a widely held view, but it follows naturally from the explicit tenets of Construction Grammar, where larger units than lexical items are regarded as form–meaning correspondences exactly like words. As such, larger schematic units can be inherited from one generation to the next exactly like words (Eythórsson & Barðdal 2011; Barðdal & Eythórsson 2012a, 2012b).

We have dealt extensively with these first four arguments against syntactic reconstruction elsewhere; hence, in the following, we focus primarily on the last problem, namely that of how to identify cognates in syntax. Nevertheless, regularity in syntactic change, and arbitrariness, will also figure in the discussion and argumentation below.

It is generally assumed in the historical linguistic community that identifying cognates in syntax is a hopeless venture. As is well known, during the early 1970s, three historical linguists launched three different reconstructions of the basic word order of Proto-Indo-European. Lehmann (1974), following Delbrück (1878, *inter alia*), claimed that Proto-Indo-European was an sov language, Friedrich (1975) claimed that it was an svo language, and Miller (1975) that it could have been an svo, sov or vso language. Lehmann founded his claims on the typological work of Greenberg, and argued for his position on the basis of typological correlations between basic word order and the word order within the noun phrase, the prepositional phrase and the adjectival phrase. Friedrich's claim is grounded in a frequency count of different word orders in early texts, in particular Homer; by means of which he found that svo dominates over sov. Finally, Miller founded his claims on the later development of the word order in the Indo-European languages, where all three word orders are attested.

These three different approaches to one and the same phenomenon, basic word order in Indo-European, attracted much antagonistic attention from

contemporary historical-comparative linguists at the time. In particular, Watkins (1976) put forward severe criticism of this whole enterprise which he regarded as a major fiasco. Although justified, Watkins' criticism, in effect, annihilated any further attempts at reconstructing syntax for decades. As a consequence, studies on syntactic reconstruction came to a halt and did not gain ground again for a long time to come.

In retrospect, seen from a modern perspective, the work of the troika from the early 1970s, Lehmann, Friedrich and Miller, may of course have been a fumble in the dark. However, given the fact that theoretical syntax was still an underdeveloped field at the time, and historical syntax even more so, any attempts at syntactic reconstruction were bound to be lightweight and imperfect. Thus, the failure of these early attempts reflects shortcomings of their models and not of the enterprise in and of itself.

Watkins himself did not suggest a systematic program for how to reconstruct syntax. Instead, he emphasized that syntactic reconstruction should be based on archaic expressions containing frozen syntax, deviations from productive patterns, and any anomalies in the language that cannot be explained synchronically. He also pointed out that in order to be able to carry out syntactic reconstruction, one has to examine the data carefully, compare linguistic units used to express similar content across the daughter languages, and in general compare cognate text traditions based on oral transmission of inherited cultural and linguistic material. In other words, Watkins did not directly address the cognacy problem.

In the same year Jeffers (1976) also problematized the issue of reconstructing syntax and claimed that there is no finite set of sentences which can be used as input for correspondence sets in syntax. He argued that one of the main problems with reconstructing syntax is that syntactic change takes place through pattern replacement, but does not necessarily involve small changes in inherited patterns, which is the kind of change needed to identify inherited patterns across daughter languages. In other words, Jeffers' claim means in essence that there can be no cognates in syntax in the same sense as in phonology, morphology and the lexicon. This position has been the prevailing view on syntactic reconstruction for decades. Below, we present examples from syntax which falsify this claim.

Watkins' main contribution to the debate on syntactic reconstruction, initiated in his early work and laid out in more detail in his 1995 book, *How to Kill a Dragon*, was that morphological clues are instrumental for identifying cognates, including cognate syntactic material (cf. also Watkins 1964; Fox 1995; Gildea 1998; Kikusawa 2003; Harris 2008). Watkins' own work on poetic formulae consistently identifies layers of cognate collocations and prefabs across

the Indo-European traditions, showing how fragments of earlier syntax can be determined and, hence, reconstructed.

Taking Watkins' method at face value means that essentially all morphosyntax constitutes a potential input for correspondence sets, and therefore provides a basis for comparative reconstruction. This is by no means an insignificant proportion of grammar: *all morphosyntax*. In addition, following Watkins' method, collocations and prefabs may provide information about word order and clause structure of earlier stages, which in turn means that such abstract units can be detected and reconstructed (cf. Comrie 1980). Thus, despite the pessimistic tone of Watkins' (1976) article, there are more possibilities inherent in his approach than he and many of his epigones may have realized. Prospects for syntactic reconstruction are therefore not as gloomy as commonly assumed. In Barðdal et al. (2013) we showed how morphosyntactic reconstruction can be expanded into the domain of information structure, and below we will argue that it can be extended into the domain of semantics as well (cf. also Barðdal 2007, Barðdal et al. 2012; Johnson et al. 2019).

Finally, let us consider the question of why syntactic reconstruction is important at all. Syntactic reconstruction is not simply a hobby of some armchair linguists who enjoy playing with historical data; it is a fundamental part of historical linguistics, as it involves putting forward grounded hypotheses on prestages of languages, and hence aims at identifying how language change comes about (cf. Ferraresi & Goldbach 2008). Syntactic reconstruction is thus a way of concretizing and "formalizing" analyses of language change. As such, syntactic reconstruction may provide important insights into the development of specific linguistic structures. We now turn to the issue of how to identify cognates in syntax.

3 Identifying Cognates in Syntax

We first discuss cognate recognition in syntax on the basis of cognate lexical material (Section 3.1), second, on the basis of cognate structure, including argument and predicate structure (Section 3.2), and finally, we show how Watkins' program can be taken one step further, namely through cognate recognition in syntax on the basis of synonymous lexical material (Section 3.3).

3.1 *Identification on the Basis of Cognate Lexical Material*
In order to extend Watkins' method to reconstruct on the basis of morphosyntax, let us now delve into the realm of argument structure, since a substantial part of our syntactic work has been focused on that domain of grammar.

TABLE 5.1   Lexical correspondence sets for Germanic 'think'

	FORM	MEANING	RECONSTRUCTED FORM
Gothic	*þugkjan*	'think, seem'	
Old High German	*thunkian*	'think, seem'	
Old English	*þyncan*	'think, seem'	**þunkjan-*
Old Saxon	*thunkian*	'think, seem'	
Old Norse-Icelandic	*þykkja*	'think, seem'	

Argument structure is understood here as the arguments selected by a predicate and their relative order. Included in this are different case frames inherent in different argument structure constructions, like Nom-Acc, Nom-Dat, Nom-Gen, etc. As valency is a major ingredient in syntax, the ability to reconstruct argument structure is, we believe, a very significant contribution to a research program aiming at syntactic reconstruction.

Let us start with an investigation of a predicate selecting for a non-canonical case frame as a part of its argument structure. For instance, the oldest verb meaning 'think, seem' in the Germanic languages has been reconstructed as **þunkjan* by historical linguists (see, for example, Kluge 2002), on the basis of the forms and the meaning in the earliest daughters. Table 5.1 gives the lexical correspondence set, on the basis of which the reconstructed form has been posited. As is evident from the table, the forms are clearly related and the meaning is the same.

Consider now some examples of the argument structure of this verb in the earliest layers:

Gothic
(1a) *þugkeiþ*   **im**   auk   ei …
      thinks.3SG   them.DAT   because   that
      'for they think that …' (Mt 6.7)

Old High German
(1b) samomichel   uuunder   mag   **temo**   *dunchen*, der …
      same.great   wonder.NOM   may.3SG him.DAT   seem.INF   who.NOM
      'He will think it an equally great wonder, who …' (Notker 1,283,9)

Old English

(1c) Ne *þynceð* **me** gerysne þæt we rondas beren eft to earde
not thinks.3SG me.DAT appropriate that we shields bear back to earth
'I do not find it appropriate that we bear our shields back home'

(Beowulf 2653)

Old Saxon

(1d) than *thunkid* **imu,** that he sie gerno forð lêstien uuillie
then seems.3SG him.DAT that he it gladly forward do.INF wishes
'Then he thinks that he will gladly wish to do it in the future'

(Heliand 2496–2501)

Old Norse-Icelandic

(1e) **oss** *þykir* eigi verr að þú sért lítt heil
us.DAT seems.3SG not worse that you are little healthy
'we don't find it worse that you are not well' (Fóstbræðra saga, Ch. 10)

As these examples show, the first argument of the argument structure, the subject, is always in the dative case (for the subject behavior of non-nominative subjects in early Germanic, see Rögnvaldsson 1991, 1995; Allen 1995; Barðdal 2000; Barðdal & Eythórsson 2003, 2012b; and Eythórsson & Barðdal 2005). The morphological case markers in Germanic are also cognate; no innovative morphology is found in the case paradigm. It is certainly true that dative and instrumental have merged at some point in proto-history (Luraghi 1987; Barðdal & Kulikov 2009), reflexes of which are documented in early West-Germanic (see, for example, Krahe 1969a). This, however, is not relevant here, as the dative with 'think' is hardly of instrumental origin. Table 5.2 gives an overview of the case marking of the subject of 'think' in both the early and the modern stages of the Germanic languages.

Observe that in the oldest languages 'think' consistently occurs with a dative subject. To be sure, accusative is also attested in Old High German, but apart from that it is only attested in the later stages. Nominative subjects with 'think' are first attested in Middle High German, and subsequently in Modern German, as well as in the Modern Germanic languages that have lost case marking. The data presented in Table 5.2 are in accordance with the well-known tendency that oblique subjects change into nominative in the course of time (Jespersen 1927; Seefranz-Montag 1982; Eythórsson 2000, 2002; Barðdal 2009, 2011a). Notice also that Old Swedish, Middle High German, Middle English and Middle Dutch, which are more or less contemporaneous with Old Norse-Icelandic, have all innovated from the original case frame. In that sense they have gone further in their development than Icelandic.

TABLE 5.2   Predicate-specific correspondence sets for the argument structure of Germanic 'think'

	ALT 1	**ALT 2**	**ALT 3**
Gothic	DAT-'thinks'		
Old English	DAT-'thinks'		
Old Saxon	DAT-'thinks'		
Old High German	DAT-'thinks'	ACC-'thinks'	
Old Norse-Icelandic	DAT-'thinks'		
Old Swedish	DAT-'thinks'	ACC-'thinks'	
Middle English	DAT-'thinks'	ACC-'thinks'	
Middle Dutch		ACC-'thinks'	
Middle High German	DAT-'thinks'	ACC-'thinks'	NOM-'think'
Modern Icelandic	DAT-'thinks'		
Modern Faroese	DAT-'thinks'		
Modern Dutch	DAT-'thinks'	ACC-'thinks'	
Modern High German	DAT-'thinks'	ACC-'thinks'	NOM-'think'
Modern Swedish			NOM-'think'
Modern English			NOM-'think'

We propose that the argument structure of 'think' can be reconstructed for Proto-Germanic, containing a dative subject, on two grounds:

– The earliest representatives of Germanic have a dative subject
– Accusative and nominative subjects are an innovation, attested first in the later layers, also in accordance with known developmental paths of oblique subjects

This proposal amounts to claiming not only that the predicate itself is cognate, but also its argument structure. We have identified the argument structure as a cognate argument structure, inherited from a common proto-stage, on the basis of three factors:

– The lexical predicate (including both its form and meaning) is cognate
– The case frame itself is cognate
– The morphological case markers are cognate

This is an example of how an argument structure construction can be identified as cognate in the preserved material. Let us now reconstruct this argument structure construction for Proto-Germanic. For that purpose we employ

*	*lxm*		
	FORM	< *þunkjan* >	
	SYN	ARG-ST <NP-DAT$_i$>	
	SEM	FRAMES	regard-fr COGNIZER i

FIGURE 5.1 A reconstruction of the argument structure of 'think' in Proto-Germanic

the formalism of Construction Grammar (Kay & Fillmore 1999; Michaelis & Ruppenhofer 2001; Boas 2003; Fried & Östman 2005; Michaelis 2009, 2012; Sag 2012; Fried 2015). Observe that the proposed reconstruction is only a partial reconstruction, based only on information about the subject argument.

The reconstruction, found in Figure 5.1, consists of three fields, a FORM field, a SYN field and a SEM field. The FORM field specifies the reconstructed form of the verb 'think' as *þunkjan*; the SYN field gives the case marking of the dative argument; the semantics of this verb-specific construction is given in terms of semantic frames (see Framenet[1]), in this case the Regard frame, where the subject is a COGNIZER, marked with an *i*, coindexed with the dative NP of the SYN field (for Frame Semantics, see Fillmore 1982, 1985; Petruck 1996; Fillmore & Baker 2009; *inter alia*). This is how verb-specific argument structure constructions may be reconstructed for earlier unattested proto-stages.

In addition to the reconstruction of verb-specific argument structure constructions as in Figure 5.1, it is also possible to reconstruct constructions at a higher, more schematic, level than the one proposed there. This may involve verb-subclass and verb-class-specific argument structure constructions that exist irrespective of the lexical items instantiating them. We refer the interested reader to Barðdal & Smitherman (2013) and Vazquez Gonzalez & Barðdal (2019) for a detailed exposition of the methodology and formalism. It follows from this that argument structure constructions, including their case frames, do not only exist at the substantive level, but also at different levels of schematicity.

Now it might be objected that dative subjects are ubiquitous with experiencer predicates cross-linguistically, and hence, dative subjects might be expected

---

1 *The Framenet Project*, available at framenet.icsi.berkeley.edu.

to arise independently in languages because of common cognitive and conceptual constraints, and therefore are not necessarily inherited. However, one must also take into account the argument of economy, or Occam's razor, which, where it is possible, privileges inheritance (no changes) over innovation (one or more changes). And of course, if the typological ubiquity of dative experiencers could be invoked as a motivation for dative subjects in the modern languages, then it should be equally valid as a motivation for a dative subject in the proto case frame. As such, there are three major arguments for assuming an inheritance here, namely the three reasons stated in the bulleted list above: cognate lexical predicates, cognate case frames and cognate case markers. Moreover, dative subjects are also a clear deviation from the canonical argument structure with a nominative subject, found with the overwhelming majority of predicates, including experiencer predicates (cf. Jónsson 2003; Nichols 2008). Also, given the large pool of potential lexical predicates with experiencer meaning, exactly which ones get assigned non-canonical subject case marking and which do not appears idiosyncratic. As such, dative subjects definitely exhibit a certain degree of arbitrariness.

We continue with a discussion of the logical basis for the arbitrariness requirement for reconstruction in Section 3.2 below. But first, let us consider another predicate with a non-canonical subject case marking other than the dative, namely the verb 'thirst' which selects for an accusative subject in the early and archaic Germanic languages. This is important because accusative subjects are far less common in the world's languages than dative subjects (see, for instance, Barðdal 2009 on Old and Modern Icelandic), involving an even higher degree of arbitrariness than datives. The lexical correspondence set for 'thirst' is given in Table 5.3, on the basis of which a verb meaning 'thirst' has been reconstructed for Proto-Germanic (cf. Kroonen 2013: 553).

TABLE 5.3   Lexical correspondence sets for the Germanic verb 'thirst'

	FORM	MEANING	RECONSTRUCTED FORM
Gothic	*þaursjan*	'thirst'	
Old High German	*dursten*	'thirst'	
Old English	*þyrstan*	'thirst'	*$þurs(t)$-
Old Saxon	*thurstian*	'thirst'	
Old Norse-Icelandic	*þyrsta*	'thirst'	

The examples below document the argument structure of 'thirst' in the earliest layers of Germanic:

Gothic

(2a) þana   galaubjandan du   mis   ni   *þaurseiþ*   hvanhun
     the.one.ACC believing   in   me   not   thirsts.3SG   ever
     'he that believes in me shall never thirst' (John, 6:35)

Old High German

(2b) thaz   **mih**   ni   *thurste*
     so.that me.ACC   not   thirsts.3SG
     'So that I do not thirst' (Tatian 87,5)

Old English

(2c) þeah   **þæt**   **folc**   *þyrste*   þære   lare
     yet   that.ACC   folk.ACC   thirsts.3SG   their teaching.GEN
     'Yet those people thirst for the teaching' (Alfred Pastoral Care 31,6)

Old Saxon

(2d) quað   that   **ina**   *thurstidi*
     said   that   him.ACC   thirsted.3SG
     'He said he was thirsty' (Heliand 5640b–5642a)

Old Norse-Icelandic

(2e) *Þyrstir*   **mig**   nú   því að eg em óvanari   erfiðinu   en   þú.
     thirsts.3SG   me.ACC   now for at I   am less.used.to hard.work than you
     'I am thirsty now because I'm less used to hard work than you.'
                                    (Þorsteins Þáttur Stangarhöggs)

As these examples show, the verb 'thirst' occurs systematically with an accusative subject in the earliest Germanic languages. There is some variation, however, at later stages, as shown in Table 5.4.

Observe that all the oldest languages display an accusative, which is preserved into later medieval stages, and surviving into modern times in Icelandic and German. The only language that shows variation in its earliest texts is Old English, with 'thirst' being documented with a dative and a nominative subject, in addition to the accusative. The dative survives into Middle English and is also documented in Middle Dutch. Except for in Old English, the nominative is

TABLE 5.4   Predicate-specific correspondence sets for the argument structure of Germanic
'thirst'

	Alt. 1	Alt. 2	Alt. 3
Gothic	ACC-'thirsts'		
Old High German	ACC-'thirsts'		
Old Saxon	ACC-'thirsts'		
Old English	ACC-'thirsts'	DAT-'thirsts'	NOM-'thirst'
Old Norse-Icelandic	ACC-'thirsts'		
Middle High German	ACC-'thirsts'		
Middle English	ACC-'thirsts'	DAT-'thirsts'	
Middle Dutch	ACC-'thirsts'	DAT-'thirsts'	
Old Swedish	ACC-'thirsts'		
Modern Icelandic	ACC-'thirsts'		
Modern High German	ACC-'thirsts'		NOM-'thirst'
Modern Faroese			NOM-'thirst'
Modern Swedish			NOM-'thirst'
Modern English			NOM-'thirst'
Modern Dutch			NOM-'thirst'

only documented in the modern stages of the Germanic languages. The reason
that the nominative is documented already in the earliest stage of English with
'thirst' is most likely due to the early mergers of the morphological case mark-
ers in that language, as opposed to in the other Germanic languages where the
case system did not collapse until later (Allen 1995; Falk 1997; Barðdal 2009).

To sum up, Table 5.4 outlines a development from accusative marking on the
subject to nominative marking. This in turn suggests that it is in fact the accu-
sative subject that is original with this verb. Again, we have identified cognate
argument structure constructions with the verb 'thirst' in Germanic. Exactly
as with 'think', this can be done on two grounds: a) the earliest representatives
of Germanic have an accusative subject, and b) dative and nominative sub-
jects are an innovation, attested first in the later layers, also in accordance with
known developmental paths of oblique subjects. As with the verb 'think', there
are three main arguments for assuming that the accusative subject construc-
tion is inherited from a common proto-stage:

*	*lxm*		
	FORM	*< þurs(t)jan >*	
	SYN	ARG-ST < NP-ACCᵢ >	
	SEM	FRAMES	need-for-intake fr NEEDER i

FIGURE 5.2   A reconstruction of the argument structure of 'thirst' in Proto-Germanic

– The lexical predicate (including both its form and meaning) is cognate
– The case frame itself is cognate
– The morphological case markers are cognate

On this basis, the argument structure construction of 'thirst' can be reconstructed, with the use of the Construction Grammar formalism, as in Figure 5.2, which has a parallel structure to that of 'think' in Figure 5.1.

The semantic frame we propose for 'thirst' is a subframe of Framenet's biological_urge frame, or more specifically the Need_for_intake_of_nourishment frame. This frame has already been suggested for 'hunger' (Barðdal & Eythórsson 2012b), on the basis of an earlier analysis involving the frame for 'eat' (Croft 2009).

As a final example in our proposed research program, let us consider 'answer' in Germanic which is lexically manifested by two different etymons, namely those reconstructed as *(and)swaran* and *andwurdjan*. Table 5.5 gives our correspondence sets and lexical reconstruction for *(and)swaran* and Table 5.6 for *andwurdjan*.

TABLE 5.5   Lexical correspondence set and reconstruction of Germanic *(and)swaran*

	FORM	MEANING	RECONSTRUCTED FORM
Old English	*andswarian*	'answer'	
Old Frisian	*andswara*	'answer'	
Old Norse-Icelandic	*svara*	'answer'	*(and)swaran*
Old Swedish	*svara*	'answer'	
Old Danish	*swarœ*	'answer'	

TABLE 5.6   Lexical correspondence set and reconstruction of Germanic *andwurdjan

	FORM	MEANING	RECONSTRUCTED FORM
Gothic	*andwaurdjan*	'answer'	
Old High German	*antwurten*	'answer'	
Old Frisian	*andwerda*	'answer'	*andwurdjan*
Old Saxon	*andwurdian*	'answer'	
Old English	*andwyrdan*	'answer'	
Middle Dutch	*antwerden*	'answer'	

Table 5.5 suggests that the distribution of the etymon *(and)swaran* is confined to North-Germanic and Ingvaeonic, while Table 5.6 shows that the etymon *andwurdjan* is found in Gothic and West-Germanic. Note that there also exists a Gothic verb *swaran* with a different meaning, 'swear (an oath)' cf. Kroonen (2013: 496). In the other languages, this meaning is expressed with a verb containing a *-ja* suffix, namely *swarjan.* Brugmann (1913) assumes a semantic development 'swear (an oath)' from 'answer before a court of justice'. This account presupposes that the meaning 'answer' is the original meaning, while 'swear (an oath)' is derived. This is further confirmed by the existence of related forms in other Indo-European languages, like Old Indic *svara-* 'sound', Oscan *sverrunei* 'spokesperson' and Old Church Slavonic *svara* 'quarrel' (de Vries 1962: 568).

Given the existence of the verb *swaran* in all three branches of Germanic, it is reasonable to assume that it also existed in Proto-Germanic. But, it is also clear from the data that this verb had a competitor, namely *andwurdjan*, with the same meaning as *(and)swaran* 'answer'. This *andwurdjan* consists of two components: a) the verb *wurdjan* (Goth. *waurdjan*, OE *wyrdan*, ON-I *orða*, etc.) meaning 'speak, put into words', derived from the noun *wurdan* 'word' (Goth. *waurd*, OE, OS, OFr. *word*, OHG *wort*, ON-I *orð*, etc.), and b) the prefix *and-*. With the addition of this prefix, *andwurdjan* came to mean 'speak against, oppose', and from there the meaning got bleached into 'answer' (cf. Icelandic *andmæla* 'oppose', composed of *and-* and *mæla* 'speak' from the noun *mál* 'speech, language'). The preserved material therefore suggests that *andwurdjan* replaced *(and)swaran* in Gothic and West-Germanic, and that this must have happened during prehistoric times. We assume, further, that the basic verb *wurdjan* selected for the Nom-Acc case frame, in the meaning 'put into words', as it did in Old English and still does in Icelandic with the inherited *orða* 'put into words'.

Let us now investigate the argument structure of *(*and*)swaran* and *and-wurdjan* in Germanic. Both verbs select for a nominative subject and a dative object in all the early Germanic languages, and they still do in the languages that have preserved morphological case marking. In the languages that have lost case morphology accusative and dative have merged into a common object (oblique) form, which synchronically amounts to an accusative. The examples in (3) are with *(*and*)swaran* and the ones in (4) are with *andwurdjan*.[2]

Old English
(3a) **him**   se   yldesta   *andswarode* (Beo. Th. 522)
     him.DAT the   oldest   answered
     'The oldest answered him'

Old Norse-Icelandic
(3b) vér   *svöruðum* **þér**   ok   sögðum … (Stj. 219, 2)
     we   answered   you.DAT   and   said
     'We answered you and said …'

Old Danish (ca. 1425)
(3c) **thæm**   scal   han   al   ene   *swore* (Skråer I.9 §9)
     them.DAT shall he   all   one   answer
     'He shall answer them all alone'

Old Swedish
(3d) budit   *swardhe*   **hanom**   swa (RK 2: 2545)
     messenger answered   him.DAT   such
     'The messenger answered him such'

Modern Icelandic
(3e) og   hann *svaraði*   **henni**   til   baka
     and   he   answered   her.DAT   to   back
     'and he answered her back'

---

2   The Gothic example (4a) is a translation from Greek, and the Old High German (4b) and
    the Old English (4e) ones are translations from Latin. In all these cases the Greek and Latin
    originals also have a dative with the verb 'answer' (see Section 3.3 below).

Modern Faroese

(3f) at   fáa   *svaraõ*   **hvørjum**   sítt
      to   get   answered   each.DAT   their.own
      'to be able to answer each and every one about their own'

Modern Danish

(3g) Jeg   *svarede*   **ham**   naturligvis   ikke.
      I   answered   him.ACC   of.course   not
      'Of course I didn't answer him'

Modern Swedish

(3h) Kan dö lycklig nu   för   osciiii *svarade* **mig**
      can die happy now   because   Osciiii answered   me.ACC
      gånger   två   idag
      times   two   to.day
      'Can die happy now since Osciiii answered me twice today'

Modern English

(3i) The girl *answered* **him**.

Gothic

(4a) þu   hwas   is,   ei   *andwaurdjais*   **guda**?
      you   who   are   that   reply.OPT   God.DAT
      'Who are you replying to (OR: against) God?' (Rom 9:20)

Old High German

(4b) Tho   *antvvurtita*   **imo**   Philippus
      then   answered   him.DAT   Philip.NOM
      'Then Philip answered him'   (Tatian, Ev. Harm., 80.3)

Old Frisian

(4c) So   ach   **him**   thi   other   *andwerdia*
      then   shall   him.DAT   the.NOM other.NOM   answer
      'Then the other must answer him' (The Skeltana Riucht XXV)

Old Saxon

(4d) **Imu**   *anduuordidun*   frôlîco   is   friund   angegin
      him.DAT   answered   cheerfully   his   friends   again
      'His friends answered him cheerfully again' (Heliand 3041)

Old English

(4e) Abraham **hire** *andwerde*
Abraham her.DAT answered
'Abraham answered her' (Gen. 16.6)

The correspondence set for *(and)swaran* is given in Table 5.7 below, on the basis of which we suggest the reconstruction in Figure 5.3, where we reconstruct the original case frame as being Nom-Dat, and not Nom-Acc. This is because Nom-Dat is found in the earliest daughters, while Nom-Acc is not found until later stages of North and West Germanic and then only in languages where the morphological accusative and dative have merged, with subsequent functional merger of the two case uses.

TABLE 5.7   Predicate-specific correspondence set for the argument structure of Germanic *(and)swaran*

	ALT 1	ALT 2
Old English	NOM-DAT	
Old Norse-Icelandic	NOM-DAT	
Modern Icelandic	NOM-DAT	
Modern Faroese	NOM-DAT	
Modern English		NOM-ACC
Modern Swedish		NOM-ACC
Modern Danish		NOM-ACC
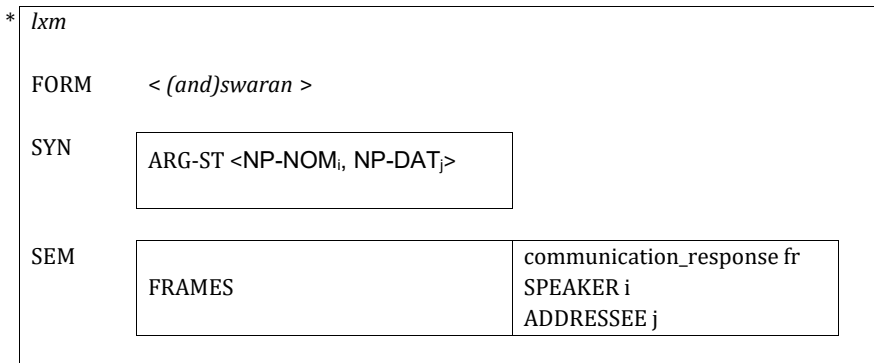
FIGURE 5.3 A reconstruction of the argument structure of 'answer' in Proto-Germanic

To recapitulate, we propose that *(*and*)*swaran* is the older verb in Germanic meaning 'answer', and that it was replaced by *andwurdjan* in East and West Germanic. This raises the question of where the Nom-Dat case frame of *andwurdjan* came from. It is obvious that since *andwurdjan* is a prefixed form of *wurdjan*, which is derived from a noun, the case frame cannot be inherited from the source. The reason is that the noun *wurdan* 'word' does not have a verbal argument structure and the argument structure of the verb *wurdjan* was most likely Nom-Acc, as discussed above.

One question that arises is whether *andwurdjan* may receive its case frame from the prefix *and-*, rather than from the synonymous *(*and*)*swaran*. This appears not to be the case. In a detailed study of the syntactic functions of prefixes in Old High German, for instance, Kuroda (2014) shows that valency is much less affected by prefixes in Old High German than in Modern German. Moreover, a case study of selected prefixed verbs in other Old Germanic languages confirms Kuroda's conclusions. For example, Go. *andbindan* 'untie' and its OE and OHG cognates, *anbindan* and *intbintan*, respectively, all select for accusative objects. Another example comes from Old Saxon, where the verb *andbītan* 'consume, partake' selects for a genitive object, illustrating beyond doubt that the dative with *andswaran* and *andwurdian* is not assigned by the prefix *and-*.

Earlier research on new verbs shows that four main tendencies are at work when new verbs are assigned case frames (Barðdal 1999, 2001, 2008, 2012):

– Case assignment by default
– Case frame borrowed with the lexical predicate it occurs with
– Case assignment inherited from a cognate source verb
– Case assignment on the basis of synonymous verbs

Default case assignment is the unmarked option and entails that a subject is assigned the nominative case and an object the accusative case, i.e. an analogical extension on the basis of the statistically most predominant pattern (cf. Barðdal 2011b). The second tendency entails that when a verb is borrowed from one language to another, it is borrowed with its case frame as well. This process has been documented in detail in a study of borrowed verbs in 15th century Icelandic (Barðdal 1999). The third tendency is typically found with verbs derived from already existing verbs, like Icelandic *aðstoða* 'assist' which is derived from its cognate *stoða* 'support' by means of the prefix *að-* 'to'. The verb *stoða* selects for Nom-Acc in Icelandic, and it seems that *aðstoða* has simply inherited this case frame from its source. Another example is the Old English verb *anbindan* 'untie' mentioned above, which has presumably inherited its case from the unprefixed *bindan* 'bind, tie' which selects for Nom-Acc in all the Germanic languages. The fourth type of case assignment is found when a new verb receives case marking on the basis of an existing synonymous verb. This third type is quite common, as has been shown in research on Icelandic

TABLE 5.8   Argument structure constructions with new verbs

Cognate verb source		Synonymous verb source	
stoða 'support'	*aðstoða* 'assist	*bjarga* 'save'	*redda* 'save'
senda 'send'	*símsenda* 'fax'	*laga* 'fix'	*ordna* 'fix'
rita 'write'	*afrita* 'copy'	*hringja* 'phone'	*bjalla* 'phone'
færa 'move'	*uppfæra* 'update'	*daga* 'dawn'	*dona* 'dawn'
þýða 'translate'	*bakþýða* 'decompile	*skera* 'cut'	*kötta* 'cut'
baka 'bake'	*afbaka* 'distort'	*binda* 'tie'	*leisa* 'lace'
skoða 'observe'	*endurskoða* 'audit'	*hreinsa* 'cleanse'	*sjæna* 'make shiny'
skipta 'divide'	*lagskipta* 'stratify'	*trufla* 'bother'	*bögga* 'bug'
setja 'put'	*gróðursetja* 'plant'	*eyða* 'delete'	*dílíta* 'delete'

(Barðdal 2001, 2008, 2012). Table 5.8 gives a few documented examples of case assignment of the latter two types (from Barðdal 2001, Appendix C), which are also most important for the present purposes.

Returning to **andwurdjan*, it is clear that default case assignment is not responsible for its case frame, as default case assignment is Nom-Acc, and not Nom-Dat (see, for instance, Barðdal 2011b on Icelandic). The second option is not relevant either since there is no reason to assume that **swaran* has been borrowed into Germanic; it is formed from a PIE root **swer-* 'speak', as all etymological dictionaries attest to, but has developed the specific meaning 'answer' within Germanic. With regard to the third option, the case frame cannot be inherited from a cognate source verb (cf. two leftmost columns in Table 5.8), since the cognate **wurdjan* must be reconstructed with a Nom-Acc case frame. The most likely source of the case frame of **andwurdjan* is thus that it comes from its synonymous verb **(and)swaran* (cf. rightmost columns in Table 5.8). In other words, when **andwurdjan* changed its meaning from 'put into words, speak' to 'answer' it also took over the case frame of the verb meaning 'answer', namely **(and)swaran*.

The goal of this section has been to illustrate in detail how syntactic structures like, for instance, argument structure constructions may be reconstructed on the basis of cognate lexical verbs across daughter languages (cf. also Harris 2008; Barðdal & Eythórsson 2012a, 2012b; Barðdal & Smitherman 2013; Barðdal et al. 2013; Barðdal 2013, 2014; Eythórsson & Barðdal 2016; Danesi, Johnson & Barðdal 2017; Dunn et al. 2017; Johnson et al. 2019; Pooth et al. 2019; Vasques, Gonzales & Barðdal 2019). The examples above, with 'think', 'thirst' and 'answer', deviate from the default Nom-Acc assignment in Germanic, and

therefore involve at least some degree of arbitrariness in the mapping between the lexical verb and its case frame.

It is commonly assumed that all linguistic units which are not motivated synchronically must have an historical explanation. It is less widely recognized that this also applies to argument structure constructions. One could, however, argue that the dative subject case of 'think' is motivated synchronically by the fact that a subset of experiencer predicates exhibit that case frame. This relation, however, is not systematic, since experiencer predicates in Germanic may occur in the nominative and the accusative, as well as the dative. One could also argue that the dative object of 'answer' is motivated by the animacy of the object referent, as dative objects are often animate.

However, in a study of the productivity of different argument structure constructions, Barðdal (2008: 75) shows for Modern Icelandic that only 45% of dative objects are animate, while corresponding figures for accusative and genitive are 26% and 21%, respectively. While languages like Modern German have more or less eliminated arbitrary case assignment from the language, with the consequence that dative as an object case is primarily used with animate objects (Wegener 1985; Primus 2012), this is not true for Icelandic, where arbitrary case assignment has been preserved to a much greater degree. Given that animate objects may be assigned dative, accusative or genitive, it appears that the exact linking between a specific lexical predicate with an animate object and its particular case frame is in part arbitrary. The verb *andwurdjan* could therefore have been assigned the Nom-Acc or the Nom-Gen case frame, as well as the Nom-Dat case frame, if animacy was the decisive factor.

In this context, let us contemplate the issue of whether there is a difference between the concepts of "syntactic reconstruction" and what Klimov (1977) labels "diachronic interpretation". In other words, is it only possible to reconstruct on the basis of arbitrary correspondence sets? For instance, if all the early and archaic Indo-European daughter languages have a transitive construction with a nominative subject and an accusative object, is it then not legitimate to reconstruct such an argument structure for the proto-language, even though there certainly is little or no arbitrariness involved? Our answer to that question is an unequivocal *yes*.

Some might now argue that a transitive construction with a nominative subject and an accusative object is not specific for Proto-Indo-European, since such a construction is extremely common across languages. This raises the question whether typological considerations should prevent us from reconstructing on the basis of preserved material. Should we, for example, not reconstruct a /p/ for Proto-Indo-European because it is cross-linguistically common? The sound

/p/ has certainly been identified in the earliest Indo-European languages on the basis of its occurrence in cognate words like 'father' and 'fish'. Since these lexical items have been inherited into the daughter languages, it is obvious that /p/ has also been inherited from an earlier proto-stage. On the basis of reconstructed words, the phonological system of Proto-Indo-European has been established (see, for example, Mayrhofer 1986), and the inventory of this reconstructed sound system uncontroversially contains /p/. Despite the fact that the sound /p/ is extremely common in the world's languages, Indo-Europeanists nevertheless confidently reconstruct it for Proto-Indo-European. The same would seem to apply to syntax.

One could now object even further and claim that a reconstruction of a Nom-Acc construction is banal. However, this is not a valid reason against such a reconstruction, as it is equally banal to say that Proto-Indo-European had a /p/ as it is to say that it had a transitive Nom-Acc construction. Thus, we conclude that a reconstruction on the basis of regularities is a legitimate enterprise, provided that this regularity is found in the daughter languages.

3.2 *Identification on the Basis of Cognate Structure*

One additional device to identify cognate argument structure constructions is through what Walkden calls the Double Cognacy Requirement (Walkden 2009, 2013). He points out that a sound, say /p/, cannot be reconstructed for Proto-Indo-European unless it is found in a cognate environment. To give an example, one cannot reconstruct /p/ on the basis of *piscis* 'fish' in Latin and *fadar* 'father' in Gothic. One has to reconstruct on the basis of *piscis* 'fish' in Latin and *fisks* 'fish' in Gothic, on the one hand, and on the basis of *pater* 'father' in Latin and *fadar* 'father' in Gothic, on the other. That is, a secure reconstruction is carried out on the basis of a cognate context, and not across cognate contexts.

Applying this requirement to argument structure, it is not only the lexical stem that may be cognate but also the case frame, and even the predicate structure. By case frame we mean valency patterns like Nom-Acc, Nom-Dat, Nom-Gen, Dat-Nom, Acc-Nom, Acc-Acc, Acc-Gen, etc. By predicate structure we refer to whether a verb is a simple verb, prefixed verb, compound verb, or a compositional predicate of some sort. As has already been discussed above, the case paradigms in Germanic are cognate; that is, the morphological markers in the case paradigms are inherited across the daughters. The case patterns themselves, i.e. the choice of subject and object case occurring together with a particular predicate, are also cognate, as evident from the fact that they are inherited into the daughters with the same lexical verbs (Barðdal & Eythórsson 2012a, 2012b; Barðdal & Smitherman 2013).

Let us illustrate all this with the compositional predicate 'be good for' from all three branches of Germanic, with examples from Gothic, Old English and Old Norse-Icelandic:

Gothic

(5a) *goþ*   *ist*   **unsis**   her   wisan
     good   is.3SG   us.DAT   here   be.INF
     'it is good for us to be here' (Luke 9: 33)

Old English

(5b) *betere*   *is*   manna   **gehwylcum** þæt him her on worulde
     better   is.3SG   men.GEN   each.DAT   that him here on world
     'Every man has it better here in the world' (Ælfric Homily 28.107)

Old Norse-Icelandic

(5c) *Betra*   *er*   **lifðum**   **og**   **sællifðum**.
     better   is.3SG   the.lived.DAT   and   the.well.lived.DAT
     'Those who live and live well have it better.' (Hávámál 70)

The examples in (5) fulfil not only the Double Cognacy Requirement, they in fact instantiate triple cognacy. Double (or triple) cognacy follows from duality (or triality) of patterning. For instance, sounds are combined into word forms, and word forms are combined into sentences. In the case of argument structure: a) the lexical material is cognate, i.e. the verb 'be' and the adjective 'good/better'; b) the predicate structure is cognate, involving a compositional predicate consisting of 'be' and an adjective; and c) the case frame is also cognate, i.e. the subject is in the dative case. To concretize, lexical items are combined into predicate structure; case markers are combined into case patterns combining with argument structures; and finally, lexical items and predicate structure are combined into argument structure through a process merging lexical items with more schematic argument structure constructions. This is how a combination of factors involving cognate structure may contribute to the identification of cognates in syntax.

For a reconstruction of more schematic argument structure constructions, see Barðdal & Smitherman (2013); Danesi, Johnson & Barðdal (2017) and Vazquez Gonzalez & Barðdal (2019), although it should be emphasized that the research program proposed here of identifying cognates does not of, course, rely on schematic reconstructions. On the contrary, it is based on substantive material being found instantiating the relevant argument structure constructions, as without substantive material, there can be no cognate lexical material. However, as we also emphasize above, lexical material is not always

needed to identify cognates, as sometimes this can also be done on the basis of grammatical cognate material, like when case morphology is cognate or when predicate structure is cognate (see also Kikusawa, this volume, and Gildea & Castro Alves, this volume).

Returning to the last discussion point in Section 3.1 above on reconstructing on the basis of regularity across the daughters, the question arises whether there is double cognacy in Nom-Acc argument structure constructions in the same sense as with compositional predicates. The case paradigms, here nominative and accusative, have already been reconstructed not only for Germanic but also for Proto-Indo-European (see, for example, Krahe 1969a, 1969b; Ringe 2006; see also Pooth et al. 2019 for the emergence, development and reconstruction of the case frames in Proto-Indo-European). A reconstruction of the transitive Nom-Acc construction would therefore not only be carried out on the basis of the existence of an inherited Nom-Acc argument structure construction in the daughter languages, but also on the basis of the morphological nominative and the morphological accusative.

A few examples of cognate lexical verbs instantiating the Nom-Acc construction in the early Germanic languages are listed in Table 5.9, some of which are quite high on the transitivity scale (Hopper & Thompson 1980). This means that not only is the case frame cognate, but also the lexical verbs. In fact, there are so many verbs that instantiate the Nom-Acc case frame, due to its property of being a default case pattern, assigned on the basis of analogical extensions of the statistically predominant pattern, that its reconstruction is almost banal, exactly as with the Proto-Indo-European /p/. However, exactly as with the Proto-Indo-European /p/, if one's intention is to reconstruct a grammatical system, both regular grammatical constructions and more idiosyncratic ones need to be reconstructed.

TABLE 5.9   Cognate Nom-Acc predicates in early Germanic

Gothic	Old English	Old Norse-Icelandic	Gloss
dreiban	*drīfan*	*drífa*	'drive'
maurþrjan	*myrðran*	*myrða*	'murder'
slahan	*slean*	*slá*	'hit'
saggjan	*senkan*	*sökkva*	'sink'
bairan	*beran*	*bera*	'carry'
drigkan	*drincan*	*drekka*	'drink'
itan	*etan*	*eta*	'eat'
saihwan	*sēon*	*sjá*	'see'
finþan	*findan*	*finna*	'find'

One could perhaps now argue that a reconstruction of the Nom-Acc case frame may be less secure, exactly because it is the default pattern. That is, this case frame can easily arise through a change from another more marked case frame. This is indeed true; therefore such reconstruction must rely heavily on the quality of the co-occurrence patterns, i.e. on double cognacy, in that both the case frame and the relevant lexical verbs must be cognates across several branches. A change from a more marked case frame to a less marked or to the default case frame is not expected to take place *with cognate verbs* in language after language within a language family, unless of course there has been a massive development of such a type of change. Such massive developments, however, if they are not already known, may be established on the basis of the comparative material, i.e. by comparing cognate verb sets within languages of one branch. Therefore, reconstruction of default argument structure constructions is also possible, provided that the double cognacy requirement is fulfilled and the procedures of the Comparative Method honored.

Finally, with regard to changes in argument structure constructions, some of these are quite well known (Barðdal 2014), including regularization processes like Nominative Sickness, whereby nominative is substituted for oblique case with subjects (Eythórsson 2000, 2002; Jónsson & Eythórsson 2005; Barðdal 2009, 2011a; Dunn et al. 2017). This process is found with both 'think' and 'thirst', discussed above, where dative and accusative have been replaced with a nominative, for instance in Modern German. In some of the other modern Germanic languages, like Modern English, Modern Swedish and Modern Dutch, this replacement happened through a different kind of process, namely the general loss of case morphology. Of course, a general loss of case morphology results in all argument structure constructions disappearing except for a generalized Nom-Acc construction, only visible with pronouns, thus distinguishing only between (nominative) subjects and (accusative) objects. As such, loss of case morphology constitutes the ultimate regularization process (cf. Barðdal 2009).

Processes resulting in the emergence of marked argument structure constructions have also been documented in the field. One such is Oblique Anticausativization, i.e. a process involving reduction in a verb's valency, first by creating synchronic correspondences between the oblique anticausative and its causative alternant, and then through the loss of the original causative alternant over time, leaving behind only the alternant with an oblique subject (Sandal 2011; Barðdal 2014; Barðdal et al. 2020). However, such a process of language change may also be detected through a rigorous investigation of the pre-stage. Any claims that there is lack of directionality in syntactic change are therefore unfounded when it comes to changes in argument structure constructions.

3.3 *Identification on the Basis of Synonymous Material*

So far in this article, we have shown how Watkins' method can be applied to argument structure constructions, where we find cognate lexical items, cognate case paradigms, cognate case frames, and even cognate predicate structure. Our data, as yet, come from Germanic, which of course represents a fairly short time span, compared to, for instance, Proto-Indo-European.

Identifying cognate argument structure constructions on the basis of cognate lexical verbs may become increasingly difficult when the time span is longer than from the modern languages to the closest reconstructed intermediate node in the family tree, because of the well-known process of *lexical substitution* (Firth 1935; D'arcy 2006; Calude & Pagel 2011; François 2011). An estimation of lexical replacement rates has been made by Pagel et al. (2007) and Pagel (2009) who suggest that the chances that a random cognate is replaced with a non-cognate word every 2,000–2,500 years is 50%, although this is highly dependent on frequency of use. That is, low-frequency words are replaced at a much faster rate and high-frequency words at a much slower rate.

Watkins (1995) certainly observes, in his work on poetic formulae, that with time, important content words are replaced, while the formula itself is maintained. This process has also been observed synchronically with idioms and set phrases (McGlone et al. 1994; Langlotz 2006). Given these observations, in combination with our present knowledge of the acquisition of case frames by new predicates, we suggest that cognate argument structure constructions may be identified despite a deeper time span. This means that cognate argument structure constructions may be distinguished, not only because they share a lexical cognate, but also by virtue of being instantiated by a synonymous verb. In other words, since lexical predicates tend to be replaced with synonymous predicates through time, it is reasonable to assume that argument structure constructions remain stable although the lexical predicate itself is renewed.

To give an example, the verb 'answer' in the Modern Germanic languages has two cognate sets, reflexes of the etymons *(*and*)*swaran* and *andwurdjan*. The facts discussed in Section 3.1 above suggest that *(*and*)*swaran* is an earlier verb with this meaning, and that *andwurdjan* replaced it in East and West Germanic. We also know that a verb meaning 'speak, put into words', usually selects for the Nom-Acc case frame in early Germanic, hence we assume that the Nom-Dat case frame with *andwurdjan* in East and West Germanic must have been assigned to *andwurdjan* on the basis of the case frame of *(*and*)* *swaran*. This is an internal reconstruction, since this development had already taken place during prehistoric times.

Now, if English and the modern North Germanic languages had also lost the verb *(*and*)*swaran*, resulting in only one cognate set for 'answer', namely the

modern reflexes of *andwurdjan*, would we then not be in total ignorance of where its case frame came from? Our answer to that question is *no*, since we know that *andwurdjan* takes a dative object by virtue of the fact that its predecessor, *(and)swaran*, took a dative object. In other words, since argument structures stay the same, while their lexical predicates are replaced with synonymous lexemes, it becomes possible to identify cognate argument structure constructions through synonymous verbs. Therefore, not only cognate lexical verbs aid in identifying cognate argument structure constructions, but also non-cognate synonymous verbs.

Let us illustrate this for Proto-Indo-European, continuing with the verb 'answer':

Ancient Greek (Homer): *hypokrinomai* 'reply, answer'
(6a) hōs **toi**   *hypokrinontai*
     how you.DAT.SG   answer.3PL
     '... how they answer you.' (Homer, *Iliad* 7.407)

Ancient Greek (Attic): *apokrinomai* 'reply, answer'
(6b) egō gar autik' *apokrinoumai* **soi**   saphōs
     I   for right.away will.answer.1SG you.DAT.SG clearly
     'for I will presently answer you distinctly.' (Aristophanes, *Clouds* 1245)

Latin: *respondeo* 'reply, answer'
(6c) **legatis**   *respondit*   diem   se   ad deliberandum
     ambassadors.DAT answered.3SG day.ACC self to deliberate
     sumpturum
     take
     'He [Caesar] replied to the ambassadors, that he would take time to deliberate.' (Caesar, *Gallic War* 1.7)

Gothic: *andwaurdjan* 'answer'
(6d) þu   hwas is,   ei   *andwaurdjais* **guda**?
     you who   are   that reply.2SG.OPT   God.DAT
     'Who are you replying to (OR: against) God?' (Rom 9:20)

Old Russian: *otvéčal-* 'answer'
(6e) I   Pskovъ   **imъ**   *otvéčalъ*
     and Pskov.NOM them.DAT replied.3SG
     'And Pskov answered them.' (Pskovskaja letopis' XVc)

Old Lithuanian: *atsakyti* 'answer'

(6f) Ar   šitaipo   **byskupui**   *atsisakai*?
     do   this   bishop.DAT   answer.2SG
     'Do you answer the bishop in this manner?'

(Bretkunas Postille I372, y. 1591)

Hittite: *āppa mema/i-* 'speak/say back'

(6g) ᵈUTU-uš   *ANA*   MUNUS.LUGAL   *āppa*   *memišta*
     sun.god-NOM   DAT   queen   back   spoke.3SG
     'The Sun God replied to the queen.' (KBo 20.82 ii 33–34)

Tocharian A: *wätk-* 'answer'

(6h) kupre   ne   säm   penu   sne   täṅklune   wätkālts
     if   he   PTCL   without   with   difficulty confidence
     wätkāṣṣ-äm ////
     answers:3SG.CONJ-CL.OBL.PL
     'If he responds to you without difficulty and with confidence' (YQ-14[11.5] b4)

Sanskrit: *prati-brū-* 'answer'

(6i) apṛcchaṃ   mātaram \   sā   **mā**   *pratyabravīt*
     asked.3SG   mother.ACC   she.NOM   me.ACC   answered.3SG
     'I asked my mother and she answered me' (Ch. 4.4.4)

The verb 'answer' takes Nom-Dat in all the earliest Indo-European languages in (6) above except for Sanskrit, where it takes a Nom-Acc. While Sanskrit is one of the oldest documented Indo-European languages, it is known to have innovated in some respects, for instance in the vowel system which has been simplified drastically (see, for example, Mayrhofer 1978). It also seems, on a comparison with the other early Indo-European languages, that predicates selecting for non-nominative subject-like arguments have been significantly reduced (cf. Danesi 2014). It is also well known that the accusative as an object case has heavily invaded the space of the dative in Sanskrit. Most importantly, however, for this example, the accusative is most likely governed by the prefix *prati-* (Leonid Kulikov, p.c.). It is therefore very likely that the accusative object of 'answer' is an innovation in Sanskrit.

Going systematically through the verbs in (6), the case marking of the addressee in Tocharian is ambiguous, as the clitic *-äm* is a general oblique form that does not distinguish between accusative, dative and the genitive, as is shown in the correspondence set in Table 5.10. The Latin, the Old Lithuanian

and the Hittite verbs might possibly have inherited the dative case of the addressee from their corresponding base verbs, as the unprefixed *spondeo* in Latin means 'promise', the unprefixed *sakyti* in Old Lithuanian means to 'say', and the simple *mema/i-* in Hittite means to 'speak'. Verbs having these meanings tend to assign dative case to the addressee in the Indo-European languages.

The situation is different with verbs like the unprefixed *krinomai* in Ancient Greek which is highly polysemous, instantiating meanings like 'judge', 'choose' and even 'distinguish'. With all three of these meanings, the object occurs in the accusative case and not in the dative. The Old Russian, *otvéčal-*, does not have an unprefixed counterpart, but with the prefix *pre-*, it means 'welcome' or 'receive in a friendly way', selecting for an accusative object. The same is true for the unprefixed Gothic **waurdjan* 'put into words', selecting for an accusative object, as already discussed in Section 3.1 above. Thus, the Greek, Slavic and Germanic data exclude an analysis involving case and argument structure assignment on the basis of cognate stems already existing in the language, while the Baltic, Italic and Anatolian data are compatible with either analysis, i.e. either case and argument structure has been assigned on the basis of already existing cognate verb stems or on the basis of synonymous verbs.

Hence, on the basis of synonymous material from at least three branches of Indo-European, Greek, Slavic and Germanic, we reconstruct Nom-Dat as the predicate-specific argument structure construction of 'answer' in Proto-Indo-European, cf. the correspondence sets in Table 5.10 and the partial reconstruction in Figure 5.4.

TABLE 5.10   Predicate-specific correspondence set for the argument structure of a Proto-Indo-European verb meaning 'answer'

	Alt. 1	Alt. 2	Alt. 3
Ancient Greek	NOM-DAT		
Latin	NOM-DAT		
Gothic	NOM-DAT		
Old Russian	NOM-DAT		
Old Lithuanian	NOM-DAT		
Hittite	NOM-DAT		
Tocharian B	NOM-DAT	NOM-ACC	NOM-GEN
Sanskrit		NOM-ACC	

*	*lxm*		
	FORM	< >	
	SYN	ARG-ST <NP-NOM$_i$, NP-DAT$_j$>	
	SEM	FRAMES	communication_response fr SPEAKER i ADDRESSEE j

FIGURE 5.4 A reconstruction of the argument structure of 'answer' in Proto-Indo-European

The reconstruction in Figure 5.4 is partial because the field for the lexical verb is empty, i.e. it does not contain any phonological material. In other words, this reconstruction is a reconstruction of the verb-specific argument structure construction Nom-Dat for a verb with the meaning 'answer' in Proto-Indo-European. While we do not know the form of the verb, we posit, on the basis of a comparison between the daughters, that a verb with this meaning must have instantiated a Nom-Dat case frame in Proto-Indo-European. The identification of this verb-specific argument structure construction is based on non-cognate synonymous predicates across the daughters, and not on a cognate lexical item. This reconstruction is certainly more schematic than traditional reconstructions based on cognate lexical material. It is nevertheless based on a known linguistic process, in which predicates get replaced by their synonyms in the course of time, while simultaneously maintaining their original argument structure constructions.

Let us consider one final example with an even more exceptional case frame, namely the Dat-Gen frame. Several Indo-European daughter languages, from at least four different branches, exhibit a predicate meaning 'lack, need' which occurs with this case frame (see also Frotscher, Kroonen & Barðdal 2020). The examples below are from Ancient Greek, Gothic, Old Russian and Lithuanian:

Ancient Greek (*prosdéō* 'need')
(7a) **hoûper   humîn**   málista   *prosdeî*
       which.GEN   you.DAT   very.much   need.3SG
       'which you need very much' (Thuc 3.13)

Gothic (*wan ist* 'is lacking')

(7b) **ainis**   **þus**   *wan*   *ist*
     one.GEN   you.DAT   lacking   is.3sg
     'you lack one thing' (Mk 10.21)

Old Russian (*lixyi* 'lack')

(7c) Zenicǫ   bo   ne   **imy**   *lixъ*   *jestь*   **světa**
     pupil.ACC because.PTCL NEG them.DAT lacking   be.3SG   light.GEN
     'they are not lacking light in their eyes'

                                         (Pandekt Antioxa po spisku XI veka)

Lithuanian (*trūkti* 'lack')

(7d) **Jam**   *trūksta*   **kantrybės**
     he.DAT   lacks.3SG   patience.GEN
     'He lacks patience'

In our view, the data in (7) motivate a reconstruction of a predicate in Proto-Indo-European with the meaning 'lack, need', as sufficiently established with data from four different branches. Such a reconstruction would be only partial, exactly as the reconstruction of 'answer' in Proto-Indo-European in Figure 5.4 above, with no phonological material. The reconstruction would be of a purely schematic argument structure construction containing the Dat-Gen case frame and the meaning 'lack, need'. As with 'answer' above, this reconstruction would be based on synonymous lexical material and not on cognate lexical material.

Furthermore, not only are partial reconstructions of predicates and their argument structures possible, as we have just demonstrated (cf. also Barðdal & Eythórsson 2012a; Barðdal & Smitherman 2013; see also Vazquez Gonzalez & Barðdal 2019 for a reconstruction of a verb-class-specific argument structure construction), but the discovery of the important developmental path whereby lexical predicates get replaced by their synonyms, while case frames are maintained, makes certain predictions in a wider historical linguistic context. One such prediction is that lexical semantic verb classes will be linked to the same argument structure constructions across time, provided of course that the case and alignment system stays more or less intact, and hence that argument structure constructions are reconstructable as such for earlier proto-stages on the basis of lexical semantic verb classes alone (cf. Barðdal 2007; Barðdal et al. 2011; Barðdal et al. 2012). In other words, this is a force for conservatism, maintaining irregularity. In this way, the insights inherent in Watkins' legacy that morphological material may be used for reconstructing abstract

syntactic units are taken one step further, potentially yielding a substantial improvement in the methodology of linguistic reconstruction.

4 Summary

The Comparative Method in historical linguistics has been successfully employed in reconstructing phonological, morphological and lexical units, not only by the Neogrammarians in the late 19th and early 20th centuries, but further into modern times. Syntactic reconstruction, in contrast, has been regarded as a precarious enterprise, an unreliable exercise, and fraught with pitfalls at every stage.

As a reaction to three different attempts at reconstructing basic word order for Proto-Indo-European, Watkins' paper from 1976 was hugely influential in the general renouncing of syntactic reconstruction. The first of these three attempts was done on the basis of Greenberg's universals, the second on the basis of word order frequencies in Homeric Greek, and the third one on the basis of the word orders attested in the early daughters. It was not until nearly 20 years later, however, in his 1995 book on Indo-European poetics, that Watkins properly proposed a research program which opened up new vistas for syntactic reconstruction, showcasing his method with an investigation of cognate dragon myths across the early Indo-European daughter languages. The core of Watkins' proposal lies in carrying out syntactic reconstruction through morphological devices, utilizing common morphosyntactic material as a means of identifying cognates. Watkins' work on poetic formulae shows how layers of cognate collocations and prefabs can be identified through morphological clues, together with important fragments of syntax from earlier periods of the Indo-European languages, which are reconstructable as such.

We maintain that the ability to reconstruct *all of morphosyntax* should by no means be considered insignificant. Moreover, the area of syntactic research which we have mostly concentrated on, that is, case and argument structure constructions, constitutes in itself a major subfield within syntactic research and syntactic theorizing. Launching a research program into how to reconstruct argument structure constructions is therefore no insignificant task, in a scientific atmosphere that has, for almost four decades, been engulfed with despondency over the perceived failure of syntactic reconstruction.

As a part of this research program, we have demonstrated how cognate argument structure constructions may be identified, with the aid of a) the lexical predicates that instantiate them, b) cognate case frames, c) cognate predicate structures, and d) cognate case morphology. For this purpose we

have compared case and argument structure constructions of three Germanic verbs, 'think', 'thirst', and 'answer', all of which have a case frame that deviates from the canonical Nom-Acc frame, and hence exhibits a certain degree of arbitrariness. The directionality of the changes is in part retrievable from documented processes and is in part revealed by a proper scrutiny of the datasets themselves.

However, we do not stop at that; our aim here has been to bring Watkins' research program one step further, and to show how cognate argument structure constructions may be identified on the basis of non-cognate synonymous predicates. This claim is based on documented processes of how new verbs acquire their case and argument structure constructions, of which one major process involves new verbs attracting case frames on analogy to already existing synonymous verbs. We have thus illustrated how cognate argument structure constructions may be identified using non-cognate lexical material through two case studies. The first involves the verb 'answer' which has two cognate sets in Germanic, but at least eight sets across Indo-European. This larger number of cognate sets is expected, given the greater time depth for Proto-Indo-European than for Proto-Germanic, and given our current knowledge of the speed of lexical replacement. We have reconstructed a predicate-specific argument structure construction, Nom-Dat, for Proto-Indo-European, on the basis of the evidence provided by synonymous predicates in the early daughters.

In our second case study, we have examined the verb 'lack, need', occurring with the Dat-Gen case frame and with synonymous non-cognate lexical material only, from four different branches of Indo-European, again illustrating the viability of this method. Ultimately, we argue that semantic spaces can be reconstructed for argument structure constructions at earlier proto-stages, on the basis of lexical-semantic verb classes, since argument structure constructions may remain stable while lexical items are replaced. In this way, Watkins' program can be taken one step further, from reconstructing on the basis of cognate lexical material to reconstructing on the basis of non-cognate synonymous material.

References

Adger, David. 2003. *Core Syntax: A Minimalist Approach*. Oxford: Oxford University Press.

Allen, Cynthia L. 1995. *Case Marking and Reanalysis: Grammatical Relations from Old to Modern English*. Oxford: Oxford University Press.

Barðdal, Jóhanna. 1999. Case and Argument Structure of some Loan Verbs in 15th Century Icelandic. *Alla tiders språk. En Vänskrift till Gertrud Pettersson november 1999*, ed. by Inger Haskå & Carin Sandqvist, 9–23. Lundastudier i Nordisk språkvetenskap A 55. Lund: Institutionen för nordiska språk.

Barðdal, Jóhanna. 2000. Oblique Subjects in Old Scandinavian. *NOWELE: North-Western European Language Evolution* 37: 25–51.

Barðdal, Jóhanna. 2001. *Case in Icelandic – A Synchronic, Diachronic and Comparative Approach*. Lundastudier i Nordisk språkvetenskap A 57. Lund: Department of Scandinavian Languages.

Barðdal, Jóhanna. 2007. The Lexical and Semantic Range of the Ditransitive Construction in the History of (North) Germanic. *Functions of Language* 14(1): 9–30.

Barðdal, Jóhanna. 2008. *Productivity: Evidence from Case and Argument Structure in Icelandic*. Amsterdam: John Benjamins.

Barðdal, Jóhanna. 2009. The Development of Case in Germanic. *The Role of Semantic, Pragmatic, and Discourse Factors in the Development of Case*, ed. by Jóhanna Barðdal & Shobhana L. Chelliah, 123–159. Amsterdam: John Benjamins.

Barðdal, Jóhanna. 2011a. The Rise of Dative Substitution in the History of Icelandic: A Diachronic Construction Grammar Account. *Lingua* 121(1): 60–79.

Barðdal, Jóhanna. 2011b. Lexical vs. Structural Case: A False Dichotomy. *Morphology* 21(3–4): 619–659.

Barðdal, Jóhanna. 2012. Predicting the Productivity of Argument Structure Constructions. *Berkeley Linguistics Society* 32(2006): 467–478.

Barðdal, Jóhanna. 2013. Construction-Based Historical-Comparative Reconstruction. *The Oxford Handbook of Construction Grammar*, ed. by Greame Trousdale & Thomas Hoffmann, 438–457. Oxford: Oxford University Press.

Barðdal, Jóhanna. 2014. Syntax and Syntactic Reconstruction. *The Routledge Handbook of Historical Linguistics*, ed. by Claire Bowern & Bethwyn Evans, 343–373. London: Routledge.

Barðdal, Jóhanna, Valgerður Bjarnadóttir, Serena Danesi, Tonya Kim Dewey, Thórhallur Eythórsson, Chiara Fedriani & Thomas Smitherman. 2013. The Story of 'Woe'. *Journal of Indo-European Studies* 41(3–4): 321–377.

Barðdal, Jóhanna & Thórhallur Eythórsson. 2003. The Change that Never Happened: The Story of Oblique Subjects. *Journal of Linguistics* 39(3): 439–472.

Barðdal, Jóhanna & Thórhallur Eythórsson. 2012a. Reconstructing Syntax: Construction Grammar and the Comparative Method. *Sign-Based Construction Grammar*, ed. by Hans C. Boas & Ivan A. Sag, 257–308. Stanford: CSLI Publications.

Barðdal, Jóhanna & Thórhallur Eythórsson. 2012b. "Hungering and Lusting for Women and Fleshly Delicacies": Reconstructing Grammatical Relations for Proto-Germanic. *Transactions of the Philological Society* 110(3): 363–393.

Barðdal, Jóhanna, Kristian E. Kristoffersen & Andreas Sveen. 2011. West Scandinavian Ditransitives as a Family of Constructions: With a Special Attention to the Norwegian V-REFL-NP Construction. *Linguistics* 49(1): 53–104.

Barðdal, Jóhanna & Leonid Kulikov. 2009. Case in Decline. *The Oxford Handbook of Case*, ed. by Andrej Malchukov & Andrew Spencer, 470–477. Oxford: Oxford University Press.

Barðdal, Jóhanna, Leonid Kulikov, Roland Pooth & Peter Alexander Kerkhof. 2020. Oblique Anticausatives: A Morphosyntactic Isogloss in Indo-European. To appear in a special issue on "Morphological Isoglosses in Indo-European" in *Poznan Studies in Contemporary Linguistics*.

Barðdal, Jóhanna & Thomas Smitherman. 2013. The Quest for Cognates: A Reconstruction of Oblique Subject Constructions in Proto-Indo-European. *Language Dynamics and Change* 3(1): 28–67.

Barðdal, Jóhanna, Thomas Smitherman, Valgerður Bjarnadóttir, Serena Danesi, Gard B. Jenset & Barbara McGillivray. 2012. Reconstructing Constructional Semantics: The Dative Subject Construction in Old Norse-Icelandic, Latin, Ancient Greek, Old Russian and Old Lithuanian. *Studies in Language* 36(3): 511–547.

Bauer, Brigitte L. M. 2000. *Archaic Syntax in Indo-European: The Spread of Transitivity in Latin and French*. Berlin: Mouton de Gruyter.

Boas, Hans C. 2003. *A Constructional Approach to Resultatives*. Stanford: CSLI Publications.

Bowern, Claire. 2008. Syntactic Change and Syntactic Reconstruction in Generative Grammar. *Principles of Syntactic Reconstruction*, ed. by Gisela Ferraresi & Maria Goldbach, 187–216. Amsterdam: John Benjamins.

Brugmann, Karl. 1913. Zu den Ablautverhältnissen der sogenannten starken Verba des Germanischen. *Indogermanische Forschungen* 32: 179–195.

Calude, Andreea S. & Mark Pagel. 2011. How Do We Use Language? Shared Patterns in the Frequency of Word Use across 17 World Languages. *Philosophical Transactions of the Royal Society B* 366: 1101–1107.

Campbell, Lyle & Alice C. Harris. 2003. Syntactic Reconstruction and Demythologizing 'Myth and the Prehistory of Grammars'. *Journal of Linguistics* 38: 599–618.

Campbell, Lyle & Marianne Mithun. 1980. The Priorities and Pitfalls of Syntactic Reconstruction. *Folia Linguistica Historica* 1: 19–40.

Comrie, Bernard. 1980. Morphology and Word Order Reconstruction: Problems and Prospects. *Historical Morphology*, ed. by Jacek Fisiak, 83–96. The Hague: Mouton.

Croft, William. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.

Croft, William. 2009. Connecting Frames and Constructions: A Case Study of "eat" and "feed". *Constructions and Frames* 1: 7–28.

Croft, William & D. Alan Cruse. 2004. *Cognitive Linguistics*. Cambridge: Cambridge University Press.

Danesi, Serena. 2014. Accusative Subjects in Avestan: 'Errors' or Noncanonically Marked Arguments. *Indo-Iranian Journal* 57(3): 223–260.

Danesi, Serena, Cynthia A. Johnson & Jóhanna Barðdal. 2017. Between the Historical Languages and the Reconstructed Language: An Alternative Approach to the Gerundive + "Dative of Agent" Construction in Indo-European. *Indogermanische Forschungen* 122: 143–188.

Daniels, Don. 2015. *A Reconstruction of Proto-Sogeram: Phonology, Lexicon, and Morphosyntax*. University of California, Santa Barbara, Ph.D. dissertation.

Daniels, Don. 2017. A Method for Mitigating the Problem of Borrowing in Syntactic Reconstruction. *Studies in Language* 41(3): 577–614.

D'arcy, Alexandra. 2006. Lexical Replacement and the Like(s). *American Speech* 81(4): 339–357.

Delbrück, Berthold. 1878. *Die altindische Wortvolge aus dem Çatapathabrāhmana dargestellt* (Syntaktische Forschungen III). Halle: Verlag des Waisenhauses.

Dunn, Michael, Tonya Kim Dewey, Carlee Arnett, Thórhallur Eythórsson & Jóhanna Barðdal. 2017. Dative Sickness: A Phylogenetic Analysis of Argument Structure Evolution in Germanic. *Language* 93(1): e1–e22.

Eythórsson, Thórhallur. 2000. Dative vs. Nominative: Changes in Quirky Subjects in Icelandic. *Leeds Working Papers in Linguistics* 8: 27–44.

Eythórsson, Thórhallur. 2002. Changes in Subject Case Marking in Icelandic. *Syntactic Effects of Morphological Change*, ed. by David Lightfoot, 196–212. Oxford: Oxford University Press.

Eythórsson, Thórhallur & Jóhanna Barðdal. 2005. Oblique Subjects: A Common Germanic Inheritance. *Language* 81(4): 824–881.

Eythórsson, Thórhallur & Jóhanna Barðdal. 2011. Die Konstruktionsgrammatik und die komparative Methode. *Indogermanistik und Linguistik im Dialog: Akten der XIII. Fachtagung der Indogermanischen Gesellschaft vom 21. bis 27. September 2008 in Salzburg*, ed. by Thomas Krisch & Thomas Lindner, 148–156. Wiesbaden: Reichert Verlag.

Eythórsson, Thórhallur & Jóhanna Barðdal. 2016. Syntactic Reconstruction in Indo-European: The State of the Art. *Franz Bopp and his Comparative Grammar Model (1816–2016)*, ed. by J. Gorrochategui, C. García Castillero & J. M. Vallejo. A special monographic volume in *Veleia* 33: 83–102.

Falk, Cecilia. 1997. *Fornsvenska upplevarverb* [Old Swedish Experiencer Verbs], Lund: Lund University Press.

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

Fillmore, Charles J. 1982. Frame Semantics. *Linguistics in the Morning Calm*, ed. by the Linguistic Society of Korea, 111–138. Seoul: Hanshin.

Fillmore, Charles J. 1985. Frames and the Semantics of Understanding. *Quaderni di Semantica* 6(2): 222–254.

Fillmore, Charles J. & Collin F. Baker. 2009. A Frames Approach to Semantic Analysis. *The Oxford Handbook of Linguistic Analysis*, ed. by Bernd Heine & Heiko Narrog, 313–339. Oxford: Oxford University Press.

Fillmore, Charles J., Paul Kay & Mary Catherine O'Connor. 1988. Regularity and Idiomaticity in Grammatical Constructions: The Case of *Let Alone*. *Language* 64: 501–538.

Firth, J. R. 1935. The Technique of Semantics. *Transactions of the Philological Society* 34(1): 36–73.

Fox, Anthony. 1995. *Linguistic Reconstruction: An Introduction to Theory and Method*. Cambridge: Cambridge University Press.

François, Alexandre. 2011. Social Ecology and Language History in the Northern Vanuatu Linkage: A Tale of Divergence and Convergence. *Journal of Historical Linguistics* 1(2): 175–246.

Fried, Mirjam. 2015. Irregular Morphology in Regular Syntactic Patterns: A Case of Constructional Re-Alignment. *Diachronic Construction Grammar*, ed. by Jóhanna Barðdal, Spike Gildea, Elena Smirnova & Lotte Sommerer, 141–174. Amsterdam: John Benjamins.

Fried, Mirjam & Jan-Ola Östman. 2005. Construction Grammar and Spoken Language: The Case of Pragmatic Particles. *Journal of Pragmatics* 37(11): 1752–1778.

Friedrich, Paul. 1975. *Proto-Indo-European Syntax: The Order of Meaningful Elements*. (Journal of Indo-European Studies Monograph 1.) Washington, D.C.: Institute for the Study of Man.

Frotscher, Michael, Guus Kroonen & Jóhanna Barðdal. 2020. Indo-European Inroads into the Syntactic-Etymological Interface: A Reconstruction of the PIE Verbal Root **menkʷ-* 'lack' and its Argument Structure. Submitted.

Gildea, Spike. 1998. *On Reconstructing Grammar: Comparative Cariban Morphosyntax*. Oxford: Oxford University Press.

Gildea, Spike. 2000. On the Genesis of the Verb Phrase in Cariban Languages. *Reconstructing Grammar: Comparative Linguistics and Grammaticalization*, ed. by Spike Gildea, 65–106, Amsterdam: John Benjamins.

Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Goldberg, Adele E. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.

Harris, Alice C. 2008. Reconstruction in Syntax: Reconstruction of Patterns. *Principles of Syntactic Reconstruction*, ed. by Gisella Ferraresi & Maria Goldbach, 73–95. Amsterdam: John Benjamins.

Harris, Alice C. & Lyle Campbell. 1995. *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: Cambridge University Press.

Harrison, S. P. 2003. On the Limits of the Comparative Method. *The Handbook of Historical Linguistics*, ed. by Brian D. Joseph & Richard D. Janda, 343–368. Oxford: Blackwell.

Hoenigswald, H. M. 1978. The Annus Mirabilis 1876 and Posterity. *Transactions of the Philological Society* 76: 17–35.

Holland, Gary. 2003. Syntactic Reconstruction. *The International Encyclopedia of Linguistics*, 2nd edition, ed. by William Frawley, Vol. 3, 448–450. Oxford: Oxford University Press.

Hopper, Paul & Sandra A. Thompson. 1980. Transitivity in Grammar and Discourse. *Language* 56: 251–299.

Jeffers, Robert J. 1976. Syntactic Change and Syntactic Reconstruction. *Current Progress in Historical Linguistics: Proceedings of the Second International Conference on Historical Linguistics*, ed. by William M. Christie, Jr. (ed.), 1–15. Amsterdam: North-Holland Publishing.

Jespersen, Otto. 1927. *A Modern English Grammar on Historical Principles*, Vol. 3. Heidelberg: Winter.

Johnson, Cynthia A., Peter Alexander Kerkhof, Leonid Kulikov, Esther Le Mair & Jóhanna Barðdal. 2019. Argument Structure, Conceptual Metaphor and Semantic Change: How to Succeed in Indo-European without Really Trying. *Diachronica* 36(4): 463–508.

Jónsson, Jóhannes Gísli. 2003. Not so Quirky: On Subject Case in Icelandic. *New Perspectives in Case Theory*, ed. by Ellen Brandner & Heike Zinsmeister, 127–163. Stanford: CSLI Publications.

Jónsson, Jóhannes Gísli & Thórhallur Eythórsson. 2005. Variation in Subject Case Marking in Insular Scandinavian. *Nordic Journal of Linguistics* 28(2): 223–245.

Kay, Paul & Charles J. Fillmore. 1999. Grammatical Constructions and Linguistic Generalizations: The *What's X doing Y?* Construction. *Language* 75(1): 1–33.

Kikusawa, Ritsuko. 2002. *Proto Central Pacific Ergativity: Its Reconstruction and Development in the Fijian, Rotuman and Polynesian languages*. Canberra: Pacific Linguistics.

Kikusawa, Ritsuko. 2003. The Development of Some Indonesian Pronominal Systems. *Historical Linguistics 2001: Selected Papers from the 15th International Conference on Historical Linguistics, Melbourne, 13–17 August 2001*, ed. by Barry J. Blake, Kate Burridge & Jo Taylor, 237–268. Amsterdam: John Benjamins.

Klein, Jared S. 2010. Review of Principles of Reconstruction, ed. by Gisela Ferraresi & Maria Goldbach (2008). *Language* 86: 720–726.

Klimov, G[eorgiĭ] A[ndreevich]. A. 1977. *Tipologija jasykov aktivnogo stroja* [The Typology of Active Languages]. Moscow: Nauka.

Kluge, Friedrich. 2002. *Etymologisches Wörterbuch der deutschen Sprache*, 24th edition, by Elmar Seebold. Berlin: Walter de Gruyter.

Krahe, Hans. 1969a. *Germanische Sprachwissenschaft*, Vol. 3: *Wortbildungslehre*. 7th edition. Berlin: Walter de Gruyter.

Kroonen, Guus. 2013. *Etymological Dictionary of Proto-Germanic*. Leiden: Brill

Krahe, Hans. 1969b. *Indogermanische Sprachwissenschaft*, Vol. 2: *Formenlehre*. Berlin: Walter de Gruyter.

Kulikov, Leonid & Nikolaos Lavidas (eds.). 2013. Proto-Indo-European Syntax and its Development, a guest-edited issue in *Journal of Historical Linguistics* 3(2): 1–152.

Kuroda, Susumu. 2014. *Die Syntaktische Funktion der Präfigierung Im Althochdeutschen*. Heidelberg: Winter.

Langlotz, Andreas. 2006. *Idiomatic Creativity: A Cognitive Linguistic Model of Idiom-Representations and Idiom-Variation in English*. Amsterdam: John Benjamins.

Lehmann, Winfred P. 1974. *Proto-Indo-European Syntax*. Austin: University of Texas Press.

Lightfoot, David W. 1979. *Principles of Diachronic Syntax*. Cambridge: Cambridge University Press.

Lightfoot, David W. 2002a. Myths and the Prehistory of Grammars. *Journal of Linguistics* 38(1): 113–136.

Lightfoot, David W. 2002b. More Myths. *Journal of Linguistics* 38(3): 619–626.

Lightfoot, David W. 2006. *How New Languages Emerge*. Cambridge: Cambridge University Press.

Luraghi, Silvia. 1987. Patterns of Case Syncretism in Indo-European Languages. *Papers from the VIIth International Conference on Historical Linguistics*, ed. by Anna G. Ramat, Onofrio Carruba & Guiliano Bernini, 355–377. Amsterdam: John Benjamins.

Mayrhofer, Manfred. 1978. *Sanskrit-Grammatik mit sprachvergleichenden Erläuterungen*. Berlin: Walter de Gruyter.

Mayrhofer, Manfred. 1986. *Indogermanische Grammatik*, Volume 1: *Lautlehre*. Heidelberg: Winter.

McGlone, Matthew S., Sam Glucksberg & Christina Cacciari. 1994. Semantic Productivity and Idiom Comprehension. *Discourse Processes* 17: 167–190.

Mendoza, Julia. 1998. Syntaxis. *Manual de Lingüística Indoeuropea* III, ed. by Francisco R. Adrados, Alberto Bernabé & Julia Mendoza, 141–246. Madrid: Ediciones Clásicas.

Mengden, Ferdinand von. 2008. Reconstructing Complex Structures: A Typological Perspective. *Principles of Syntactic Reconstruction*, ed. by Gisella Ferraresi & Maria Goldbach, 97–119. Amsterdam: John Benjamins.

Michaelis, Laura A. 2009. Sign-based Construction Grammar. *The Oxford Handbook of Linguistic Analysis*, ed. by Bernd Heine & Heiko Narrog, 155–176. Oxford: Oxford University Press.

Michaelis, Laura A. 2012. Making the Case for Construction Grammar. *Sign-Based Construction Grammar*, ed. by Hans C. Boas & Ivan A. Sag, 31–68. Stanford: CSLI Publications.

Michaelis, Laura A. & Josef Ruppenhofer. 2001. *Beyond Alternations: A Constructional Model of the German Applicative Pattern*. Stanford: CSLI Publications.

Miller, D. Gary. 1975. Indo-European: *VSO*, *SOV*, *SVO*, or all three? *Lingua* 37: 31–52.

Miranda, Rocky. 1976. Comments on Jeffers. *Current Progress in Historical Linguistics: Proceedings of the Second International Conference on Historical Linguistics*, ed. by William M. Christie, Jr., 12–14. Amsterdam: North-Holland Publishing.

Nichols, Johanna. 2008. Why Are Stative-Active Languages Rare in Eurasia? A Typological Perspective on Split Subject Marking. *The Typology of Semantic Alignment*, ed. by Mark Donohue & Søren Wichmann, 121–139. Oxford: Oxford University Press.

Pagel, Mark. 2009. Human Language as a Culturally Transmitted Replicator. *Nature Reviews Genetics* 10: 405–415.

Pagel, Mark, Quentin D. Atkinson & Andrew Meade. 2007. Frequency of Word-Use Predicts Rates of Lexical Evolution throughout Indo-European History. *Nature* 449: 717–720.

Petruck, Miriam. 1996. Frame Semantics. *Handbook of Pragmatics*, ed. by J. Verschueren, J-O Östman, J. Blommaert & C. Bulcaen, 1–13. Amsterdam: John Benjamins.

Pires, Acrisio & Sarah G. Thomason. 2008. How Much Syntactic Reconstruction is Possible? *Principles of Syntactic Reconstruction*, ed. by Gisela Ferraresi & Maria Goldbach, 27–72. Amsterdam: John Benjamins.

Primus, Beatrice. 2012. *Semantische Rollen*. Heidelberg: Carl Winter.

Ringe, Don. 2006. *From Proto-Indo-European to Proto-Germanic*. Oxford: Oxford University Press.

Roberts, Ian. 2007. *Diachronic Syntax*. Oxford: Oxford University Press.

Rögnvaldsson, Eiríkur. 1991. Quirky subjects in Old Icelandic. *Papers from the Twelfth Scandinavian Conference of Linguistics*, ed. by Halldór Á. Sigurðsson, 369–378. Reykjavík: Institute of Linguistics, University of Iceland.

Rögnvaldsson, Eiríkur. 1995. Old Icelandic: A Non-Configurational Language? *NOWELE: North-Western European Language Evolution* 26: 3–29.

Sag, Ivan. 2012. Sign-Based Construction Grammar: An Informal Synopsis. *Sign-Based Construction Grammar*, ed. by Hans C. Boas & Ivan A. Sag, 69–202. Stanford: CSLI Publications.

Sandal, Catrine. 2011. Akkusative subjekt og antikausativitet i norrønt [Accusative Subjects and Anticausativization in Old Norse-Icelandic]. University of Bergen MA thesis.

Seefranz-Montag, Ariane von. 1982. *Syntaktische Funktionen und Wortstellungs veränderung: Die Entwicklung 'subjektloser' Konstruktionen in einigen Sprachen*. Munich: Fink.

Serzant, Ilja A. 2015. An Approach to Syntactic Reconstruction. *Perspectives on Historical Syntax*, ed. by Carlotta Viti, 117–154. Amsterdam: John Benjamins.

Smitherman, Thomas. 2015. Reconstructing Non-Canonical Argument Structure for Proto-Indo-European: Methodological Questions and Progress. *Perspectives on Historical Syntax*, ed. by Carlotta Viti, 95–116. Amsterdam: John Benjamins.

Tomasello, Michael. 2003. *Constructing a Language: Usage-Based Theory of Language Acquisition.* Cambridge, Ma.: Harvard University Press.

Viti, Carlotta. 2014. Reconstructing Syntactic Variation in Proto-Indo-European. *Indo-European Linguistics* 2: 73–111.

Vazquez Gonzalez, Juan G. & Jóhanna Barðdal. 2019. Reconstruction the Ditransitive Construction for Proto-Germanic: Gothic, Old English and Old Norse-Icelandic. *Folia Linguistica Historica* 40(2): 555–620.

Vries, Jan de. 1962. *Altnordisches Etymologisches Wörterbuch.* 2nd Ed. Leiden: Brill.

Walkden, George. 2009. *The Comparative Method in Syntactic Reconstruction.* Cambridge University M.Phil. thesis.

Walkden, George. 2013. The Correspondence Problem in Syntactic Reconstruction. *Diachronica* 30(1): 95–122.

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

Watkins, Calvert. 1964. Preliminaries to the Reconstruction of Indo-European Sentence Structure. *Proceedings of the IX International Congress of Linguists*, ed. by H. G. Lunt, 1035–1045. The Hague: Mouton.

Watkins, Calvert. 1976. Towards Proto-Indo-European Syntax: Problems and Pseudo-problems. *Papers from the Parasession on Diachronic Syntax*, ed. by Sanford B. Steever, Carol A. Walker & Salokoko S. Mufwene, 306–326. Chicago: Chicago Linguistic Society.

Watkins, Calvert. 1995. *How to Kill a Dragon: Aspects of Indo-European Poetics.* Oxford: Oxford University Press.

Wegener, Heide. 1985. *Der Dativ im heutigen Deutsch.* Tübingen: Narr.

Willis, David. 2011. Reconstructing Last Week's Weather: Syntactic Reconstruction and Brythonic Free Relatives. *Journal of Linguistics* 47(2): 407–446.

Winter, Werner. 1984. Reconstructional Comparative Linguistics and the Reconstruction of the Syntax of Undocumented Stages in the Development of Languages and Language Families. *Historical Syntax*, ed. by Jacek Fisiak, 613–625. The Hague: Mouton.

PART 2

Directionality

∴

On the Origins of the Ergative Marker *wã* in the Viceitic Languages of the Chibchan Family

Sara Pacchiarotti

1 Introduction

Bribri and Cabécar are Chibchan languages spoken in Costa Rica in contiguous territories.[1] Both languages display two distinct ergative markers, diachronically unrelated in any obvious way. The canonical, standard ergative marker is *tö* in Bribri (1) and *të* in Cabécar (2). A further ergative marker *wã* seems to be construction-determined and is more restricted in terms of usage. It is found, among others, in the perfect construction, cf. (3) and (4).

(1) *ye'   tö   ú   sũ'*   BRIBRI
     1SG   ERG   house   see.PFV.REM
     'I saw the house.'[2]

---

1 I am grateful to Alí García Segura, native Bribri speaker, and to Severiano Fernández Torres, native Cabécar speaker, for their invaluable help and patience, without whom this article would not have been possible. I am much indebted to Spike Gildea who guided my reasoning in the right direction during the entire writing process. I am grateful to Doris Payne, Miguel Angel Quesada Pacheco, Eric Campbell, Scott DeLancey and two anonymous reviewers for insightful comments, critiques and discussion on earlier drafts of this article. All errors and shortcomings are my own.

2 The marking of nasality follows the conventions established in Constenla et al. (1998), not those of Jara Murillo & García Segura (2013). However, unlike Constenla et al. (1998) I use the symbol <ã>, not <a̱>, to mark nasality. The data in this article comes from the dialectal variety of Bribri spoken in Coroma. In my description, unlike others, falling tone is indicated as <â>, whereas high tone is indicated as <á>. Low tone is not marked in the orthography. It should be noted, however, that the tonal system of Bribri is poorly described and more work is needed in order to fully understand it. The examples report tonal transcriptions found in the original text wherever applicable, adapted to the tonal conventions set out above. In addition, elicited examples are transcribed differently from examples found in other sources. In elicited examples, I transcribe the reduced set of short personal pronouns (1SG [ɟ]/[ɲ], 2SG [b]/[m], 3SG [i]/[j], 1PL [s]) as prefixes, and the negation morpheme *kẽ́* as a pro-clitic. I refer the interested reader to Chevrier (2017) for phonological arguments in favor of this classification. The abbreviation (EL) means data is from elicitation. Abbreviations for Bribri and other languages considered in this article are as follows: 1= first person, 2= second person, 3= third person, ASP= aspectual marker, COMPL= complementizer, CPL= completive aspect;

(2) *yís   të   jũ   sṹ-l*   CABÉCAR
    1SG   ERG   house see-PFV.REM
    'I saw the house.'[3] (EL)

(3) *ye'   wã   ú   sṹ-ule*   BRIBRI
    1SG   ERG   house see-PTCP
    'I have seen the house.' (EL)

(4) *yís   wã   jũ   sṹ-le*   CABÉCAR
    1SG   ERG   house see-PTCP
    'I had seen the house.' (EL)

In previous literature, it has been argued that the Agent NP followed by *wã* in
Bribri has a *by*-phrase (i.e. oblique) status (Constenla et al. 1998: 112). In con-
trast, I demonstrate that the NP marked by the postposition *wã* has the same
subject properties as an NP marked by the canonical ergative postposition *tö*.[4]
For instance, the latter controls co-reference of a 3rd person anaphoric pro-
noun in a following coordinate clause (5).

(5) *Alí   tö   Trini wöalátsë'   ẽnã ie'   ulú-n-ẽ*   BRIBRI
    A.   ERG   T.   kiss.PFV.REM and 3SG.PRX.H become.angry-MVC-PFV
    'Alí$_i$ kissed Trini$_j$ and he$_i$/*she$_j$ became angry.' (EL)[5]

DST= distal, ERG= ergative postposition, EXCL= exclusive, GEN= genitive, H= human, HAB=
habitual, IDP= ideophone, INF= infinitive, INT= intensifier, IPFV= imperfective, IPFV II= ad-
ditional imperfective suffix expressing habitual or near future meanings, MVC= middle voice
cluster, NEG= negation, NF= non-focusable (in reference to pronouns), OBL= oblique (i.e.
pronoun from the oblique set), PFV= perfective, PL= plural, POS= positional (existential verb
which specifies the position in which its argument is found), POSS= possessive suffix, PRX=
proximal, PSSR= possessor, PST= past, PTCP= participle, REC= recent, REM= remote, SG=
singular.

3 High tone is indicated as <á> and low tone is not marked in the orthography.
4 Bribri is morphologically (i.e. "surface") ergative but syntactically accusative. By 'Subject'
   here I mean the syntactic grouping of S and A, as is most typically found in nominative-
   accusative languages. The emic category of Subject is shown to exist in Bribri by properties
   such as control of co-reference and Subject to Object raising (see Dickeman-Datz 1984).
5 This example and others similar to this one in terms of semantic/ pragmatic oddity have
   been elicited following the work of Hoff (1995: 362). In examples 5 and 6, the pronoun *ie'*
   can refer to a male or female referent. If the person who becomes angry is 'Trini', there are
   two possibilities: either the NP 'Trini' needs to be repeated in the coordinate clause, or the
   demonstrative pronoun *e'* 'that' needs to be used.

The NP marked by *wã* shows the same ability (6).

(6) *Alí wã Trini wöalátsë-ule ẽnã ie'*   *ulú-n-ẽ*   BRIBRI
    A. ERG T.   kiss-PTCP   and 3SG.PRX.H become.angry-MVC-PFV
    'Alí$_i$ has kissed Trini$_j$ and he$_i$/*she$_j$ became angry.' (EL)

The same behavior is observed in Cabécar, as shown in (7) and (8).

(7) *Trini të Severiano wöalats-á   jẽnã jie' ulú-nã*   CABÉCAR
    T.   ERG S.   kiss-PFV.REC and 3SG become.angry-MVC.PFV.REC
    'Trini$_i$ kissed Severiano$_j$ (and) then she$_i$/*he$_j$ became angry.' (EL)

(8) *Trini wã Severiano wöalatsë-le jẽnã jie'*
    Trini ERG S.   kiss-PTCP   and 3SG
    ulú-nã-wã   CABÉCAR
    become.angry-MVC.PFV.REC-ASP
    'Trini$_i$ had kissed Severiano$_j$ and she$_i$/*he$_j$ (had) become angry.' (EL)

In the typological literature, a system with two non-allomorphic and dia-chronically unrelated ergative markers is called "differential ergative marking" (McGregor 2009) and is reported to be quite uncommon (see Arkadiev 2017 for a survey). The ergative marker *wã* also constitutes, in these two Chibchan languages, an instance of case syncretism with the possessor in possessive predications. It is important to stress that the term *possessor* should not be conflated with *genitive* case marking. In this article, the term *genitive* is restricted to case marking in attributive possession (i.e. on a noun within a possessive NP), while the term *possessor* is restricted to the subject of a possessive predication (i.e. a full clause which expresses a possessive relation). In Bribri and Cabécar, the case syncretism is syntactic rather than semantic; the subject possessor bears the morpheme *wã*, cf. (9) and (10), whereas the semantic possessor in attributive possession is unmarked in both languages, cf. (11) and (12).

(9) *ye' wã kró tso'*   BRIBRI
    1SG PSSR rooster exist
    'I have a rooster.' (EL)

(10) *yís wã jóshkoro tsố*   CABÉCAR
    1SG PSSR rooster exist
    'I have a rooster.' (EL)

(11) *sa'*   *ú*   BRIBRI
      1PL.EXCL   house
      'Our house' (EL)

(12) *sá*   *jũ*   CABÉCAR
      1PL.EXCL   house
      'Our house' (EL)

From a diachronic perspective, one might wonder how the synchronic syncretism between a possessor and an ergative marker came to be and what could be the ultimate source of the ergative marker *wã*. In other languages, other case markers are among the most widely attested sources of ergative case markers, followed by demonstratives and pronouns (McGregor 2009: 499). In particular, possessor markers are the third most widely attested source for ergative markers (Palancar 2002: 41). Languages that display this phenomenon are found in the Eskimo-Aleutian, Tibeto-Burman, Mayan and, peripherally, Caucasian language families.

   In the light of such typological considerations, I will demonstrate that the diachronic source of the ergative marker *wã* in the perfect construction is the possessor marker found in possessive predications. Furthermore, based on a syntactic comparison of possessive predications within the Isthmic branch of the Chibchan family, I will argue that the ergative marker *wã* can be traced back to Proto-Chibchan as a possessor marker, whereas the more canonical ergative marker *tö* reconstructs as such to Proto-Chibchan *tV (Constenla 2008: 131).

   The article is organized as follows: §2 offers a genealogical characterization of the Chibchan family and a concise typological sketch of Bribri and Cabécar. §3 is divided into several sub-sections and deals with: (i) the internal reconstruction of the evolution of the possessor marker *wã* into an ergative marker in the perfect construction in Bribri, as well as arguments for directionality (§3.1); (ii) the existence of a parallel reanalysis in Cabécar (§3.2); and (iii) the identification of cognate constructions for the perfect and possessive predication in Bribri (§3.3). §4 traces the presence of the possessor marker in alienable possessive predications in other Chibchan languages of the Isthmic branch. This section shows that the source of the possessor marker *wã* (later reanalyzed as an ergative marker in some of the languages) can be traced back to a Proto-Chibchan word meaning 'something' or 'belonging'. The comparative method will be used to show evidence of cognacy between the proto-form and the reflexes in different languages (details in Appendices A and B). §5 concludes the article.

2 Genealogical Classification of the Chibchan Languages

The Chibchan family was established by means of historical comparative reconstruction by Uhle (1890) and was later confirmed by Constenla (1981, 1988, 1989, 2012) and Holt (1986). Further work applied lexicostatistics to this family (Constenla 1985). Compared to other language families, this family is quite heterogeneous and only a nucleus of basic vocabulary can be reconstructed (Pache 2016). Constenla (1991) places the Chibchan family in what he calls the geographically "Intermediate Area", which is divided as follows: (a) part of the Venezuelan-Antillean Area; (b) the Ecuadorian-Colombian sub-area of the Andes Area; and (c) the Colombian-Central American Area, which is in turn divided into (c1) the Central Sub-area and (c2) the North Sub-area and East Sub-area.

The Chibchan family is part of the Central Sub-area (c1). Figure 6.1 offers a map with the geographical location of modern Chibchan languages comprised in the Colombian-Central American Area (c). The symbol † means that a given language is now extinct.

According to Constenla (1991), in the Colombian-Central American Area (c), there are, among others, the following general linguistic features: (i) exclusively SOV order as the basic order of the transitive clause (as opposed to

1 Paya	13 Kuna
2 Rama	14 Kogi
3 Guatuso	15 Damana
4 †Huetar	16 Ika
5 Cabécar	17 †Atanques
6 Bribri	18 Chimila
7 Teribe/Térraba	19 Barí
8 Boruca	20 Tunebo
9 †Dorasque	21 †Nutabe
10 †Chángena	22 †Duit
11 Guaymí	23 †Muisca
12 Bocotá	
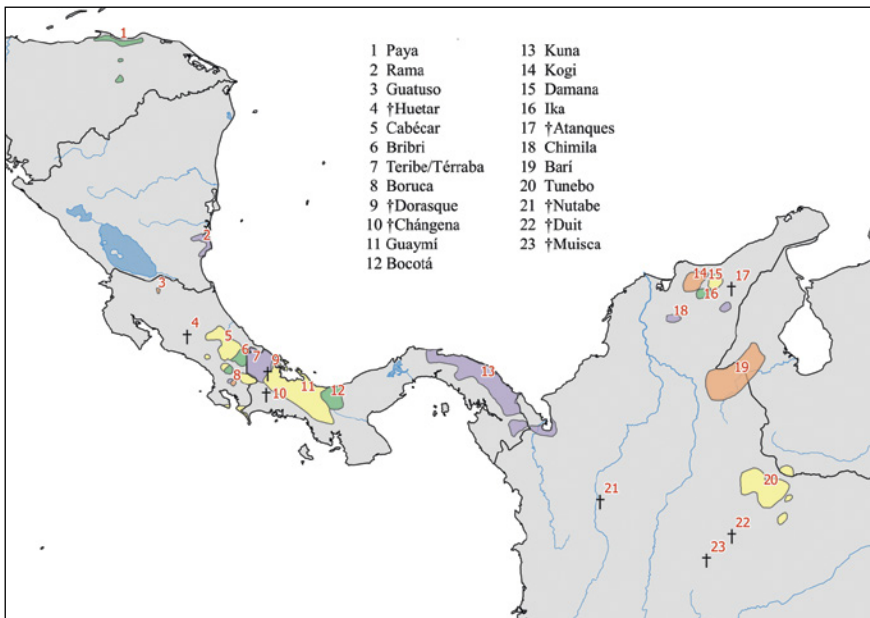
FIGURE 6.1 Approximate distribution of Chibchan languages (taken from Pache 2018: 2)

the Mesoamerican area and the Venezuelan Antillean area); (ii) exclusive use of postpositions; (iii) almost complete absence of gender oppositions in pronouns; (iv) scarcity of inflections; and (iv) absence of accusative case marking. The languages of the Central Sub-area (c1), show, among others, the following features: (i) ergative or active-stative case systems and (ii) absence of inflection for person marking. The marking of genitive case, or possession, by means of suffixation or postpositions is present in a third of the languages. In some, this kind of marking is restricted to some possession relations, usually alienable possession or those in which the possessor is a person.

So far, the reconstruction of Proto-Chibchan has been concerned mainly with phonology and only marginally with nominal and verbal morphology. Constenla (2008: 129 ff.) argues that it is very likely that Proto-Chibchan did not have nominal inflection or, if it did, it was extremely reduced. As for verbal morphology, the following morphemes, among others, have been reconstructed: (i) /*-e/ for imperfective and /*-o/ for perfective aspect; (ii) a marker of non-finite verb form /*-ka/; and (iii) a marker of middle voice /*-de-/. Proto-Chibchan probably had sov order, along with the following orders of constituents: noun-adjective, noun-numeral, noun-postposition and possessor-possessed. As for the case-marking system, Tunebo /ta~t/, Bribri /tʊ/ (<tö>), Cabécar /tɪ/ (<të>) and Guatuso /ti/ favor the possibility, according to Constenla (2008), of reconstructing an ergative postposition /*tV/, with an alternation in the vowel portion.

There have been several versions and revisions of the Chibchan family tree. I will refer to the most recent proposal by Constenla (2008: 127), named 'Paya-Chibchan', in Figure 6.2.

As shown in Figure 6.2, Proto Paya-Chibchan includes four branches: Votic, Isthmic, Magdalenic and Paya (which forms its own branch). The languages that concern the present article are Bribri and Cabécar (Viceitic), and only for purposes of comparison Teribe and Térraba (Naso), and Guaymí and Buglere (Guaymiic), all of which are found in the larger Isthmic branch. Peripherally, this article will also refer to Muisca as well as Cuna, Rama, Damana, Paya and Boruca. The choice of languages for syntactic comparison in §3 relied on two main criteria: (i) the language belongs to the Isthmic branch and (ii) there are sufficient and accessible materials on the language.[6]

---

6   My primary sources for Bribri are Constenla et al. (1998) (CBB) and elicitation with the native consultant Alí García Segura (EL), plus Jara Murillo (1993) (IHB) and Jara Murillo & García Segura (2008) (SOA) as secondary sources; for Cabécar: Margery Peña (1989) (DCE) and the native consultant Severiano Fernández Torres (EL); for Teribe: Quesada (2000) (GT); for Térraba: Constenla (2007) (LT); for Guaymí: Quesada Pacheco (2008) (GG) and Murillo Miranda (2010) (OSN); for Buglere: Quesada (2012) (GB). Besides these, my source for Muisca
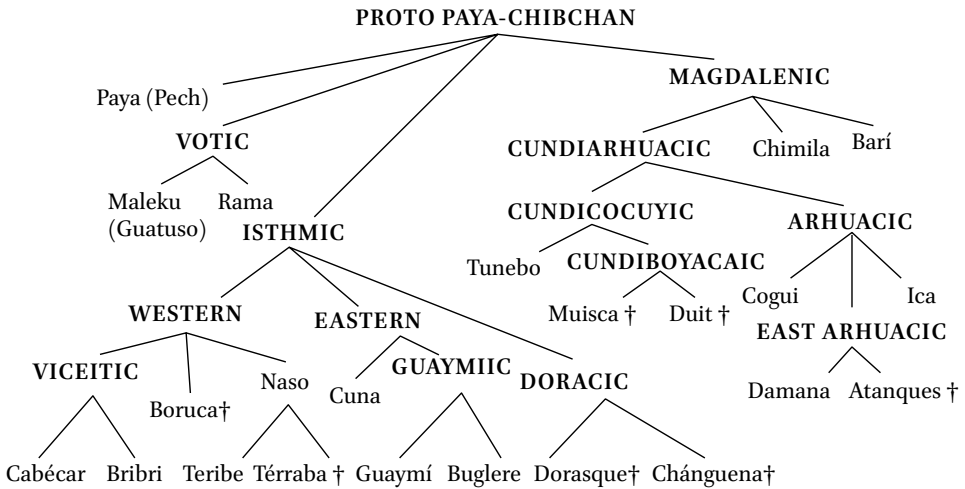
FIGURE 6.2   Macro-family Paya-Chibchan (reproduced from Constenla 2008: 127)

Both Bribri and Cabécar show an ergative-absolutive alignment system in most areas of simple clause syntax. Word order is rigidly PV[7] and the ergative phrase can go either before or after the indivisible PV constituent. In both languages, there is only one set of pronouns for all roles: S/P is always unmarked (both NPs and pronouns) and A is case-marked (both NPs and pronouns). Both languages have almost entirely suffixing verbal morphology, but neither has obligatory indexation of A or S/P arguments in the verb. In Bribri, there is optional bound marking of the absolutive category only when it is a 3rd person plural, animate entity. In Cabécar, verbal morphology can optionally indicate the plurality of S/A in transitive or intransitive clauses and that of P by means of two different suffixes only in certain tenses and aspects. In both languages, the verbal paradigm is based on voice, active vs. middle, and aspect, perfective vs. imperfective. The canonical ergative marker in Bribri is *tö* or its allomorphic variants *dör* and *r*. Other ergative forms are obtained by the contraction of personal pronouns plus the ergative postposition *tö* (such as *yö*, formed by the first person singular pronoun *ye'* plus *tö*). In Cabécar, the canonical ergative marker

---

is mainly Ostler (1994) (STM). Throughout the article, every example includes the abbreviations established in this section along with the page number from where the example was taken, e.g. (GG: 145). The abbreviation is based on the title of a given work (i.e. GG= Gramática de la lengua guaymí).

7   S, A and P are understood here primarily as core syntactic arguments. More specifically, following Comrie (1978), S stands for the sole argument of an intransitive verb; A stands for the most actor-like argument of a transitive verb; P stands for the most patient-like argument of a transitive clause.

is *të* or its variant *te*. Finally, both languages present the additional ergative marker *wã*, which is not an allomorph of *tö* or *të*.[8] Constructions in which *wã* marks the subject in both languages are: possessive predications, the perfect construction, the transitive perfective negative construction and the caused motion construction. In Bribri, the occurrence of this ergative marker is lexically determined with some deponent verbs which take two overt core arguments. In Cabécar, the marker seems to be used in constructions with negative polarity besides the transitive perfective negative. The present article is concerned mainly with the perfect construction and the alienable and inalienable possessive predications, to which we now turn.

3 The Perfect Construction

In the Coroma variety of Bribri presented in this article, the perfect construction is used to express a past event whose relevance continues into the present time, similarly to the English 'have' perfect. Formally, it is a construction in which A is marked by the postposition *wã*, the preverbal P is unmarked, and the verb shows the suffix *-ule* as in (13).

(13) *ye' wã kêkë-pa tsê-ule   ñítöl-ök i'   kũêkĩ*   BRIBRI
      1SG ERG elder-PL listen-PTCP fight-INF this because
      'I have heard the elders fight because of this.' (IHB: 63)[9]

The suffix *-ule* has been previously described in the literature as a "marker of anteriority" (Margery 1989: lxv, Constenla et al. 1998: 91). I would like to suggest that, outside of the perfect construction, the suffix *-ule* functions as a past participle. This suffix derives adjectives from verbs (14) and the resulting forms (i.e. participles) have generally the same distributional properties as simple

---

8   In Bribri, there is a verbal suffix formally identical to the postposition *wã*. According to some authors, this suffix emphasizes the completion of an action (Jara Murillo & García Segura 2009: 137), while according to others (Constenla et al. 1998: 27), it indicates a movement of penetration, and conveys an aspectual value of punctuality. Another formally identical suffix *-wã* can index absolutive animate NPs in the verb. Constenla et al. (1998) claim the existence of another verbal suffix, *-wa*, meaning descending movement or complete affectedness of the absolutive. Finally, there is an unbound form *wa*, which is the instrumental/comitative postposition. Some of these forms may be etymologically related to the postposition *wã* discussed here. Similarly, in Cabécar, *wa* is the instrumental postposition. Moreover, in Cabécar, the verbal suffix *-wã* is used in combination with *-nã* in deriving verbs from adjectives.

9   The glosses of all examples obtained from oral tradition texts are my own.

adjectives (i.e. they can combine with an existential verb[10] and modify an NP, as in (15)).

(14) *ie'pa   ãñíb-ule   dur,   bőka   bërë*   BRIBRI
      3PL   hide-PTCP   exist.POS.SG   two   quiet
      'They are standing hidden, both of them quiet.' (IHB: 95)

(15) *kő   yő-ule-wã   bua'   kő   ã*   BRIBRI
      mouth   do-PTCP-CPL   good   basket in
      'The well sewed (lit: 'done') mouth of the basket' (IHB: 166)

In the perfect construction, if the clause is intransitive, S is not marked for case (16).

(16) *sa'   kapë́-ule   wí   shőő*   BRIBRI
      1PL.EXCL sleep-PTCP   there   IDP
      'We have slept there (points at the mountain).' (SOA: 33)

In seeking cognates for the relevant grammatical morphology of this construction, consider another construction also attested in Bribri. The construction in (17) is formally just slightly different from the transitive perfect construction presented in (13).

(17) *e'   tso'   ie'   wã   sű́-ule*   BRIBRI
      that.DST exist 3SG.PRX.H ERG   see-PTCP
      'He has seen that' (lit: 'that has he seen.').' (IHB: 117)

The construction in (17) also appears in the order in (18).

(18) *ie'   wã   e'   tso'   sű́-ule*   BRIBRI
      3SG.PRX.H ERG   that.DST exist   see-PTCP
      'He has seen that.' (EL)

---

10 In Bribri, there are two existential verbs: *ta'* and *tso'*. Additionally, there is a considerable number of existential verbs which specify the position in which the absolutive argument is found (seated, standing, suspended, lying down, floating, being vertically attached, buried, etc.) (see Constenla et al. 1998: 67, Jara Murillo & García Segura 2009: 91). There is also an invariant copula, *dör*, formally distinct from the existential verbs.

Moreover, it is possible to use this construction with verbs which imply a change of state, as in (19).

(19) *ie'*   *wã*   *kró*   *tso'*   *kṍt-ule*   BRIBRI
      3SG.PRX.H   ERG   rooster   exist   kill-PTCP
      'He has a killed/dead rooster.' (EL)

The only formal difference between (13) and (17) is the presence in (17) of the existential *tso'*. Except for the order in which they appear, all other components are identical: the ergative is marked by *wã*, the absolutive is unmarked and precedes the existential and the past participle form is stranded at the end of the clause, after the ergative phrase. The meaning of the construction in (17) is perhaps better understood by looking at (19), in which the same construction is used with a change-of-state verb. The meaning of (19) is close to English 'he has a rooster killed', or better said, 'he possesses a rooster that is in the state of having been killed', where the possessor 'he' might or might not be the person who killed the rooster.

Taking into account the order presented in (18) and (19), the constructions in (13) and (15) can be schematized as in Figure 6.3 and Figure 6.4, respectively. The subscripted *i* indicates that the participial form refers to the state in which the absolutive NP is found.

In §3.1, I will argue that construction A in Figure 6.3 is a later development from construction B in Figure 6.4 and that both originated in a possessive predication. This explains why the case-marking of the perfect construction, in which A is marked by *wã*, differs from that of other main clause transitive events, where A is usually marked by *tö*.

[NP]$_{ERG}$	[NP]$_{ABS}$	V$_{TR\_PTCP}$
NP *wã*	NP$_i$-Ø	V$_{TR}$-*ule*$_i$

FIGURE 6.3 Construction A (cf. 13)

[NP]$_{ERG}$	[NP]$_{ABS}$	EXIST	V$_{TR\_PTCP}$
NP *wã*	NP$_i$-Ø	*tso'*	V$_{TR}$-*ule*$_i$

FIGURE 6.4 Construction B (cf. 15)

3.1   *From Possession to Perfect Aspect: Bribri*

The purpose of this section is to show that the alienable possessor subject marker *wã* found in alienable possessive predications was re-analyzed as an ergative marker in the perfect construction.[11] Therefore, the evolution of a perfect construction from a possessive construction is the mechanism responsible for the innovation of the construction-determined ergative marker *wã*. The following are the steps that would have taken place, all still attested in Modern Bribri.

Stage 0: possession. In alienable possessive predications (20), the verb used in the construction is the existential auxiliary *tso'* and the possessor must be marked by the postposition *wã*.

(20) *Alí   wã   kró   tso'*   BRIBRI
     A.   PSSR   rooster   exist
     'Alí has a rooster.' (lit: 'Alí's rooster exists.') (EL)

The possessor marked by the postposition *wã* in the alienable possessive predication displays subject properties, just as any S/A argument would. First, it controls co-reference of a 3rd person anaphoric pronoun in a following possessive NP. In (21), 'house' can refer only to the house of the possessor.

(21) *ie'   wã   yámĩ   tso'   ie'   ú   ã*   BRIBRI
     3SG.PRX.H   PSSR   friend   exist   3SG.PRX.H   house   in
     'He$_i$ has a friend$_j$ in his$_i$ house.' (EL)

If the house belongs instead to the possessed, the NP 'friend' must be repeated (22). Alternatively, the demonstrative *e'* 'that' could be used.

---

11   An anonymous reviewer suggested to me that, as the discussion immediately below will show, the NP marked by Possessor marker *wã* behaves as an external possessor. This means that the Possessor NP marked by *wã* is not a sub-constituent of the possessed NP, but a distinct constituent. Although I acknowledge the validity of this observation, I will refrain from calling *wã* an external possessor only to avoid possible terminological confusion with extended meanings related to external possession in the typological literature (see for instance, cases of possessor raising included under the broad term 'external possession' in Payne & Barshi 1999).

(22) *ie'*   *wã*   *yámĩ* *tso'* *yámĩ* *ú*   *ã*   BRIBRI
3SG.PRX.H PSSR friend exist friend house in
'He_i has a friend_j at his_j house.' (EL) (lit: 'He has a friend at the friend's house.')

Second, the NP marked by *wã* controls co-reference under coordination as an A argument marked by the ergative postposition *tö* would (23).

(23) *ie'*   *wã*   *yámĩ* *bák*   *bua'* *ie'*   *ú*   *ã* BRIBRI
3SG.PRX.H PSSR friend exist.PST good 3SG.PRX.H house in
ẽnã   *ie'*   *mĩ́-n-ẽ*
and   3SG.PRX.H   go-MVC-PFV
'He_i had a good friend_j at his_i house, and he_i left.' (EL)

If 'the good friend' is the one who leaves the house, there are again two options: the repetition of the NP 'friend' in the coordinate clause or the use of the demonstrative *e'* as in (24).

(24) *ie'*   *wã*   *yámĩ* *bák*   *bua'* *ie'*   *ú*   *ã* BRIBRI
3SG.PRX.H PSSR friend exist.PST good 3SG.PRX.H house in
ẽnã   *e'*   *mĩ́-n-ẽ*
and   that.DST   go-MVC-PFV
'He_i had a good friend_j at his_i house and he_j left.' (EL)

It should be noted that in (23) and (24), *wã* is an integral and essential part of the construction. If the first coordinate clause were introduced without *wã*, the possessive meaning would be lost and only a predicate attributive reading would be possible (25).

(25) *ie'*   *yámĩ*   *bák*   *bua'*   BRIBRI
3SG.PRX.H friend   exist.PST   good
'His friend was good.'

The necessity to repeat the NP which is not the possessor to disambiguate cases of co-reference when both A and P are 3SG can be observed also in the caused motion construction.

(26) *Alí* *wã* *Serveriano* *mĩ́-n-ẽ*   *ie'*   *ú*   *ã*   BRIBRI
A. ERG S.   go-MVC-PFV 3SG.PRX.H house in
'Alí_i took Severiano_j to his_i house.' (EL)

The only possible interpretation in (26) is that the house is Alí's, which is the NP marked by *wã*. If the house belongs to Severiano, the proper name needs to be repeated in the possessive NP as in (27).

(27) *Alí wã   Severiano mĩ́-n-ẽ   Severiano ú   ã   BRIBRI*
     A. ERG S.   go-MVC-PFV   S.   house in
     'Alí_i took Severiano_j to his_j house.' (lit: 'to Severiano's house') (EL)

Thus, the possessor NP marked by *wã* displays the subject properties typical of Bribri. In the alienable possessive predication (Figure 6.5), the possessor NP and the possessed NP do not form a single constituent, because the possessor is capable of controlling the reference of an anaphoric pronoun in a conjoined clause, while the possessed NP cannot.[12]

It is worth mentioning that the construction in Figure 6.5 originates from a source other than a transitive predication (i.e. an existential genitive construction of the type 'X's Y exists') and acquires transitive features (i.e. control of co-reference of the subject NP). In this respect, Creissels (2014) argues that diachronically, the transitivization of non-transitive existential possessive predications (including the "genitive" subtype) is widely attested, for instance, in Maltese, Israeli Hebrew and Amharic (cf. also Stassen 2009: 208 ff.). In particular, according to Creissels (2013: 469) the possessed, unlike the figure in a spatial relationship, has some similarities to a patient of a prototypical transitive predication in terms of control. For this reason, if an existential construction extends its use to possession, the syntactically ambiguous zero marking of the possessed is reanalyzed as encoding object rather than subject function.

[NP]_PSSR	[NP]_PSSD	EXIST
NP *wã*	NP-Ø	*tso'*

FIGURE 6.5 Alienable possessive predication

---

12   An anonymous reviewer observed that the external possessor status of the NP marked by *wã* can be contrasted with internal possessors in English. Attributive (NP-internal) possessors in English are not able to bind reflexives. In [*John's father*] *saw himself in the mirror*, 'himself' can refer only to *father*, and never to the NP-internal possessor *John's*. This clearly contrasts with the subject properties displayed by the subject possessor marked by *wã*.

Stage 1: possession of the possessed NP in a modified state. In this stage, the past participle is added after the existential. This verbal form occupies the position of an adjective: both follow the existential verb. As happens in stage 0, in stage 1 a primary possessive interpretation of the construction is possible if the existential is followed by an adjective, as shown in (28).

(28) *ie'*   *wã*   *kró*   *tso'*   *ôjkẽ́*   BRIBRI
     3SG.PRX.H   PSSR   rooster   exist   fat
     'He has a fat rooster.' (EL)

Therefore, in this stage, the alienable possessive predication in stage 0 can be expanded as shown in Figure 6.6. The adjective after the existential specifies a property of the possessed NP, as indicated by the subscript *i*.

[NP]$_{PSSR}$	[NP]$_{PSSD}$	EXIST	ADJ
NP *wã*	NP$_i$-Ø	*tso'*	ADJ$_i$

FIGURE 6.6   Alienable possessive predication modified by an adjective

The function of the past participle in this construction is identical to that of an adjective in that it modifies the state in which the possessed NP is found, as in (19) repeated as (29).

(29) *ie'*   *wã*   *kró*   *tso'*   *kṍt-ule*   BRIBRI
     3SG.PRX.H   PSSR   rooster   exist   kill-PTCP
     'He has a killed/dead rooster.' (EL)

In this stage, the Possessor marked by *wã* need not be co-referential with the Agent of the event that led to the state: the possessor might or might not be the one who caused the death of the rooster. A close translation of (29) would be 'he possesses a rooster that is in the state of having been killed'. An example of possessor co-referential with the Agent of the event appears in (30).

(30) *ie'*   *wã*   *kró*   *tso'*   *sṹ-ule*   BRIBRI
     3SG.PRX.H   PSSR   rooster   exist   see-PTCP
     'He has a rooster seen.' (EL)

Conceivably, in this stage the existential starts to be reinterpreted as an auxiliary: it no longer plays the role of expressing possession as in stage 0. Rather, it accompanies the participial form carrying the semantic content of the verb ('kill', 'see'). The construction with the participial verb form in (29) and (30) (cf.

[NP]$_\text{PSSR}$	[NP]$_\text{PSSD}$	EXIST	V$_\text{TR\_PTCP}$
NP *wã*	NP$_i$-Ø	*tso'*	V$_\text{TR}$_*ule$_i$*

FIGURE 6.7   Alienable possessive predication modified by a participle

[NP]$_\text{ERG}$	[NP]$_\text{ABS}$	AUX	V$_\text{TR\_PTCP}$
NP *wã*	NP$_i$-Ø	*tso'*	V$_\text{TR}$-*ule$_i$*

FIGURE 6.8   Perfect construction with auxiliary

Figure 6.7) is identical to the alienable possessive predication modified by an adjective (Figure 6.6).

Probably, this is the transitional stage in which reanalysis covertly takes place. The reanalysis is prompted most likely by cases such as 'he has a rooster seen' in which the owner of the rooster and the 'seer' must be the same person. Then, co-referentiality is generalized to ambiguous cases such as 'he has a killed rooster/he has killed a rooster'.[13] Cases such as 'he has a rooster seen' favor the interpretation in which the owner of the rooster and the 'seer' are the same person. Thus, the construction in Figure 6.7 becomes re-analyzed as follows.

Stage 2: from modified possession to perfect aspect. In this stage, the possessive interpretation still possible in stage 1 (cf. (28), (29)) is no longer available. The NP marked by *wã* is necessarily the Agent, and not necessarily a possessor. The invariant existential auxiliary is dropped, probably because the participial form already carries the semantic information of the event being described. Thus, example (30) of stage 1 'he has a rooster seen', meaning 'he possesses the rooster and he is the one who saw the rooster', comes to mean 'he has seen a rooster' (31).

(31) *ie'*   *wã*   *kró*   *sṹ-ule*   BRIBRI
      3SG.PRX.H   ERG   rooster   see-PTCP
      'He has seen a rooster.' (EL)

---

13   The (possibly odd) English translations 'he has a killed rooster' and 'he has a rooster seen' are not meant to suggest a causative interpretation, but only a 'perfect' interpretation. The constructions presented so far can never be used in Bribri to express a causative event.

[NP]$_{ERG}$	[NP]$_{ABS}$	V$_{TR}$
NP *wã*	NP$_i$-Ø	V$_{TR}$-*ule*$_i$

FIGURE 6.9   Perfect construction without auxiliary

The meaning of the construction at this point is only about aspect: it conveys the idea of an action that occurred in the past, but with relevance in the present. The perfect construction with the auxiliary of stage 1 (Figure 6.8) becomes the perfect construction without the auxiliary in Stage 2 (Figure 6.9).

As in previous stages, in stage 2 the ergative NP marked by *wã* continues to display Subject properties, such as control of co-reference under coordination (32).

(32) *Alí wã aláköl pakló-ule ẽnã ie'   ulú-n-ẽ*   BRIBRI
A. ERG woman hug-PTCP and 3SG.PRX.H become.angry-MVC-PFV
'Alí$_i$ has hugged the woman$_j$ and (then) he$_i$/*she$_j$ became angry.' (EL)

An additional test which shows a clear S/A pivot in some areas of Bribri's syntax, and thus serves to show subject properties of the NP marked by *wã*, is found outside main clauses (Dickeman Datz 1984: 124 ff.). An embedded complement clause (CC) in absolutive position (33) can be moved to the right, leaving in its place the resumptive absolutive pronoun *i* (34).

(33) *ye' wã [Alí wã kró   kṍt-ule   e']$_{CC}$ sṹ-ule*   BRIBRI
1SG ERG A. ERG rooster kill-PTCP that.DST see-PTCP
'I have seen that Alí has killed the rooster.' (EL) (lit: 'I (that) Alí has killed the rooster that I have seen.')

(34) *ye' wã ĩ-sṹ-ule   [tö   Alí wã kró   kṍt-ule]$_{CC}$* BRIBRI
1SG ERG 3SG-see-PTCP COMPL A. ERG rooster kill-PTCP
'I have seen that Alí has killed the rooster.' (EL) (lit: 'I have seen it that Alí has killed the rooster.')

The S/A argument of the postposed clause, i.e. Alí in (34), can be raised to become the absolutive within the matrix clause (35). When this happens, the S/A argument of the complement clause, which has become the absolutive in the matrix clause, leaves a zero in the complement clause and the verb appears in its infinitive form.[14]

---

14   This is true only for certain verbs of perception and cognition (i.e. 'see'). Other verbs retain their conjugated form in the complement clause after the S/A argument has been raised.

(35) *ye' wã Alí sű-ule*   [∅   *kró*   *tt-ők*]_{CC}   BRIBRI
     1SG ERG A.   see-PTCP   rooster kill-INF
     'I have seen Alí killing roosters.' (EL)

As shown in examples (33) to (35), the NP marked by the ergative postposition *wã* behaves in the same way as an NP marked by the ergative postposition *tö* would (see Dickeman Datz 1984: 123 ff. for equivalent examples with *tö* in the transitive complement clause and for examples with an unmarked S in the intransitive complement clause).

The evolution of the perfect construction from the alienable possessive predication seems to be a recent innovation in Bribri: all the stages of the reanalysis are still synchronically present in the language. The stages and constructions of each stage are summarized in Table 6.1.

TABLE 6.1   Evolution of the perfect construction from the alienable possessive predication

Stage	Meaning	Cognate constructions			
0	Possession (*wã* marks the PSSR subject) possession + modifier (adj)	[NP]_{PSSR} NP *wã*	[NP]_{PSSD} NP-∅	EXIST *tso'*	
		[NP]_{PSSR} NP *wã*	[NP]_{PSSD} NP_i-∅	EXIST *tso'*	Adj Adj
1	Possession of the possessed NP in a modified state (PTCP) Resultative meaning *Possible* co-referentiality of A and PSSR *wã* marks the PSSR subject who might also be A	[NP]_{PSSR} NP *wã* [NP]_{ERG} NP *wã*	[NP]_{PSSD} NP_i-∅ [NP]_{ABS} NP_i-∅	EXIST *tso'* AUX *tso'*	V_{TR_PTCP} V_{TR}_*ule_i* V_{TR} V_{TR}_*ule_i*
2	Modified possession becomes perfect aspect *wã* marks A, which might also be PSSR Loss of existential/auxiliary Only possible reading is past tense with present relevance	[NP]_{ERG} NP *wã*	[NP]_{ABS} NP_i-∅	V_{TR} V_{TR}_*ule_i*	

The main argument to support the directionality presented in this section is based on similar attested directions of change. In this respect, Heine & Kuteva (2006) argue for the evolution of perfects from possessive constructions based on the Action schema [X has Y] in several branches of Indo-European (Italic, Germanic, Celtic, Albanian, partially Slavic and Baltic) and in some non-Indo-European languages spoken in Europe (e.g. Finnish). Other examples in the literature of perfects diachronically deriving from possessive constructions include: Benveniste (1966), who describes the origin of the transitive perfect in Classical Armenian from a possessive construction, and Haig (2008), who argues for the emergence of ergativity in Middle Iranian from external possession.

Along the same line, Trask (1979: 398) claims that ergative syntax has arisen from possessive sources in Kurdish, Middle Persian, Old Armenian, Eskimo-Aleut and North and South Caucasian languages. In particular, Trask argues that while 'Type A' ergativity results from a passive made obligatory, 'Type B' ergativity "results from the incorporation into the inflectional paradigm most often of a stative de-verbal adjective, incorporated by means of a possessive construction" (1979: 402). More specifically, in Trask's view, Type B ergativity depends on the existence, in a given language, of a de-verbal adjective, sov word order (this correlation remains unexplained) and the lack of a verb 'have'. Crucially, in European languages in which stative de-verbal adjectives became part of the verbal system through initially possessive predications, no Type B ergativity arose due to the presence of a verb 'have' in the construction, whose possessor was coded as nominative and remained such when it was reinterpreted as subject in the new perfect construction.

However, according to Trask, in languages that lack a verb 'have', possessive predications very often feature an overtly marked possessor expressed by an oblique case – usually genitive, dative or locative. The reanalysis of such a possessor marker as an agent would then bring about ergative case marking. Therefore, a clause of the shape *To me/of me/at me* (*is*) *a window broken* being reinterpreted as 'I have broken a window' would inevitably lead to Type B ergativity. This depicts exactly the features of Bribri, a language with a participle which can function as a de-verbal adjective, no 'have' verb, sov order and an overtly marked possessor in possessive predications.[15]

A second main argument in support of the directionality outlined here is that in closely related languages such as Cabécar, a similar pattern of reanalysis can be observed. Arguably, the long-lasting language contact and

---

15   While beyond the scope of the present article, it should be noted that Type B ergativity has been extended to several other constructions in this language where it now competes with the canonical ergative marker *tö*.

intermarriage between these two communities might have favoured a simi-lar reanalysis.[16]

3.2   *From Possession to Perfect Aspect: Cabécar*

In Cabécar, *wã* marks both alienable (36) and inalienable (37) possessor subjects.

(36) *yís   wã   jóshkoro   tsố*   CABÉCAR
      1SG   PSSR   rooster   exist
      'I have a rooster.' (EL)

(37) *yís   wã   kutá   tsố   tkế̃l*   CABÉCAR
      1SG   PSSR   sister exist four
      'I have four sisters.' (DCE: 310)

The meaning of past perfect or pluperfect tense/aspect is expressed by the fol-lowing construction.[17]

(38) *yís   wã   i   sṹ-le*   CABÉCAR
      1SG   ERG   3SG   see-PTCP
      'I had seen it.' (EL)

The construction in (38) is identical in all its constituent parts to that of Bribri (cf. (31)). The meaning, however, is past rather than present perfect. Don Severiano Fernández Torres (p.c.) indicates that a construction such as (39) is most likely to be used by elders nowadays rather than by younger generations. This is because the construction is somewhat archaic due to the presence of the existential *tsố*. The meaning of (39) is closer to 'I had it in the state of being seen'.

---

16   Constenla & Ibarra (2009) reconstruct the approximate distribution of the indigenous languages and people of Costa Rica and border areas of Nicaragua and Panama before the arrival of the Spanish conquerors. Bribri and Cabécar have been spoken in contiguous areas of the Talamanca mountain chain since the XVIth century and still are nowadays.

17   The examples offered by Margery (1989) of perfect constructions in Cabécar are different in structure from those presented here. I have presented a native speaker of Cabécar (Don Severiano Fernández Torres) with the data concerning perfect constructions found in Margery (1989: lxxix) and he expressed that those are not in fact perfect constructions but rather constructions with an aspectual meaning that differs from the perfect. The Cabécar examples in this section are the result of elicitation with Don Severiano Fernández Torres. It must be noted that Don Severiano Fernández Torres also speaks Bribri.

(39) *yís*   *wã*   *i*   *tsṍ*   *sṹ-le*   CABÉCAR
    1SG   ERG   3SG   exist   see-PTCP
    'I had it seen.' (EL)

By comparing (38) and (39) with Bribri examples presented in the previous sections, it is possible to identify the following lexical and morphological cognates: (i) in both languages the possessor marker is *wã*; (ii) the existential is *tso'* in Bribri and *tsṍ* in Cabécar; and (iii) the morpheme suffixed to the verb to make a participle is *-ule* in Bribri and *-le* in Cabécar. The order in which the elements appear in the perfect constructions is also identical: A precedes P which is followed by a verb phrase composed of the existential (later dropped) plus the participial form of the verb. Therefore, in a language with a possessive predication identical to that of Bribri, we also find an identical perfect construction which conveys a past perfect meaning. This suggests a coincident path of reanalysis. In Cabécar, the source construction could have been either the alienable or inalienable possessive predication, given that both host a marked possessor.

However, the reanalysis in Cabécar yielded a construction which conveys past perfect meaning. The present perfect construction in this language (40) raises questions.

(40) *yís*   *të*   *i*   *sṹ-le*   CABÉCAR
    1SG   ERG   3SG   see-PTCP
    'I have seen it.' (EL)

This construction is formally identical to the one presented in (38), which expresses a past perfect meaning, except that it contains the canonical ergative marker *të* instead of *wã*. As in the case of the past perfect construction, (41) is possible albeit archaic according to the synchronic intuition of the speaker.

(41) *yís*   *të*   *i*   *tsṍ*   *sṹ-le*   CABÉCAR
    1SG   ERG   3SG   exist   see-PTCP
    'I have it seen.' (EL)

The difference between present and past perfect constructions in Cabécar relies on the choice of the ergative marker: *wã* (38–39) for past perfect and *të* (40–41) for present perfect. As in Bribri, a participle which becomes a main verb, as in (38) and (40), has its origin in an originally more complex construction, which featured an existential/auxiliary. The construction in (41) possibly originated on the basis of the structure in (39). The post-reanalysis

non-canonical ergative marker *wã* was replaced via analogical extension: the canonical ergative marker *të* extends into the construction. This created two competing constructions from a single source, which somehow acquired two distinct meanings, i.e. past and present perfect. Afterwards, the existential was dropped and the participle became the main verb in both constructions. The shift of meaning from present to past perfect for the construction that uses *wã* as an ergative marker (38) is not fully comprehended at this point, but it is presumably due to competition with the perfect construction that has the canonical ergative marker *të*.

As expected, evidence from possessive predications and perfect constructions found in the sister language Cabécar corroborates the direction of re-analysis posited for Bribri in §3.1. These facts invite the assumption that before the split into two distinct languages, Proto-Viceitic had a perfect construction such as that in Figure 6.10.

Obviously, there must also have been in the ancestor language some sort of possessive predication in which the adjective or the participial verb form modifying the state of the possessed NP was optional, as in Figure 6.11.

[NP]$_{ERG}$	[NP]$_{ABS}$	EXIST	V$_{TR\_PTCP}$
NP *wã	NP$_i$-Ø	*tsõ	V$_{TR\_}$*ule$_i$

FIGURE 6.10   Perfect construction of Proto-Viceitic[18]

[NP]$_{PSSR}$	[NP]$_{PSSD}$	EXIST	(ADJ$_i$/V$_{TR\_PTCP}$)
NP *wã	NP$_i$-Ø	*tsõ	(ADJ$_i$/V$_{TR\_}$*ule$_i$)

FIGURE 6.11   Optionally modified possessive construction of Proto-Viceitic

18   The reconstruction of *tsõ is disputable. The problem is obviously the final vowel, which is nasalized and with high tone in Cabécar (*tsǿ*), but oral in Bribri, followed by a glottal stop (*tso'*). According to Constenla (1981: 183), the reflexes of proto-nasality were maintained in all environments in both languages. As a consequence, the loss of nasality in Bribri needs to be explained. A possible explanation is offered by Constenla, who argues that cases in which a language presents the expected reflex of proto-nasality and the other does not "can be explained as resulting from analogic changes which have incorporated a stem derived by means of /*˜/ into the class of stems consisting of a root without any stem formative or of a root plus another stem formative, or vice versa" (1981: 181). Thus, my proposed reconstruction assumes that Bribri *tso'* is the result of analogical replacement of the proto-nasality stem formative by the /-ʔ/ stem formative.

The conclusion reached at the end of this section is that the ergative marker *wã* arose from the possessor marker *wã*. The next section intends to show that this evolution can be claimed through undisputable cognacy among constructions.

3.3   *Corroborating Actual Cognacy among Constructions*

After the analysis put forth in §3.1 and §3.2, one might argue that this is merely a case of seeing similarities within constructions and imposing a hypothetical scenario of reanalysis based on constructions that are not demonstrably cognates. The purpose of this section is to show that in fact this argumentation can be proven correct if we examine each of the individual pieces which appear in the constructions under investigation. We started out with construction A (Figure 6.3), i.e. the perfect construction without an auxiliary (repeated as Figure 6.9), and construction B (Figure 6.4), the perfect construction *with* an auxiliary (repeated as Figure 6.8).

Two fundamental pieces that constructions A and B have in common are the ergative marker *wã* and the suffix *-ule*. We should then inquire where else in the grammar we find modern reflexes of each of these two components. We saw that the latter component is a participle which behaves much like an adjective, cf. (14) and (15). As for the former component, the postposition *wã* is found in the alienable possessive predication (Figure 6.5), which also features the existential *tso'*. This construction can be modified by an adjective (Figure 6.6) or by a participle (Figure 6.7). The construction in Figure 6.7 is formally identical to the construction in Figure 6.4.

A fair question to posit at this point is the nature of the relationship between all these constructions. Given the extensive similarity in form, one possibility is that they are simply synchronic variants of one another, that is, they are all synchronically the same construction. However, the semantics of the perfect construction (Figure 6.3, Figure 6.9) are quite different from the possessive semantics of the other constructions (Figures 6.4–6.8), and this shift in semantics correlates with the presence versus absence of one formal element, the existential verb *tso'*. As such, it appears necessary to posit at least two distinct constructions synchronically. On the other hand, their near identity in form, both morphological and syntactic, can hardly be considered a product of chance – they must come from a common origin, that is, they must be cognate.

Accepting that constructions A and B share a common source, there are three logical possibilities: (i) the alienable possessive predication came from the perfect construction; (ii) the perfect construction came from the alienable possessive predication; or (iii) both the alienable possessive predication

and the perfect construction came from some third source. The third possibility remains out of the available evidential range. The other two possibilities entail different directions of change. A reasonable way to provide an answer to the issue of directionality is to consider which directions of change are attested in the literature when both perfect constructions and some sort of possessive predications are cognate. As argued in §3.1, the only attested direction of change points to possibility (ii). By relying on this attested direction of change, §3.1 has posited directionality for the cognate constructions presented in this section.

Now that the synchronic syncretism between the possessor marker and the ergative marker has been explained, we can turn to the second question set forth in the introduction: the ultimate source of the Possessor marker *wã*. The underlying possession schema of the alienable possessive predication in Bribri and Cabécar seems to be the Genitive schema [X's Y exists] (Heine 1997). However, there is not any readily available synchronic source for this possessor marker in Bribri or Cabécar. This means that *wã* does not serve, synchronically, other functions besides marking possessor and ergative case in these two languages. This postposition is *not* used, for instance, to express Source, Location, Goal, Comitative or Instrumental meanings, which would possibly have linked it to other possession schemas proposed by Heine (1997).[19]

The absence of an expected synchronically available source in both languages invites comparison among other languages of the Isthmic branch. The purpose of the next section is to discover, on the basis of historical and comparative evidence, if the postposition *wã* is found in cognate forms in languages within the larger Isthmic branch, and if so, with what functions.

4   Alienable Possessive Predications in the Isthmic Branch

The alienable possessive predication cognate in Bribri and Cabécar clearly reconstructs to Proto-Viceitic (see §3.2). Now the question is whether there

---

19   In Bribri, Source is expressed by the postposition *mĩk*, Location and Goal by the postposition *ã*, Comitative meaning by the postposition *tã* and instrumental by *wa*. In Cabécar, Goal is expressed by the postposition *ĩã*, Location is expressed by the postposition *ska* or *na*, Comitative meaning by the postposition *da*, Instrumental by *wa* and Source by the postposition *mĩ*. It is evident that the instrumental postposition *wa* in both languages looks a lot like *wã*. However, nasality was distinctive in Proto-Chibchan (Constenla 1981) and languages that have a cognate form of *wã*, such as Guaymí, have *be* and *biti* as a comitative and instrumental postpositions, respectively (Quesada Pacheco 2008: 81).

exists a similar construction in other languages within the Isthmic group and whether such a construction features a cognate form of the Possessor Subject marker *wã*.

As the attentive reader will know by now, the alienable possessive predication of Bribri looks like (42).

(42) *ye'* *wã* *báka* *tso'* BRIBRI
1SG PSSR cow exist
'I have a cow.'

In this respect, it should be noted that Constenla & Margery (1979: 33) and Constenla et al. (1998: 105) report for Bribri an example identical to (42) but with the verb *tã'* 'exist' as the main predicate. However, Alí García Segura (p.c.) argues that if *tã'* is used, the construction in (42) would be slightly unusual in terms of its semantics. It would mean something close to 'I have one exemplar of cow and I am not going to sell it, rather I keep it to show it to people'. Because of this semantic awkwardness, it seems that in the Coroma variety of Bribri the verb used in alienable possessive predications can be only the existential *tso'* or its positional variants, i.e. *tchër* in (43). The alienable possessor subject must be marked by the postposition *wã*.

(43) *pë'* *wã* *kró* *tchër* BRIBRI
person PSSR rooster exist.POS.SG
'The person [the king of the dogs] had a rooster.' (IHB: 198)

In Cabécar, on the other hand, both existentials *tã* and *tső* can be used to express alienable possessive relations. In this case too, the alienable possessor subject is marked by means of the postposition *wã*, as in (44) and (45).

(44) *yís* *wã* *nã́glö* *tső* *tãĩ́* CABÉCAR
1SG PSSR money exist much
'I have a lot of money.' (DCE: 35)

(45) *yís* *wã* *nã́glö* *tã* *tãĩ́* CABÉCAR
1SG PSSR money exist much
'I have a lot of money.' (DCE: 35)

In Guaymí, alienable possessive predications are verbless. The orders possessed-possessor or possessor-possessed (Quesada Pacheco 2008: 72) can

be used to express alienable possession. With both orders, the possessor is marked with the genitive case marker *gwe*, cf. (46) and (47).[20]

(46) *Chiti   gwe   nu   krämä*   GUAYMÍ
     Chiti   PSSR   dog   three
     'Chiti has three dogs.' (or: 'the three dogs of Chiti') (GG: 141)

(47) *mädä   bänrabe   ti   gwe*   GUAYMÍ
     horse   nice   1SG   PSSR
     'I have a nice horse.' (or: 'my nice horse') (GG: 141)

There is a third strategy in Guaymí to mark alienable possession: the order possessor-possessed, in which the possessed is marked by a morpheme identical to the dative marker, namely *e*. However, (46) and (47) are the only two possibilities listed by Quesada Pacheco (2008) to build an alienable possessive *predication* in Guaymí. This seems to suggest that the Guaymí construction possessor possessed followed by *e* is used only in attributive possession and not in possessive predication.[21]

In Buglere, existential and possessive constructions are built without any verb overtly expressed in the clause (Quesada 2012: 85). In this language, the alienable possession relation can be expressed by two different strategies: (i) with the order possessor-possessed plus a possessive suffix *-a~-ia* on the possessed (48); or (ii) with the order possessor-possessed plus the genitive postposition *no* marking the possessor (49).

(48) *cha   chunu-a*   BUGLERE
     1SG   pig-POSS
     'My pig' (or: 'I have a pig.') (GB: 75)

(49) *chunu   cha   no*   BUGLERE
     pig   1SG   GEN
     'My pig' (or 'I have a pig.') (GB: 75)

---

20   The grapheme ä corresponds to [ɔ] in Guaymí's orthography.
21   Miguel Angel Quesada Pachecho informs me that several Chibchan languages have a postposed genitive marker in attributive possession (i.e. in possessive NPs). This marker is similar in several languages, among these: Guaymí *e* (or *i*), Teribe *ĩ* (see below), Cabécar and Bribri *-i* (sometimes with vowel harmony, cf. Bribri *kal* 'tree' > *keli* 'a tree of X') and Muisca *-e*. This issue will not be addressed here, as the section deals with possessive predications.

In other languages of the Isthmic group, such as Teribe and Térraba, the alienable possessor is not marked by any postposition. In Teribe (Quesada 2000), there are two sets of personal pronouns: "nominative" and "oblique". The possessive pronouns used in attributive possession are the oblique pronouns (50). However, in order to have a possessive predication, the pronouns from the nominative paradigm must be used for the possessor subject (51). The order in attributive possession and the possessive predication is possessor-possessed. As in Guaymí and Buglere, possessive predications do not contain an overtly expressed verb.

(50) *bor*   *u*   TERIBE
       1SG.OBL   house
       'My house' (adapted from GT: 55)

(51) *ta*   *u*   TERIBE
       1SG   house
       'I have a house.' (adapted from GT: 141)

Another strategy used only to express alienable possession in Teribe is the marker *ĩ*, which functions also as a copula. This marker is used only with full NPs when the possession relation being predicated is alienable as in (52). In this case, the order is possessed-possessor.

(52) *kwomgla*   *e*   *Juan*   *ĩ*   TERIBE
       horse   that   Juan   PSSR
       'The horse, it is Juan's.' (GT: 139)

In Térraba, the same phenomenon described for Teribe occurs: one of the two sets of pronouns present in the language is used for attributive possession but not for possessive predication. This set, called "non-focusable" (Constenla 2007), appears to be cognate with Quesada's (2000) "oblique" set. According to Constenla (2007), in the case of alienable possessive predications both the "focusable" (Quesada's "nominative") and "non-focusable" (Quesada's "oblique") sets of pronouns can be used with different nuances in meaning, as in (53) and (54). Finally, the existential in possessive predications is often dropped (55).

(53) *t'a*   *cuomgrá*   *t'óc*   *c'ue*   *shcó*   TÉRRABA
       1SG   horse   exist   that   in
       'I have horses there.' (LT: 128)

(54) *bor   cuomgrá   t'óc   c'ue   shcó*   TÉRRABA
     1SG.NF   horse   exist   that   in
     'There are horses of mine there.' (LT: 128)

(55) *t'a   cuomgrá   c'robṍ*   TÉRRABA
     1SG   horse   two
     'I have two horses.' (adapted from LT: 128)

The structure of alienable possessive predications in Guaymí, Buglere, Teribe, Térraba, Bribri and Cabécar is summarized in Table 6.2.

Table 6.2 shows that the cognate alienable possessive predications found in the Viceitic subgroup have some structural similarities with other languages within the Isthmic branch. All the languages in Table 6.2 display the order possessor-possessed. Some also allow the order possessed-possessor. In particular, in Guaymí, Buglere and Teribe when the order is possessed-possessor, the latter retains extra morphological marking and no verb is used in the possessive predication. None of the languages outside the Viceitic group requires the presence of a verb to form a possessive predication. Finally, Térraba is the only language that optionally builds a possessive predication with

TABLE 6.2   Alienable possessive predications within the Isthmic group

		Alienable possessive predications					
Guaymiic	Guaymí	NP$_{PSSR}$	NP$_{PSSD}$		NP$_{PSSD}$	NP$_{PSSR}$	
		NP *gwe*	NP-∅		NP-∅	NP *gwe*	
	Buglere	NP$_{PSSR}$	NP$_{PSSD}$		NP$_{PSSD}$	NP$_{PSSR}$	
		NP-∅	NP *a*		NP-∅	NP *no*	
Viceitic	Bribri	NP$_{PSSR}$	NP$_{PSSD}$	EXIST	–		
		NP *wã*	NP-∅	*tso'*			
	Cabécar	NP$_{PSSR}$	NP$_{PSSD}$	EXIST	NP$_{PSSR}$	NP$_{PSSD}$	EXIST
		NP *wã*	NP-∅	*tsṍ*	NP *wã*	NP-∅	*tã*
Tiribí	Teribe	NP$_{PSSR}$	NP$_{PSSD}$		NP$_{PSSD}$	NP$_{PSSR}$	
		NP-∅	NP-∅		NP-∅	NP *ĩ*	
	Térraba	NP$_{PSSR}$	NP$_{PSSD}$	(EXIST)	–		
		NP-∅	NP-∅	*t'óc*			

[NP]$_{PSSR}$	[NP]$_{PSSD}$	EXIST
NP *wã	NP$_i$-Ø	*tsõ

FIGURE 6.12   Alienable possessive predication of Proto-Viceitic

an existential verb. Given the evidence available in Table 6.2, the Alienable Possessive Predication of Bribri and Cabécar in Figure 6.12 can only be plausibly reconstructed to Proto-Viceitic.

Two questions led this section: (i) whether *wã* is present in other Isthmic languages and (ii) whether these languages have an alienable possessive predication similar or identical to that of Bribri and Cabécar. The answer to latter question is obviously no, while the answer to the former is maybe.

In terms of potentially cognate elements among languages in this sample, one should take into account case markers on the possessor and the presence of existential verbs. Case markers present on the possessor are: Bribri and Cabécar *wã*, which are identical cognates, Guaymí *gwe* which could be cognate with *wã*, Buglere *no*, which does not seem to be cognate, and Teribe *ĩ*, which also does not seem to have enough phonological similarity to be cognate with either *wã* or *gwe*. As for the existentials, potential cognacy could be argued for *tso'* (Bribri) and *tső* (Cabécar) with *t'oc* (Térraba), which is phonetically [t'ók].[22] By positing the existence of a /*ts/ sequence in the proto-language, the reflexes would be /ts/ in Bribri and Cabécar and /z/ in Térraba before /*ə/ or /tʃ/ before /*i/ and /*u/ (Constenla 1981: 241). It is worth noting, however, that Constenla finds only one etymology for this sequence in Térraba and the vowels that follow the sequence do not include /*o/, which is the one needed for this particular reconstruction. For this reason, it is not possible to determine whether these forms are in fact cognates.

As for the other existential, it is worth comparing Cabécar *tã* (and Bribri *tã'*, although not used in alienable possession) with the verb *tä* in Guaymí, meaning 'be, stay', which is phonetically [tɔ]. There is another set of words which undergo exactly the same sound change: Bribri and Cabécar *kã́~kő* and Guaymí *kä* 'place'. Constenla (1981: 277, 283) actually reconstructs the form for 'time, space' for Proto-Chibchan as *ka. Thus, even if *tä* is not used to build alienable possessive predications in Guaymí, the cognate forms *tã* in Cabécar and *tã'* in Bribri could be reflexes of a proto-form *ta which has probably undergone a semantic shift from 'be, stay' to 'exist' in the Viceitic languages.

---

22   The glottalization of *t'* [tʔ] in Térraba is not predicted by regular sound change and would therefore have to be considered idiosyncratic.

Given this evidence, and considering that the comparison shows only a feeble possibility of cognacy between *wã* and *gwe* as possessor markers, one might conclude at this point that the alienable possessive predication of Proto-Viceitic was probably an innovation within this subgroup, either parallel in Bribri and Cabécar or influenced by language contact between these bordering languages. This speculation abides by the principle of parsimony, also known as Occam's razor: since a certain feature is attested only in a single clade within a given sub-branch, the uniqueness of this feature finds its "simplest" explanation in innovation.

This reasoning is, nevertheless, challenged by an inherent contradiction. The argument of innovation claims that *wã* is a 'new' piece of grammar which was not present in the proto-language. It follows from this statement that the source for this innovation should be readily available: we should find synchronic evidence for the origin of this postposition in the languages that have it, or somewhere else within the languages of the Isthmic branch. As we saw in this and the previous section, this is not the case: no synchronic source can be found for *wã*. In addition to this, the alienable possessor marker *gwe* in Guaymí could in fact be cognate with the alienable possessor marker *wã* in Bribri and Cabécar. If this were the case, independent, parallel innovation could be posited for two clades, Guaymiic and Viceitic, to avoid abandoning the argument of parsimony. Alternatively, one could posit the loss of this morpheme in all other languages within the branch, so that Guaymiic and Viceitic would become the conservative sub-branches.

Determining whether *wã* is in fact cognate with Guaymí *gwe* is pivotal to allow one to choose between innovation and conservatism. As happens in Bribri and Cabécar, *gwe* in Guaymí does not seem to have an immediate synchronic source or to be related to other postpositions present in the language (Quesada Pacheco 2008: 80–86). Similarly to the Viceitic languages, *gwe* in Guaymí, at least in the variety of Panama, is also an ergative marker.

Based on additional comparative data from Muisca, a Chibchan language from Colombia, I will argue that: (i) Guaymí *gwe* and Bribri and Cabécar *wã* are in fact cognates; (ii) they most likely come from a noun meaning 'something', 'thing', 'property' or 'belonging'; and (iii) the argument of conservatism is better grounded than the argument of innovation.

Muisca is a Chibchan language of the Magdalenic branch, a sister branch of Isthmic, which used to be spoken in Colombia and became extinct in the 18th century. Muisca had an order of possessor-possessed in attributive possession, cf. (56). Some nouns appearing in the possessor slot underwent final-vowel truncation (Ostler 1994: 208), i.e. *muysca > muysc* in (56), and *zepaba > zepab* in (57).

(56) *muysca   cubun   >   muysc-∅   cubun*   MUISCA
person   language   person-GEN   language
'Person's language (i.e. Muisca)' (STM: 208)

Possessive predications were built with a copula following the possessed (57).

(57) *ze-pab-∅   ipqua   gue*   MUISCA
my-father-GEN   something   is
'It is something of my father.' (lit: 'my father's something is.') (STM: 208)

The form *<Ipqua>*, which I will argue is the source for the possessor markers in languages of the Isthmic branch, appears in other manuscripts as *<ipcua>* (González de Pérez 1980: 96).[23] According to Constenla (1981: 146), the graphemes <c> and <qu> in Muisca represent [k]. In general, <c> was used before <a>, <o> and <u>, while <qu> before <i> and <e>. Because of the existence of *<ipcua>* in other Muisca manuscripts, it is safe to assume that the phonetic realization of this word was most likely [ipkua] or [ipkwa].[24]

A similar change is observed in the available literature on grammaticalization. Heine & Kuteva (2002: 296) offer examples from several languages in which (presumably) a noun meaning 'thing' is the source of a genitive marker (or, as the authors call it, a marker of attributive possession, something akin to the English 'of'). Languages for which this change is attested include Thai, Khmer, Japanese and Kxoe. Heine & Kuteva state that more research is needed to understand the nature and genetic/aeral distribution of this grammaticalization pattern. Relevant to the present discussion, the grammaticalization of

---

23   Two external reviewers have noticed that this Muisca form might be bi-morphemic. This observation is supported by two facts. First, if it were to be further analyzed, the most likely syllable division would be *i-pcua* (Nicholas Ostler, p.c.) given that *pcua* is an extremely common syllable in Muisca (e.g. *pcua* 'tongue, marrow, pip' and *pcuapcua* 'hat' (Adam 1878: 100)). Second, the initial vowel of *i-pcua* seems to be subject to phonological change (cf. *opcua, epcua, upcua* (González de Pérez 1980: 96). I do not currently have an etymology for the possibly bi-morphemic form *i-pcua*.

24   In other sources, such as the prescriptive grammar of Lugo (1619) and López García Molins (1995), *ipqua* is said to be a 'genitive case marker'. However, Nicholas Ostler (p.c.) informs me that this is a misunderstanding, possibly due to the wish of the Catholic friar Bernardo de Lugo to find in Muisca the same categories of Latin, such as the genitive case. The word *ipqua* is the Muisca equivalent of the interrogative pronoun 'what'. When it has this function, it must be followed by an interrogative particle (*ua* or *o* as in *ipqua ua* or *ipquo*). This interrogative pronoun can also be used after a noun in the genitive (i.e. a noun which sometimes undergoes final-vowel truncation), and in that case it means 'something' (i.e. X's something > X's property).

'something' > possessor marker (> ergative marker) in Chibchan relies on similar attested cases.[25]

Appendix A offers a historical reconstruction of a possible proto-form from which Muisca *ipcua*, Guaymí *gwe*, Bribri and Cabécar *wã* could have originated. The reconstruction is based on the comparative Chibchan phonology of Constenla (1981), as updated in later publications (Constenla 1989, 2008). As readers will notice, the reconstruction based on the comparative method in Appendix A is far from perfect. The actual synchronic form in a given language cannot always be predicted by regular sound change and in many instances, idiosyncratic sound changes need to be posited. However, there is additional evidence in favor of an ancient possessor subject marker in other languages of the family.

Several other languages within the Chibchan family seem to have retained fragments of the proto-form *i-pkwə. In Boruca (Western Isthmic), when the possessor within a possessive NP is a proper name, a personification or an animal, it can be marked by *i* or *ígui* (Castro 2008: 67), as in (58).

(58) *Juan   ígui   ú*   BORUCA
     J.   PSSR   house
     'Juan's house' (BT: 68)

The form <*ígui*> (phonetically [ígi]) is a variation of <*égui*> which results from the blend of two segments: <*éc*>, a postposition which indicates possession, plus the pronominal segment <*i*> (Miguel Angel Quesada Pacheco, p.c.). Although this remains simply a speculation, it could be that Boruca *éc~íc* is an eroded reflex of *ipkwə. Since *i > i in Boruca (Constenla 1981: 199) and *p > ∅ before *k (Constenla 1981: 220), regular sound change would give /ik/ (<ic>) as a result. The loss of the rest of the proto-form would then be idiosyncratic.[26]

Other Chibchan languages form their 'have' verb by combining the verb 'be' with a prefix that has the form of *kw-* or *k-* plus some sort of vowel. This is the case of Rama, a Chibchan language from Nicaragua. In Rama, the verb 'have' is formed by the verb *aakar* 'be' plus the prefix *kw-* (Stassen 2009: 632). This

---

25   Doris L. Payne has suggested to me that *ipcua* 'something' might be a possessive classifier rather than a generic noun meaning 'something'. Unfortunately, this question remains open, given the limited data available on Muisca.

26   Given the scenario outlined so far in terms of Proto-Chibchan cognates of *i-pkwə, the *ĩ* marker, found in Teribe in the construction in which the possessor follows the possessed and is marked by *ĩ* (see Table 6.2), might also be a heavily eroded reflex of *i-pkwə. I do not currently have, however, an explanation for the nasality of the vowel /i/, which is not predicted by regular sound change. Therefore, this remains only a speculation.

prefix is not used elsewhere as a postposition or a pre-verb in the language and synchronically the verb 'to have' should be considered monomorphemic (Craig p.c. in Stassen 2009). Rama has relational preverbs synchronically and diachronically derived from postpositions (Craig & Hale 1988: 313) which become prefixed to verbs by some sort of incorporation process, yielding a construction similar to an applicative. As such, cognates to etymological (and synchronic) postpositions can sometimes be found as prefixes on verbs. An example of alienable possessive predication with the verb 'have' in Rama is offered in (59).

(59) *ngainguk*   *hap*   *i-kwaakar*   RAMA
      money   some   3SG-have
      'He has some money.' (Craig 1990: 61)

Damana, a Chibchan language from Colombia, displays a similar pattern (Stassen 2009: 632). The verb 'have' is *kʉnʉn,* formed by *nʉn* 'be' plus the dative or benefactive *kʉ*, phonetically [kə] (Trillos Amaya *et al.* 1989: 47). This tendency is also found in Paya, the northernmost Chibchan language from Honduras: the existential *šu* combines with the benefactive prefix *kà-* to obtain the verb 'have' (Holt 1999: 77). In the light of what has been noted about relational preverbs, the hypothesis is that the proto-form *i-pkwə yields *kw-* in Rama, *kʉ-* in Damana, and *kà-* in Paya.[27] These erstwhile preverbal postpositions could then have become prefixed to the verb. In all cases, there would be extreme phonological erosion, presumably idiosyncratic, consistent with grammaticalization or lexicalization.

For the sake of completeness, I should mention that an anonymous reviewer proposed that *gue* 'be' in (57) may be the source of the Possessor markers *wã* in Bribri and Cabécar and *gwe* in Guaymí. According to this hypothesis, we would need to posit the existence, in the proto-language, of a construction such as *[Possessed-∅ Possessor-∅]=*gue which was then reanalyzed as *[Possessed-∅] [Possessor-*gue]. In this reanalysis, a clause final existential particle is reanalyzed as a possessive particle syntactically attached to the Possessor NP. After the reanalysis, a change in the order of the two NPs would be posited, where the Possessor NP marked by *gue would come before

---

27   Reflexes of the consonant portion of the Rama, Damana and Paya prefixes appear to be regular. In Paya, *k >∅ word-initially and *k >k elsewhere (Constenla 1981: 250). In Rama, *k> k (Constenla 1981: 257). In Damana, *k > g intervocalically and *k>k elsewhere (Constenla 1981: 313). Reflexes of the vowel portion of these prefixes do not appear to be regular in the case of Paya and Damana.

the Possessed NP, possibly for topicality reasons. In support of this analysis, Heine & Kuteva (2002: 127) present several languages in which an existential verb has grammaticalized into a possession marker in possessive predications. According to Heine & Kuteva this type of reanalysis requires the possessee NP to be the subject and the possessor NP to be a genitival modifier of the subject. I propose one main counterargument to this hypothesis. The exact phonetic value of *<gue>* in Muisca is uncertain, due to Spanish orthographic issues at the time in which Spanish Catholic friars transcribed Muisca (see Constenla 1981: 147). There are at least six possible phonetic values for Muisca *<gue>*: [gwe], [we], [ge], [gwɨ], [wɨ] and [gɨ] (Nicholas Ostler, p.c.). This uncertainty creates problems in the reconstruction of a possible proto-form (see Appendix B). In general, the reconstruction presented in Appendix B seems to pose slightly more complications than the reconstruction for my hypothesis (cf. Appendix A). By no means is this second hypothesis untenable, but as presented above, the first hypothesis seems to have more supporting evidence.

Based on the evidence advanced so far, rather than a shallow innovation restricted to the Viceitic group within the Isthmic Branch, it is now necessary to reconstruct a proto-form, and then to posit retention in Guaymí, Bribri and Cabécar, along with possible relics in other languages (Boruca, Rama, Damana, Paya). The complete loss of this proto-form in the other languages would not be entirely unexpected: there would have been a time lapse of centuries for this morpheme to be lost across the other languages (something that cannot be stated for the 'innovation' hypothesis).

However, the conservatism hypothesis sets forth a question. This question has to do with the marking of attributive possession versus possessive predication and with how a possessor marker gained subject properties in some of the languages involved in the comparison. In particular, in the Viceitic subgroup, alienable and inalienable attributive possession do not display any kind of marking on the (NP-internal) possessor, cf. (11) and (12). In possessive predications however, we find that the possessor subject is marked by *wã*, which presumably comes from a proto-form meaning 'something' in Proto-Chibchan. Most importantly, the possessor NP marked by *wã* in possessive predications displays subject properties. For instance, in attributive possession in Bribri, the possessor and the possessed form an indivisible constituent (i.e. in a possessive NP), independently of whether the possession is alienable or inalienable. Thus, compare (60) and (61).

(60) *Alí chíchi sḗ-r-ke tër kêkra-ë* BRIBRI
　　 A. dog live-MVC-IPFVII exist.POS.SG always-INT
　　 'Alí's dog is always lying down.' (EL)

(61) *Alí   kêkra-ë   chíchi   sé-r-ke   tër   BRIBRI
      A.   always-INT   dog   live-MVC-IPFVII   exist.POS.SG
      *'Ali's dog is always lying down.' (EL)

However, the possessor subject marked by *wã* and the possessed object in possessive predications do not form a constituent. The same adverb that cannot go between possessor and possessed in (61) can separate them in (62).

(62) Alí   wã   kêkra-ë   nũköl   tso'   BRIBRI
      A.   PSSR   always-INT   money   exist
      'Alí always has money.' (EL)

In Guaymí, a possessive predication looks identical to attributive possession because there is no overtly expressed verb. Unlike the Viceitic languages, the possessor is marked by *gwe* in both attributive possession and possessive predications (Quesada Pacheco 2008: 72). The only data I have been able to find which tells something about constituency in a possessive predication is from Young & Givón (1990: 211), when the sentential negation occurs between the possessed and the marked possessor.

(63) krägä   nyaka   nun-gwe   GUAYMÍ
      medicine   NEG   1PL-GEN
      'We have no medicine.' (lit: 'Medicine is not ours.')

One example is far from sufficient to draw conclusions about constituency in Guaymí: synchronic data is needed to determine whether attributive possession and possessive predication are really identical in this language in all aspects, including, for example, intonation, and whether the Possessor and the Possessed form a constituent in either or both attributive possession and/or possessive predication.

   Whatever the case might be, in Bribri and Cabécar, attributive possession lacks marking on the possessor, while possessive predications feature a special marker for the possessor subject. In this respect, Stassen (2009: 27) states that the ways of marking possession on NPs and in clauses do not seem to abide by any predictable match: there are languages in which attributive possession derives from possessive predication (clauses) and languages in which it does not. In the latter case, attributive possession and possessive predication display divergent morphosyntactic patterns:

   It is theoretically possible that, in some languages, a distinction is made between the clausal and the phrasal syntax of possession. In particular,

> it is possible that the phrasal syntax of possession is more grammatical-
> ized than the clausal syntax of possession is. In such languages, the pos-
> sessor NP and the possessee NP may form a constituent in attributive
> possession, while they do not form a constituent in predicative posses-
> sion. (Stassen 2009: 114)

This theoretical possibility is the case in Bribri: the possessor and the possessed
form a constituent in attributive possession but *not* in possessive predication.

Within the Viceitic group, the question is then how the inherited proto-
form ended up only in possessive predications and when and how the inherit-
ed proto-form, reanalyzed as a possessor, developed subject properties outside
of a possessive NP.

This section has reached the conclusion that the conservatism hypothesis
wins over the innovation hypothesis: the languages which synchronically have
a possessor marker in alienable possessive predications have inherited and re-
analyzed it possibly from a proto-form which originally meant 'something', or
'thing' or 'property'.

Future research should address whether other Isthmic languages might
have preserved relics of the proto-marker *i-pkwə in inalienable possessive
predications. As a preview, in Bribri, inalienable possession can be expressed
by constructions which feature the copula *dör* or the existential *tã'*. The in-
alienable possessor is usually not marked by the postposition *wã*. However, in
the domain of objects of personal use, the possessor can be optionally marked
by *wã*, at least in the case of certain objects, such as 'house' in (64).

(64) *ie'*   *(wã)*   *ú*   *tã'*   BRIBRI
     1SG   PSSR   house   exist
     'I have a house.' (i.e. I do not rent this house, rather I live in it)[28] (EL)

In Cabécar, on the other hand, the inalienable possessor must be marked by
the postposition *wã*. Unlike Bribri, the existential *tsố* can be used, cf. (65),
along with *tã*, cf. (66), to express inalienable possession.

(65) *yís*   *wã*   *kuta*   *tsố*   *tkḗl*   CABÉCAR
     1SG   PSSR   sister   exist   four
     'I have four sisters.' (DCE: 310)

---

28   A similar example, in which *wã* is not indicated as optional, is found in Dickeman-Datz
    (1984: 119).

(66) *mõgú   wã   wöbla   shabóo   tã   CABÉCAR
     owl   PSSR   eyes   almost.full exist
     'Owls have really big eyes.' (DCE: 271)

As far as the Viceitic languages are concerned, Cabécar appears to be the more conservative in that (i) it has preserved the possessor marker *wã* in both types of possessive predication; and (ii) there is no distinction in the verbs used to express alienable and inalienable possession.

5   Conclusions

This article has demonstrated how the syncretism between the ergative marker and the possessor subject marker in the Viceitic languages of the Chibchan family came to be. Further, the article has posited a possible origin for this morpheme, *wã*, in Proto-Chibchan *i-pkwə 'something'. Although several questions still remain unanswered, digging into diachrony has proved pivotal to understanding why Bribri and Cabécar display the relatively uncommon phenomenon of differential ergative marking.

In §3 (and subsections therein), I demonstrated that the possessor marker *wã* was reanalyzed as an ergative marker in Bribri and Cabécar: this was possible because the perfect construction in which *wã* is found historically came from a possessive predication, a well-established evolutionary pattern. The discovery of this path of reanalysis brought about inquiries with respect to the ultimate origin of this marker, as well as a methodological dilemma concerning the choice between innovation and conservatism. Comparative work has shown that this marker is not found in more closely related languages within the larger branch: a reasonable assumption was then to claim that it represents an innovation within a sub-group or two of the larger branch. This assumption, in turn, made room for a potential contradiction: if the innovation hypothesis were right, then there should be a readily available synchronic source for *wã*, whether in the languages that have it or in more closely related ones. The absence of such a synchronic source proved fundamental to giving more credit to the hypothesis that the morpheme is a conservative relic.

This possibility led to further inquiries outside of the immediate larger branch of which the languages under survey were part, into more distant branches (cf. §4). The identification of a cognate form in Muisca, a language as far away as Colombia, invited the reconstruction of a proto-form meaning 'thing', 'something' or 'property' in Proto-Chibchan which would have been

reanalyzed in several Chibchan languages as a possessor marker, and in three of these, further as an ergative marker. This reanalysis is supported by attested changes in the same direction ('thing' > possessor marker). The reconstruction of this form was supported by the presence of cognate forms scattered throughout the Chibchan family which abide (mostly) by regular sound change rules in their reflexes of the proto-form (see Appendix A). Additional evidence was found in other Chibchan languages which could have incorporated a heavily eroded relic of the proto-form at the beginning of certain verbs (i.e. the grammaticalization of preverbal postpositions into verbal prefixes).

In addition, because the ergative marker *wã* came from a possessor (i.e. from an 'oblique' case), the presence of 'Type B' ergativity has been determined to exist within the Chibchan family, at least in Bribri, Cabécar and Guaymí. Given that this type of ergativity arises from the reanalysis of an oblique case marker (such as a possessor), one future inquiry should be concerned with whether the standard Proto-Chibchan ergative marker *tV is an instance of 'Type A' ergativity, that is, of a passive made obligatory, or of something else.

Although we now know the source of the ergative marker *wã* in perfect constructions, this marker is also found in several other constructions in both languages, such as the transitive perfective negative and the caused motion construction, among others. Future studies will need to probe the spread of this innovative ergative marker into these other constructions.[29]

Appendix A: Reconstruction of the Proto-Form *i-pkwə and Its Reflexes in Different Chibchan Languages

The hypothesis illustrated in this appendix is that the forms presented in Table 6.3 represent a cognate set of reflexes of the Proto-Chibchan form *i-pkwə '(some)thing'.

Two observations are in order before presenting the regular (and irregular) sound changes which operated to give the reflexes of the proto-form *i-pkwə in each of the languages in Table 6.3. First, when comparing and reconstructing Chibchan languages,

---

29   Many unsolved questions remain with respect to ergative marking in the Viceitic languages of the Chibchan family. For instance, it is unclear at the present time why the innovative ergative marker *wã* appears in negative domains in Cabécar and Bribri, although in the latter to a lesser extent. An anonymous reviewer interestingly suggested that the presence of wã in negative contexts indicates some sort of pragmatic force associated with this ergative marker, possibly connected to the existential sense presumably present in the original construction.

TABLE 6.3   Potential cognates for a proto-form *i-pkwə in some Chibchan languages

	Proto-form 'thing'/'something'
	*i-pkwə
MUISCA	[i-pkwa] (<ipkua, ipcua>)
CABÉCAR	[wã] (<wã>)
BRIBRI	[wã] (<wã>)
GUAYMÍ	[gwe/kwe] (<gwe/kwe[a]>)

a   *Gwe* is found in the variety of Guaymí spoken in Costa Rica. *Kwe* occurs in the variety spoken
     in Panama, of which the Costa Rican variety is considered a dialect (Murillo 2010).

it is often the case that a given root appears with different vowel endings (Constenla 1989, 2008) in different languages. These vowel endings are akin to 'thematic vowels' and have no function other than combining with a given root.[30] Constenla (2008: 130) offers several examples of this phenomenon and acknowledges that it can be a serious problem when reconstructing within Chibchan. For instance, for the meaning 'salt, sea', there are reflexes of the proto-root *dahg plus: (i) the vowel ending /*-e/ in three Chibchan languages from Central America (Paya /táːké/, Bribri /daʤî/, Cabécar /daʤî/); (ii) the vowel ending /*-u/ in three Chibchan languages from Colombia (Cogui /nəkku/, Damana /ningu/, Ika /nəggɨ/); and (iii) a sequence of the vowel endings /*-u/ and /*-a/ in two other Chibchan languages from Colombia (Muisca /nɨgua/, Tunebo /ɾauwa/). Different vowel endings are also common language-internally (for example, in Cuna /nue/, /nui/ and /nua/ all mean 'good'). In the reconstructions I propose here final vowels often do not form a regular correspondence set, but this is just one specific instance of a more general phenomenon in the family.

Second, for this historical reconstruction I rely entirely on the comparative Chibchan phonology of Constenla (1981). The vowel system reconstructed in 1981 comprised 8 vowels (/*i/, /*ɪ/, /*e/, /*a/, /*ə/, /*o/, /*ʊ/, and /*u/). This system was subsequently amended (Constenla 1989, 2008) and reduced to 5 vowels (/*i/, /*e/, /*a/, /*o/, /*u/). The proto-phonemes /*ɪ/, /*ʊ/ and /*ə/ proposed in 1981 were later ascribed to /*i/, /*u/ and /*a/ respectively. For the purposes of this reconstruction, then, the proto-form *i-pkwə would be ascribed to *i-pkwa.

---

30   This is my translation of 'formativos vocálicos' in Constenla (1989, 2008). In fact these
      endings include more than just vowels. Constenla (2008: 130) enumerates at least the fol-
      lowing: /*-a/, /*-e/, /*-i/, /*-o/, /*-u/, /*-ʔ/, /*~/, /*-ke/, /*-te/, /*-ka/, /*-ba/.

In each of the following tables, the first column shows the proposed proto-form, the second column illustrates the reflexes of each phoneme of the proto-form *i-pkwə in a given daughter language and the regular sound changes, as described in the comparative Chibchan phonology of Constenla (1981). Page numbers in parentheses refer exclusively to this source. The third column indicates the kind of sound change which took place. Sound changes which are language specific and cannot be reconstructed by claiming regular sound change are indicated in parentheses as idiosyncratic. The fourth column summarizes all the stages of the evolution from the proto-form to the modern reflex. A remark which applies to all tables is that Proto-Chibchan probably had asyllabic allophones of /i/ and /u/ at the beginning of syllables when these proto-vowels were followed by another vowel (Constenla 1981: 208). This is why, in all tables in Appendix A (and B), /u/ followed by another vowel has been transcribed as [w].

As can be seen in Table 6.4, the reflex of /*i-pkwə/ in Muisca is conservative with respect to other languages and abides by regular sound change. As can be seen in Table 6.5, in Cabécar, if we assume that the proto-form was bi-morphemic and that the initial *i- was lost, we are left with the form *kwã > *wã. This kind of simplification (/kw/>/w/) is attested as a tendency in the languages of South and Central America (Holmer 1947: 56). For Bribri, the same reflexes presented for Cabécar apply, with the exception that Constenla (1981: 194) indicates that /*ə/ has /à/ as a reflex (with a low tone) in Bribri. Conceivably, because of centuries of intimate contact with Cabécar, an intermediate form /wà/ could have become nasalized; alternatively, the nasalization originated independently as a compensatory strategy for the erosion of /i-kwà/. As for the phonological erosion that has been posited as an idiosyncratic sound change in Bribri and Cabécar, this sort of reduction is widely attested in cases of

TABLE 6.4   Reflexes of the proto-form *i-pkwə in Muisca

Muisca			
Protoform /*i-pkwə/	Regular sound changes yielding reflexes	Type of sound change	Evolution
	/*i/ > /i/ in non-final position (p.199)	–	/i-pkwə/
	/*p/> /p/ before /*k/ (p.220)	–	/i-pkwə/
	/*k/> /k/ (p.228)	–	/i-pkwə/
	/*u/>/u/ in non-final position (p.204)	–	/i-pkwə/
	/*ə/>/a/ after /*u/ (p.195)	vowel lowering	/i-pkwa/

TABLE 6.5   Reflexes of the proto-form *i-pkwə in Proto-Viceitic

Proto-Viceitic

Protoform /*i-pkwə/	Regular sound changes yielding reflexes	Type of sound change	Evolution
	/*i/ > /i/ (p.199)	–	/i-pkwə/
	/*p/ > /∅/ before /*k/ (p.219)	cluster reduction	/i-kwə/
	/*k/ > /k/ (p.227)	–	/i-kwə/
	/*u/ > /u/ (p.203)	–	/i-kwə/
Cabécar	/*ə/ > /à/ or /ã/ (p.194)	vowel lowering, low tone or nasalization	/i-kwã/
	–	phonological erosion (idiosyncratic)	/wã/
Bribri	/*ə/ > /à/ (p.194)	vowel lowering, low tone	/i-kwà/
	–	phonological erosion (idiosyncratic)	/wà/
	–	nasalization due to contact or loss of segment (idiosyncratic)	/wã/

grammaticalization (Lehmann 1985; Traugott & Heine 1991; Heine et al. 1991; Hopper & Traugott 1993; Bybee et al. 1994; Heine & Kuteva 2007; *inter alia*) – *wã* is a highly grammatical morpheme and, as such, because of frequency of use, it is reasonable to assume that erosion happens at a higher rate compared to lexical morphemes.

In the case of Guaymí (see Table 6.6), the idiosyncratic change of vowel raising is supported by the description of Quesada Pacheco (2008: 26), who shows that vowel harmony in the form of raising is common in this language. Here, we would have to posit that vowel harmony occurred before the loss of the initial *i- or, alternatively, we would have to reconstruct a different vowel ending for Guaymí based on the attested tendency of Chibchan languages to show different vowel endings for the same root (Constenla 1989, 2008). Therefore, Guaymí *gwe/kwe* could be the result of the bimorphemic root *i-pkw- plus the vowel ending *-e which gave /e/ as a reflex in Guaymí (Constenla 1981: 278).

TABLE 6.6   Reflexes of the proto-form *i-pkwə in Guaymí

	Guaymí		
Protoform ***/i-pkwə/**	**Regular sound changes** **yielding reflexes**	**Type of sound change**	**Evolution**
	/*i/ > /i/[a] (p.279)	–	/i-pkwə/
	/*p/> /∅/ before /*k/ (p.284)	cluster reduction	/i-kwə/
	/*k/> /k/ (p.284)	–	/i-kwə/
	/*u/> /u/ (p.280)	–	/i-kwə/
	/*ə/>/a/ after /*u/ (p.277)[b]	–	/i-kwa/
	–	vowel harmony and raising (idiosyncratic)	/i-kwe/
	–	loss of initial vowel (idiosyncratic)	/kwe/~/gwe/

a The reflex /i/ in Guaymí occurs in the general 'elsewhere' environment. /*i/ gave /e/ as a reflex after a bilabial and /ɪ/ before /*ʔ/ or /*k/ if not preceded by a voiced consonant (Constenla 1981: 279).

b Constenla (1981: 278) states that in one postposition /*ə/ gave /e/ and in other two postpositions either /*ə/ or /*a/ gave /e/. He hypothesizes that the phonological environment could be in word final position if unstressed. However, in these three cases, /e/ is never preceded by /*u/.

Appendix B: Reconstruction of the Proto-Form *kwə~*kwe 'exist' and Its Reflexes in Different Chibchan Languages

The hypothesis illustrated in this appendix is that the forms presented in Table 6.7 represent a cognate set of reflexes of the Proto-Chibchan form *kwə~*kwe 'exist'.

The great variation in vowel endings for certain roots across Chibchan languages (see Appendix A) represents the same problem for this reconstruction (see below). Two additional observations are relevant for the reconstruction in Appendix B. First, I have added to the cognate set, the Cuna form *<kue>,* an independent verb with the meaning of 'be, take place' (Holmer 1947: 156). Although in Cuna the hypothesized reanalysis of *kwə~*kwe into a possessive particle has not taken place, the form is similar enough to the others in meaning and shape for it to be considered as a possible cognate. Second, the phonetic realizations of *<gue>* in Muisca vary greatly, and it is impossible to determine exactly which one of the six possibilities is the correct reflex for

TABLE 6.7   Potential cognates for a proto-form *kwə~*kwe in some Chibchan languages

	Proto-Form 'exist, be' *kwə~*kwe
Muisca	[gwe], [we], [ge], [gwɨ], [wɨ], [gɨ] (<gue, guɣ[a]>)
Cabécar	[wã] (<wã>)
Bribri	[wã] (<wã>)
Guaymí	[gwe/kwe] (<gwe/kwe>)
Cuna	[kwe][b] (<kue>)

a   A symbol similar to <ɣ> is found in the Lugo grammar of Muisca (1619) and it is considered to represent a vowel intermediate between [e] and [i], probably [ɨ].
b   Often realized phonetically as [ɣwe] or [we].

constructing the correspondence (see discussion in §4). For reasons of convenience, I have based the reconstructions in this Appendix on the assumption that Muisca *<gue>* is phonetically [gwe]; this choice has no consequences for assessing the plausibility of the reconstruction.

As in Appendix A, in each of the following tables, the first column shows the proposed proto-form, the second column illustrates the reflexes of each phoneme of the proto-form *kwə~*kwe in a given daughter language and the sound changes which took place according to regular sound change, as described in the comparative Chibchan phonology of Constenla (1981). Page numbers in parentheses refer exclusively to this source. The third column indicates the kind of sound change which took place. Sound changes which are language specific and cannot be reconstructed by claiming regular sound change are indicated in parentheses as idiosyncratic. The fourth column summarizes all the stages of the evolution from the proto-form to the modern reflex. As a final remark, although most forms in Table 6.7 present an initial /g/ instead of /k/, the reconstruction of a *g posits serious problems in terms of reflexes.[31] Therefore, *k has been preferred.

In the scenario outlined by the reconstruction in Appendix B, Muisca (see Table 6.8) is the most problematic case. The main problem is that for all the proto-vowels that could reasonably be posited for the proto-form (i.e. *a, *e, *ə) Muisca has /a/ as a reflex (Constenla 1981: 192 ff.). Positing a Muisca form without /u/, such as [ge] or [gɨ],

---

31   Based on Constenla (1981), reconstructing a *g for the forms in Table 6.7 would give the following non-expected reflexes: /h/ in Cabécar, /ŋ/ in Guaymí and /s/ in Cuna.

TABLE 6.8   Reflexes of the proto-form *kwə ~ *kwe in Muisca

Muisca			
Protoform /*kwə/	**Regular sound changes yielding reflexes**	**Type of sound change**	**Evolution**
	/*k/> /k/ (p.228)	–	/kwə/
	/*u/>/u/ in non-final position (p.204)	–	/kwə/
	/*ə/>/a/ after /*u/ (p.195)	–	/kwa/
	–	vowel raising (idiosyncratic)	/kwe/
	–	voicing (idiosyncratic)	/gwe/

TABLE 6.9   Reflexes of the proto-form *kwə ~ *kwe in Proto-Viceitic

Cabécar			
Protoform /*kwə/	**Regular sound changes yielding reflexes**	**Type of sound change**	**Evolution**
	/*k/> /k/ (p.227)	–	/kwə/
	/*u/> /u/ (p.203)	–	/kwə/
Cabécar	/*ə/>/à/ or /ã/ (p.194)	vowel lowering, low tone or nasalization	/kwã/
	–	phonological erosion (idiosyncratic)	/wã/
Bribri	/*ə/>/à/ (p.194)	vowel lowering	/kwà/
	–	phonological erosion (idiosyncratic)	/wà/
	–	nasalization due to contact or loss of segment (idiosyncratic)	/wã/

is not of much help, because the loss of asyllabic /u/ would also need to be posited as an idiosyncratic sound change.

The same observations made in Appendix A for Bribri also apply to this reconstruction. The intermediate form /wà/ in Table 6.9 could have become nasalized due to contact with Cabécar, or the nasalization could have originated independently, as a compensatory strategy for the loss of /k/ (erosion).

In the case of Guaymí and Cuna (see Tables 6.10 and 6.11 respectively), an alternative solution to the idiosyncratic vowel raising would be to posit the presence of two proto-forms with an alternation in the final vowel portion: *kw-ə (> *kwa in the revised comparative phonology of Constenla 1989 and 2008 in which *ə is ascribed to *a) for Muisca, Cabécar and Bribri, and *kw-e for Guaymí and Cuna. In both of these

TABLE 6.10   Reflexes of the proto-form *kwə ~ *kwe in Guaymí

Guaymí			
Protoform ***/kwə/** **or */kwe/**	**Regular sound changes** **yielding reflexes**	**Type of sound change**	**Evolution**
	/*k/> /k/ (p.284)	–	/kwə/
	/*u/> /u/ (p.280)	–	/kwə/
	/*ə/>/a/ after /*u/ (p.277)[a]	–	/kwa/
	–	vowel raising (idiosyncratic)	/kwe/

a   See fn. 33 in Appendix A.

TABLE 6.11   Reflexes of the proto-form *kwə ~ *kwe in Cuna

Cuna			
Protoform ***/kwə/** **or */kwe/**	**Regular sound changes** **yielding reflexes**	**Type of sound change**	**Evolution**
	/*k/> /k/ (p.264)	–	/kwə/
	/*u/> /u/ (p.262)	–	/kwə/
	/*ə/>/a/ (p.261)	–	/kwa/
	–	vowel raising (idiosyncratic)	/kwe/

languages *e > e (Constenla 1981: 261, 278). This would then be a very common case of different vowel endings for a given root in different Chibchan languages (see discussion in Appendix A).

The reconstruction in Appendix B has an advantage compared to that in Appendix A. Positing that the origin of *wã* in Bribri and Cabécar is to be found in a verbal particle could potentially explain why these languages have a verbal suffix *-wã*, identical in form to the possessor marker *wã*. However, in terms of function, the meanings of the synchronic suffix *-wã* include, depending on the author, "completion of an action", "movement of penetration" and "punctuality" (cf. fn. 8). These meanings do not seem, at first glance, to be derivable from a former existential particle.

Additionally, if Muisca *ipqua* and Boruca *ic-* are in fact cognates (as my reconstruction in Appendix A suggests), then the alternative reconstruction in Appendix B does not explain the presence of a prefix *i-* not attested elsewhere and present, presumably, in at least two branches (Magdalenic and Isthmic).

References

Adam, Lucien. 1878. *Etudes sur six langues américaines: dakota, chibcha, nahuatl, kechua, quiche, maya*. Paris: Maisonneuve.

Arkadiev, Peter. 2017. Multiple Ergatives: From Allomorphy to Differential Agent Marking. *Studies in Language* 41(3): 717–780.

Benveniste, Émile. 1966. *Problèmes de linguistique générale*. Paris: Gallimard.

Bybee, Joan L., Revere D. Perkins & William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. Chicago: University of Chicago Press.

Castro, Damaris. 2008. Brunkahk Tek: An Extinct Language. *Letras* 43: 51–74.

Chevrier, Natacha. 2017. Analyse de la phonologie du bribri (chibcha) dans une perspective typologique: Nasalité et géminée modulée. Doctoral dissertation. Université Lumière, Lyon II.

Comrie, Bernard. 1978. Ergativity. *Syntactic Typology: Studies in the Phenomenology of Language*, ed. by Winfred P. Lehmann. Austin: University of Texas Press.

Constenla Umaña, Adolfo. 1981. Comparative Chibchan Phonology. Ph.D. dissertation, University of Pennsylvania.

Constenla Umaña, Adolfo. 1985. Clasificación lexicoestadística de las lenguas de la familia chibcha. *Estudios de Lingüística Chibcha* 4: 155–197.

Constenla Umaña, Adolfo. 1989. Subagrupación de las lenguas chibchas: Algunos nuevos indicios comparativos y lexicoestadísticos. *Estudios de Lingüística Chibcha* 8: 17–72.

Constenla Umaña, Adolfo. 1991. *Las lenguas del Área Intermedia: Introducción a su estudio areal*. San José: Editorial de la Universidad de Costa Rica.

Constenla Umaña, Adolfo. 2007. *La lengua de Térraba*: *Esbozo fonológico y gramatical y pequeño diccionario*. San José: Editorial de la Universidad de Costa Rica.

Constenla Umaña, Adolfo. 2008. Estado actual de la suclasificación de las lenguas chibchenses y de la reconstrucción fonológica y gramatical del protochibchense. *Estudios de Lingüística Chibcha* 27: 117–135.

Constenla Umaña, Adolfo & Enrique Margery Peña. 1979. *Bribri II*. San José: Editorial de la Universidad de Costa Rica.

Constenla Umaña, Adolfo & Eugenia Ibarra Rojas. 2009. Mapa de la distribución territorial aproximada de las lenguas indígenas habladas en Costa Rica y en sectores colindantes de Nicaragua y de Panamá en el siglo XVI. *Estudios de Lingüística Chibcha* 28: 109–112.

Constenla Umaña, Adolfo, Feliciano Elizondo Figueroa & Francisco Pereira Mora. 1998. *Curso básico de Bribri*. San José: Editorial de la Universidad de Costa Rica.

Craig, Colette & Ken Hale. 1988. Relational Preverbs in Some Languages of the Americas: Typological and Historical Perspectives. *Language* 64(2): 312–344.

Craig, Colette. 1990. *Rama Kuup: Gramática Rama*. Centro de Investigaciones y Documentación de la Costa Atlántica. Printed in Eugene, Oregon.

Creissels, Denis. 2013. Control and the Evolution of Possessive and Existential Constructions. *Argument Structure in Flux: The Naples-Capri Papers*, ed. by Elly van Gelderen, Jóhanna Barðdal & Michela Cennamo, 461–476. Amsterdam/Philadelphia: John Benjamins.

Creissels, Denis. 2014. Predicative Possession and Existential Predication. Paper presented at the Ecóle thématique LACITO, Méthodes en linguistique: des données empiriques aux hypothèses typologiques held in Roscoff, France, July 2014.

Dickeman Datz, Margaret. 1984. Split Ergativity and Subject in Bribri. *Estudios de Lingüística Chibcha* 3: 113–134.

González de Pérez, María Stella. 1980. *Trayectoria de los estudios sobre la lengua chibcha o muisca*. Bogotá: Instituto Caro y Cuervo.

Haig, Geoffrey. 2008. *Alignment Change in Iranian Languages: A Construction Grammar Approach*. New York: Mouton de Gruyter.

Heine, Bernd. 1997. *Possession: Cognitive Sources, Forces, and Grammaticalization*. Cambridge: Cambridge University Press.

Heine, Bernd, Ulrike Claudi & Friederike Hünnemeyer. 1991. *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press.

Heine, Bernd & Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

Heine, Bernd & Tania Kuteva. 2006. *The Changing Languages of Europe*. Oxford: Oxford University Press.

Heine, Bernd & Tania Kuteva. 2007. *The Genesis of Grammar*. Oxford: Oxford University Press.

Hoff, Berend J. 1995. Configurationality and Nonconfigurationality in the Carib Language of Surinam. *International Journal of American Linguistics* 61(4): 347–377.

Holmer, Nils M. 1947. *Critical and Comparative Grammar of the Cuna Language.* Göteborg: Etnografiska Museet.

Holt, Dennis Graham. 1986. The Development of the Paya Sound System. Doctoral dissertation, University of California at Los Angeles.

Holt, Dennis Graham. 1999. *Pech (Paya).* München: LINCOM Europa.

Hopper, Paul J. & Elizabeth Closs Traugott. 1993. *Grammaticalization.* Cambridge: Cambridge University Press.

Jara Murillo, Carla Victoria. 1993. *I ttè: Historias Bribris*. San José: Editorial de la Universidad de Costa Rica.

Jara Murillo, Carla Victoria & Alí García Segura. 2008. *Cargos tradicionales del pueblo bribri: Sĩõ'tãmĩ / Óköm / Awá*. San José: Instituto Costarricense de Electricidad.

Jara Murillo, Carla Victoria & Alí García Segura. 2009. *Se' ẽ' yawö bribri wa: Aprendemos la lengua Bribri*. San José: UNICEF/Universidad de Costa Rica.

Jara Murillo, Carla Victoria & Alí García Segura. 2013. *Se' ttŏ bribri ie*: *Hablemos en bribri*. San José: E-digital *ED.*

Lehmann, Christian. 1985. *Thoughts on Grammaticalization*. München: Lincom Europa.

López García-Molíns, Angel. 1995. *Gramática muisca*. München: Lincom Europa.

Lugo, Bernardo de. 1619. *Gramática en la lengva general del Nvevo Reyno, llamada mosca*. Bogotá: Ediciones Sol y Luna.

Margery Peña, Enrique. 1989. *Diccionario cabécar–español, español–cabécar.* San José: Editorial de la Universidad de Costa Rica.

McGregor, William B. 2009. A Typology of Ergativity. *Language and Linguistics Compass* 3(1): 480–508.

Murillo Miranda, José Manuel. 2010. La oración simple ngäbére. *Forma y función* 23(1): 39–69.

Ostler, Nicholas. 1994. Syntactic Typology of Muisca: A Sketch. *Languages in the Andes*, ed. by Peter Cole, Gabriella Hermon, Mario Daniel Martin. University of Delaware: Latin American Studies.

Pache, Matthias. 2016. External Relations of Chibchan. Paper presented at the Intermediate Area Workshop I, the Chibchan and Arawakan Families: Linguistic and Cultural Traits, Lyon DDL, Université Lumière Lyon 2, April 2016.

Pache, Matthias. 2018. Contributions to Chibchan Historical Linguistics. Ph.D. dissertation, Universiteit Leiden.

Palancar, Enrique L. 2002. *The origins of Agent markers*. Berlin: Akademie Verlag.

Payne, Doris L. & Immanuel Barshi. 1999. *External Possession*. Amsterdam/Philadelphia: John Benjamins.

Quesada Pacheco, Miguel Angel. 2008. *Gramática de la lengua guaymí (ngäbe)*. Munich: Lincom Europa.

Quesada, Juan Diego. 2000. *A Grammar of Teribe*. Munich: Lincom Europa.

Quesada, Juan Diego. 2012. *Gramática del buglere*. Quito: Abya-Yala.

Stassen, Leon. 2009. *Predicative Possession*. Oxford: Oxford University Press.

Trask, Robert Lawrence. 1979. On the Origins of Ergativity. *Ergativity: Towards a Theory of Grammatical Relations*, ed. by Frans Plank, 385–404. London: Academic Press.

Traugott, Elizabeth Closs & Bernd Heine. 1991. *Approaches to Grammaticalization*. Amsterdam/Philadelphia: John Benjamins.

Trillos Amaya, María, Gerardo Reichel-Dolmatoff & Carolina Ortíz Ricaurte. 1989. *Sierra Nevada de Santa Marta*. Vols. 1–2. Bogotá: Colciencias.

Uhle, Max. 1890. Verwandtschaften und Wanderungen der Tschibtscha. *Congreso Internacional de Americanistas* session VII, Berlin, 1–6 October, 1890, 466–489. Berlin: Librairie W. H. Kühl.

Young, Philip D. & Talmy Givón. 1990. The Puzzle of Ngóbére Auxiliaries: Grammatical Reconstruction in Chibchan and Misumalpan. *Studies in Typology and Diachrony: Papers Presented to Joseph H. Greenberg on his 75th Birthday*, ed. by Joseph H. Greenberg, William Croft, Keith M. Denning & Suzanne Kemmer, 209–243. Amsterdam/Philadelphia: John Benjamins.

Voice, Transitivity and Tense/Aspect: Directionality of Change in Indo-European (Evidence from Greek and Vedic)

Nikolaos Lavidas and Leonid Kulikov

Abstract

The aim of this article is to examine the directionality of change in Voice in relation to Tense/Aspect, foremost based on evidence from Greek as well as additional evidence from Early Vedic. Starting with the hypothesis that in (standard) Proto-Indo-European a number of innovations resulted in the introduction of some elements of the Perfect-Stative inflection into the Present (cf. Kulikov & Lavidas 2013), we study the directionality of change in Voice. We show that the original relationship between Tense/Aspect and Voice determines the directionality of change in Voice in Greek. Basing our study on the analysis of Vedic active Perfects that are intransitive and belong with middle Presents, we claim that this initial relationship between Voice and Tense/Aspect can be reconstructed on the basis of some tendencies and changes found in several Indo-European dialects, in particular in Greek forms. We also argue that the relationship between Tense/Aspect and Voice in the diachrony of Greek depends on the new features acquired by the voice morphology as well as on the development of the categories Tense and Aspect.

1 Introduction

The hypothesis that the Indo-European (IE) categories *Middle* and *Perfect* are historically related (and probably originate in one single proto-category) goes back as far as Kuryłowicz (1932) and Stang (1932). This hypothesis is based on the fact that the middle voice and the active Perfect endings share a number of characteristics in ancient IE languages (for further discussion see, for instance, Di Giovine 1990–1996; Kulikov 1999). The category of *Stative* has also been appended to the Middle and Perfect relationship (Oettinger 1976; Jasanoff 1978; Di Giovine 1990–1996; Kümmel 1996; Gotō 1997; for the relationship between Perfect, Stative and Middle in PIE, see also Kuryłowicz 1964; Kortlandt 1979, 1981). Kulikov (1999) has drawn attention to a particular type of relationship

between Tense and Transitivity that can be called "split causativity." This tendency, observable for some verbs in Vedic and Homeric Greek, characterizes basic correlations between the tense oppositions (Present/Perfect), on the one hand, and Transitivity of the verbal form, on the other:

> Verbal formations of the Present system : transitive-causative
> Verbal formations of the Perfect system : intransitive

This tendency could result in a number of secondary (and, at first glance, inexplicable) uses of some forms, such as, for example, the use of Perfect forms in the function of intransitive Present forms, as in the case of the early Vedic "Perfecto-Presents." This formation includes not only the handbook example *véda* 's/he knows', but also a less studied group of forms with a long reduplication syllable, which are mostly or exclusively employed with present resultative (stative) meanings; e.g., *jar* 'become awake' – *jāgā́ra* 'is awake' (← 'has awoken'), *dī* 'shine' – *dīdā́ya* 'shines'. Notice that a few such Perfects sporadically use the long reduplication to emphasize both the present (stative) meaning as opposed to the preterital (pret.) usages of the Perfect of the same root and their prevailingly intransitive syntax; cf. *tan* 'stretch': *tatā́na* 'has stretched (pret.), stretches (pres.)', ~ *tātā́na* 'stretches (pres.)' (cf. ex. (2) below) or *vṛt* 'turn': *vavárta* 'has turned (pret.), turns (pres.)' ~ *vāvárta* 'turns (pres.)'. See Delbrück (1888: 297); Kümmel (2000: 21–22, with fn. 10, 191–194, 208–211, 227–230, 462–469 et passim); Kulikov (2005: 439). Most importantly, in this case we are dealing with, in fact, the embryo of a separate tense category "perfecto-present" (Kulikov 2005: 450, note 18).

Given this assumption (cf. Kulikov & Lavidas 2013), the active/middle opposition would have been not relevant to Perfect forms in early Proto-Indo-European (PIE) (Figure 7.1). Accordingly, leaving aside the difficult issue of the chronological localization of the emergence of the PIE aorist, we can tentatively present the structure of the early PIE verbal system as stage I in the scheme below.[1] In (standard) PIE, a number of innovations (resulting from a contamination and/or analogical rebuilding of endings belonging to different sets) resulted in the introduction of some elements of the Perfect-Stative inflection into the Present (Kortlandt 1979). These forms must have retained the

---

1 With regard to the Aorist, there are some reasons to believe that the Aorist is a more recent addition to the early PIE verbal system, probably going back to some nominal formations (Kortlandt 2009, 2010). Yet, much remains unclear about the exact status and origin of this tense form in the PIE verbal system. However, this issue goes beyond the scope of the present article and cannot be discussed here at length.
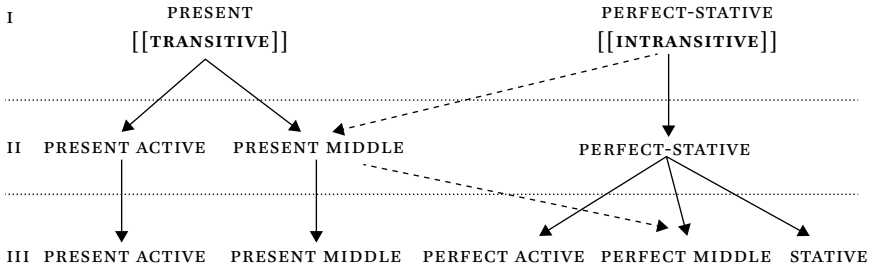
FIGURE 7.1   *The emergence and development of the middle in IE* (adopted from Kulikov 2006). *Splits* denote the reanalysis of a category (for instance, the reanalysis of the Present, that was identified with the transitive patterns, into two types of Present, active and middle Present); *Lines* express the development of a category, and *dotted lines* innovation [contamination and/or analogical rebuilding of endings belonging to different sets] with the extension of a category (for instance, with the introduction of some elements of the Perfect-Stative inflection into the Present).

functional connection with the statives that were associated with intransitive syntax. Notice that this feature of the Perfect-Stative fits well with the recent analysis by Dahl (2010) of the meaning of the Vedic (and, to some extent, PIE) Perfect as largely based on the resultative semantics. This connection may be one of the mechanisms that had given rise to the middle voice used to mark several intransitive derivations. In some IE dialects (Stage III), the active/middle distinction was introduced into the Perfect under the influence of the Present (Renou 1925: Ch. 5–8; Jasanoff 1978: 16, 81f.; Kümmel 2000: 94).[2]

The aim of the present study is to show *how this original relationship between Transitivity/Voice and Tense determines the directionality of changes in Voice in Greek*. Taking as a starting point the analysis of Vedic active Perfects that are used intransitively, and, from the syntactic point of view, can be grouped with middle Presents, we will argue that *this initial relationship between Transitivity/Voice and Tense/Aspect is reflected in Greek in the form of some tendencies and relics*. Moreover, we will show that the correlation between Tense/Aspect and Voice/Transitivity triggers the development of some new features of the category of voice as well as a number of new developments of the categories Tense and Aspect. In the following section, we provide further evidence for

2   We thank the reviewer for the remark that the contrast between the more ancient transitive Presents and intransitive Perfects is also represented in the distinction between Greek intransitive Perfects, which are more ancient, and transitive new formations. Cf., for instance, *pépeika* (persuade:PF.1SG) 'I have persuaded (somebody)' vs. *pépoitha* (persuade:PF.1SG) 'I trust, I am persuaded' (Luraghi, Pompei & Skopeteas 2005).

*the correlation between Tense/Aspect and Transitivity/Voice, which is the basis
of our hypothesis.* Section 3.1 describes the status of Tense/Aspect in Homeric
Greek as well as the changes in this category in Classical and Koine Greek. In
Section 3.2, the historical relationship between Transitivity/Voice and Tense/
Aspect is presented through a discussion of their parallel development and
directionality of changes in Greek. Section 4 summarizes the main conclusions
of this study.

2   Correlations between Tense/Aspect and Transitivity/Voice in Ancient Indo-European: Preliminary Remarks

The prevailing intransitivity of the Perfect forms in Homeric Greek was re-
peatedly observed in historical grammars and studies of the Greek verb
(Wackernagel 1904; Chantraine 1927; Bader 1972; Kulikov 1999), cf. (1a) vs. (1b).
The same holds true for the observation that *active* Perfects are intransitive
and belong with *middle* Presents (see Kulikov 1999).[3]

(1)   a. *pántas   mén   rh'   élpei*
          all:ACC.PL   PTC   PTC   hope:ACT.PRS.3SG
          'She holds out hope to all.' (lit. '**makes** all **hope**') (Hom. Od. 2.91; 8th c. BC)

       b. *mál'   éolpas   enì   phresí*
          very   hope:ACT.PF.2SG   in   mind:DAT.SG
          'Certainly, you **hope** in your mind ...' (Hom. Il. 21.583; 8th c. BC)

In Vedic, a number of verbs such as *tan* 'stretch' or *ukṣ/vakṣ* 'grow, increase' dis-
play a comparable distribution of syntactic patterns: the forms of the Perfect
system mostly appear in intransitive (anticausative) constructions, whereas
the corresponding forms of the Present system mainly attest transitive syntax;
see Kulikov 1999: 26ff. for details. Cf. (2a) vs. (2b).

(2)   a. *ahám   rudrā́ya   dhánur   ā́   tano-mi*
          I:NOM   Rudra:DAT.SG   bow:ACC.SG   to   stretch:PRS-1SG.ACT
          'I stretch the bow for Rudra.' (RV 10.125.6)

---

3   Cf., for example, Velten (1931: 239, fn. 32): "Active Perfect forms with an intransitive meaning –
    often used as a Present like *dédorka* 'I see' – occur commonly beside medio-passive Presents
    [...] the Perfect itself is of durative character and serves as a device of durativation."

b. *dūrā́t*   *sū́ryo*   *ná*   *śocíṣā*   **tatān-a**
   from.afar   sun:NOM.SG   like   flame:INS.SG   stretch:PF-3SG.ACT
   'From afar [Agni] is/has extended (resultative), like the sun, with [his] flame.' (RV 6.12.1)

The ratio of syntactic patterns (transitive/intransitive) attested with the verb *tan* 'stretch' which is one of the most instructive examples, is shown in Table 7.1 and 7.2 (adopted from Kulikov 1999): the bigger font used for Vedic forms in Table 7.2 shows that transitive-causative Presents and intransitive Perfects are more common than intransitive Presents and transitive-causative Perfects.

Drawing on such phenomena, we can assume that the initial relationship between Tense/Aspect and Voice is evidenced in the Indo-European dialects in the form of certain (weak) tendencies and archaisms. Furthermore, the Indo-European dialects follow their own path of development both with regard to the features of the new active vs. middle opposition and the relation of Transitivity/Voice to Tense/Aspect. This means that the directionality of change of the hypothesized relationship between Tense/Aspect and Transitivity/Voice depends on the new features (or values) that the voice morphology acquires,

TABLE 7.1   Transitivity and Tense in Vedic: the ratio of transitive/intransitive constructions in the Present and Perfect in the R̥gveda

	Intransitive	Transitive
Present	7	≈ 40
Perfect	≈ 25	≈ 15

TABLE 7.2   Transitivity and Tense in Vedic: Predominantly transitive-causative Presents vs. intransitive Perfects[a]

	Present		Perfect
	Present indicative	Present subjunctive	
Intransitive	*tanóti* etc.	–	*tatā́na* etc.
Transitive-Causative	*tanóti* etc.	*tanavāvahai* etc.	*tatā́na* etc.

a   For the sake of simplicity, we do not show in this table other forms of the Present system.

as well as on the development of Tense/Aspect (and its relation to Voice) in the IE dialects. In the following section, we will examine the nature of the relationship between Tense/Aspect and Transitivity/Voice in Homeric Greek, and we will provide evidence for the directionality of the development of this relationship at the later stages of Greek. We will argue that the PIE relationship between Transitivity/Voice and Tense/Aspect is reflected in Greek in the form of some tendencies, and that the directionality of the change of the relationship between Tense/Aspect and Transitivity/Voice depends on the new features acquired by the voice morphology as well as on the development of the categories Tense and Aspect. As we locate and trace the contribution of the new features to the direction of changes within the linguistic system, we are also able to provide a more adequate and reliable reconstruction of the PIE stage.

3   The Development of the Hypothesized PIE Relationship between Tense/Aspect and Transitivity/Voice: Correlations between Tense and Transitivity/Voice in Greek

3.1   *Tense and Aspect in Greek*

In this section, we present data from Greek aiming at analyzing the development of the hypothesized PIE relationship between Tense/Aspect and Transitivity/Voice.[4] The development of this relationship is complex because both Transitivity/Voice and Tense/Aspect follow various paths in different ancient IE dialects. However, the study of changes in the features of Tense/Aspect and Transitivity/Voice in Greek can reveal many aspects of both the PIE verbal system and the nature of changes in IE languages.

Based mainly on Moser (2005, 2008, 2014), the main characteristics of the Tense/Aspect system of Homeric Greek can be summarized as follows: The Perfect is often not distinguished from the Present in Homeric Greek and has the same interpretation as the corresponding Present;[5] see (3). Furthermore, the Perfect almost never expresses a resultative meaning in Homeric Greek (on Perfect in Homeric Greek, see also Monro 1891, Schwyzer & Debrunner 1950 and Chantraine 1953). The Present in Homeric Greek displays greater variation than the Perfect with regard to its interpretations

---

4 We refer to the following periods of Greek: Homeric Greek: 8th c. BC; Classical Greek: 5th–3rd c. BC; Hellenistic and Roman/ Koine Greek: 3rd c. BC–4th c. AD; Early Byzantine Greek: 5th–8th c. AD.

5 Schwyzer & Debrunner (1950: 227) argue that it is (only) the Middle which alternates with the Perfect.

and functions, but verbs denoting achievements are rare with the Present. Two verbs with the same function differing in lexical aspect, or actionality (Aktionsart) – accomplishment vs. activity, for instance – are not found in the same tense, but the accomplishment is attested in the Aorist and the activity in the Present (*kálupsen – anapálletai* in Ex. 4).[6]

(3) *epeì   polù   **boúlomai**   autền   oíkoi*
    because   much   want:PRS.1SG   3SG.F.ACC.SG   house:DAT.SG
    *ékhein.   Kaì   gár   rha   Klutaimnèstrēs   **probéboula***
    have:INF   and   PTC   PTC   Klytemnestra:GEN.SG   prefer:PF.1SG
    'Because I very much want to have her at home. For I prefer her to Klytemnestra.' (Hom. Il. 1.112–113; 8th c. BC – Moser 2008)[7]

(4) *hos   d'   hoth'   hupò   phrikòs   Boréō*
    as   PTC   when   by   gust:GEN.SG   north.wind:GEN.SG
    anapálletai   *ikhthùs   thín'*   *en*
    shudder:PRS.3SG   fish:NOM.SG   sand:DAT.SG   in
    phukióenti,   mélan-dé   he
    seaweed.covered:DAT.SG   black:ACC.SG-and   3SG.M.ACC.SG
    *kûma   **kálupsen**   hòs   plēgeìs*
    wave:ACC.SG   cover:AOR.3SG   thus   hit:AOR.PRT.NOM.SG
    anépalt'(o)
    shudder:AOR.3SG
    'As when the fish shudders in a gust of the north wind in the seaweed-covered sandy sea and the black wave covers it, thus did he shudder when hit.' (Hom. Il. 23.692–694; 8th c. BC – Moser 2008)

Verbs in the Aorist in Homeric Greek can appear without the past tense augment, which is obligatory in Classical Greek (see below). Moreover, many verbs in Homeric Greek are attested in only one of the stems, either Present, Aorist, or Perfect. According to Moser (2014), Homeric dictionaries supply the entire paradigm but point out that some forms are not attested in Homeric

---

6 *kálupsen* here denotes the event triggering the *anapálletai*-event.

7 Latacz (2003) considers *probéboula* to be a resultative: "Ganz recht, ich hab' sie Klytaimestra vorgezogen."

TABLE 7.3   The number of verbs (types) that appear in Present, Aorist, and Perfect in
Homer's Iliad[a]

	Present (3sg – indicative)	Aorist (3sg – indicative)	Perfect (3sg – indicative)
Number of verbs			
(types)	34.78%	54.62%	10.60%
(1472)	(512)	(804)	(156)

a   The tables are based on searches in electronic corpora of Homeric, Classical, and Koine
Greek: PROIEL (http://www.tekstlab.uio.no:3000/), Perseus Digital Library (http://www
.perseus.tufts.edu) and Perseus under Philologic (http://perseus.uchicago.edu/), The Homer
Chicago (http://digital.library.northwestern.edu/homer/), and TLG online (http://www.tlg
.uci.edu/).

Greek.[8] In Table 7.3, we present a corpus study that we conducted with regard
to the types (lemmas) that appear in the Present, Aorist, and Perfect in Homer.
This comparison is based on the hypothesis that the Present-Aorist-Perfect
tripartite opposition in Homeric Greek was based on Aktionsart and denoted
duration/ non-terminativity, instantaneity/ terminativity, and stativity, respec-
tively (see below). Table 7.3 shows that there is a statistically significant dif-
ference between the number of types (lemmas) that appear in the different
tenses in Homer.

Moser (2005, 2008, 2014) claims that the morphological tense oppositions
in Ancient (Homeric and Classical) Greek express the opposition between the
stative and dynamic interpretations. In later stages of the language, this dy-
namic interpretation is further differentiated with respect to telicity and dura-
tion. Cf. Moser (2014: 76) and Figure 7.2:

---

8   With regard to the counter-examples, cf. Moser (2014: 76–77):
"This is not to say that Homeric Greek presents a homogeneous picture of an
Aktionsart-based system. On the contrary, the aspectual system is already well-established,
as shown in Napoli (2006). The epic, however, is a multi-layered text, not only with elements
from different dialects but also with elements from different periods. At the time the Iliad
and the Odyssey were written down, they had already been circulating as oral poetry for
centuries. Due to the formulaic nature of oral epics, some of the earlier linguistic characteris-
tics were preserved (see, for instance, Horrocks 2007). It is those elements differing from the
norm that can point us to older stages in the history of the language."

All this leads to a conclusion similar to that of Sihler (1995, 445), accord-
ing to whom the Homeric situation – in conjunction with the data from
other Indo-European languages – suggests that, in earlier stages, the
Present-Aorist-Perfect tripartite opposition was based on Aktionsart, ex-
pressing respectively duration/ nonterminativity, instantaneity/termina-
tivity, and stativity.

Situations

Stative
(Perfect)

Dynamic

Durative/atelic   Instantaneous/telic
(Present)   (Aorist)
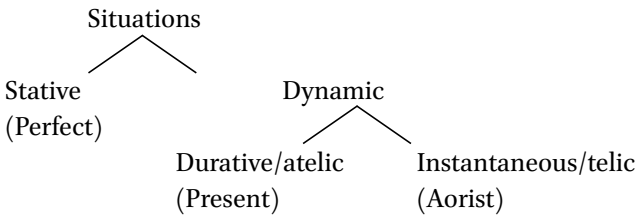
FIGURE 7.2   The system of Tenses in Homeric Greek. From Moser (2014: 76)

TABLE 7.4   The ratio of verbs (types) in the active vs. mediopassive Perfect in Homer (Iliad
             and Odyssey)

Perfect active – Homer, Iliad	(119/252) 47.22%
Perfect active – Homer Odyssey	(83/226) 36.73%
Perfect active – Homer (total)	**(202/478) 42.26%**
Perfect mediopassive – Homer, Iliad	(133/252) 52.78%
Perfect mediopassive – Homer Odyssey	(143/226) 63.27%
Perfect mediopassive – Homer (total)	**(276/478) 57.74%**

Given the systematic correlation between Tense/Aspect and Transitivity/Voice
and the correlation between Perfect and Middle (see Section 2 and Table 7.4),
we can hypothesize a parallel directionality for the development in Tense/
Aspect and Transitivity/Voice in Greek.

   Before proceeding to the status of Voice in Homeric Greek and the analysis
of its relation to Aktionsart and Tense/Aspect, we refer here to the develop-
ment of Tense/Aspect in Greek. The aim is to provide a basis for a comparison
to the development of Voice in Greek. Note that it is beyond the scope of this
study to analyze the characteristics and development of Tense/Aspect in all
periods of Greek. We focus on the evidence available from Homeric Greek, as
well as changes attested in Classical and Koine Greek, to track the directional-
ity of the relevant changes.

   In Classical Greek, all verbs can have forms based on the Perfect, Present,
and Aorist stem. Uses that show a dependence on Aktionsart in Classical Greek

are absent from Koine Greek (Moser 2008).[9,10] In Classical Greek, the Future forms still do not express the perfective–imperfective opposition, which will become available in Koine Greek with the periphrastic formation of the Future. However, the uses of the Perfect and Aorist as Presents have been lost in Classical Greek – but with some archaic exceptions (for instance, *péphuka* 'be by nature'); in addition, the past augment is obligatorily used in the Aorist forms and the Perfect acquires a resultative interpretation.

Koine Greek (and Early Byzantine Greek) demonstrate the most significant changes in the Tense/Aspect system (and we will observe that the same holds true for the relationship between Tense/Aspect and Voice as well). In Koine Greek, the synthetic Perfect and Future are replaced with periphrastic forms, see Blass, Debrunner & Rehkopf (1975 [1984]: §§ 340–356), Moser (1988). All new periphrastic constructions express the opposition of perfective–nonperfective, creating the new status of the Tense/Aspect system: verbal forms are linked to the aspect (grammatical aspect/outer aspect) and not to the Aktionsart.

To summarize the discussion above, (a) in Homeric Greek, the morphological tense oppositions show certain relationships with Aktionsart (as archaisms), and (b) in Koine Greek, we observe several features of the innovative relation of the verbal forms to the (grammatical) aspect. In Section 3.2, we will show that Voice in Greek changes in a similar way to the morphological tense oppositions (which were linked to Aktionsart/lexical aspect in Homeric Greek but to (grammatical) aspect in Koine Greek). We will show that the mediopassive morphology in Homeric Greek was linked to the absorption of any argument (agent, cause, benefactive), but changed into an intransitivity marker, blocking the presence of an object in the accusative, in later stages (Lavidas 2012; Lavidas et al. 2012). This means that intransitivity was not rigidly connected with the (mediopassive) voice in Homeric Greek. This is particularly evident in the case of active forms with passive interpretation or mediopassive

---

9   Cf. van Gelderen (2004: 203ff), who argues for a similar change in the history of English where the setting of a parameter is switched from having (inner/ lexical) aspect as unmarked to tense as unmarked.

10 Cf. Moser (2014: 77):

      "In Classical Greek, the picture has changed quite dramatically. Verbs now possess full paradigms, with instances of practically every form attested in the very large corpus of texts. These forms are used with considerably greater freedom in order to express the speaker's vantage point, i.e., grammatical aspect. But aspect has not gained complete independence from Aktionsart: the latter still plays an important role in determining choice."

forms that were transitive and could take a direct object in the accusative, with an autobenefactive interpretation (see below).

3.2 *Transitivity/Voice and Tense/Aspect in Greek*

The relation of Voice to Tense/Aspect is evident in Homeric Greek. Homeric Greek has a system of distinctions between three voices: active, middle, and passive, but the middle and the passive morphology are distinguished *only in the Future and Aorist* (Aorist stem); Table 7.5.

TABLE 7.5   Middle and passive endings in the Future and Aorist (Aorist stem)

	Future	Aorist
Middle endings	*-somai*	*-samēn*
Passive endings	*-thē-somai*	*-(th)ē-n* *-thē-samēn*

From a purely morphological perspective, an additional non-active morpheme, *-thē*, can be distinguished in the passive type of the Future and Aorist (Aorist stem): *lu-thḗ-somai* (unbind/release-PASS-MP.FUT.1SG) vs. *e-lú-thē-n*[11] (AOR-unbind/release-PASS-1SG, cf. Chantraine 1953, 1961). In the Future, the morphemes *-thē-* and *-omai* express the non-active, whereas in the Aorist, only the morpheme *-thē-* expresses the non-active. Moreover, the suffix *-ē-* of the Aorist (as in *ekseplág-ē-n* 'I was struck with terror or amazement') should have started as an Aorist suffix for active verbs of the *-mi*-conjugation (*-ē-* started as an Indo-European suffix of stative verbs). We focus here on the directionality of change in its functions. However, at a second stage, *-ē* was used for encoding intransitive derivatives (reflexives, anticausatives – but not passive) with verbs that take active endings and, at a third stage, as a non-active (reflexive, anticausative, and passive) suffix (although it was less frequent in the passive function than *-thē-*, cf. Allan 2003); see (5).

---

11   Humbert (1945) has argued that the new formation of the passive Aorist with *-thē-* has been completed in Homer, that the passive Future with *-thēsomai*, unknown in Homer and Herodotus, is not in evidence before Aeschylus. The passive Future was formed on the basis of the Aorist.

The suffix *-ē-* is used mainly in reflexive and anticausative constructions in Homer: of the 22 instances mentioned by Delbrück (1897), only two have a purely passive interpretation (*eplḗgēn* 'I was smitten' / *etúpēn* 'I was beaten'); see (6). All of the remaining examples are not passive; many, in fact, denote spontaneous change-of-state (*pagênai* 'become solid/stiffen', *ragênai* 'break', *tmagênai* 'be po divided') as is shown in (7). Hence, there appears to be a connection between the early suffix *-ē-* and anticausativity.[12]

(5) *ekséplēksa* (transitive)   /   *ekseplág-ē-n* (intransitive)
     'I struck with terror or amazement.'   'I was struck with terror or
                                              amazement.'

(6) *hélkea*   *pánta*   *mémuken*   *hóss'*
     wounds:NOM   all:NOM   heal:ACT.PF.3SG   which:NOM
     etúpē
     strike:PASS.AOR.3SG
     'All his wounds have been closed up where he was struck.' (Hom. Il. 24.
     420–421; 8th c. BC)

(7) *ouranóthen*   *d'*   *ár'*   ***huperrágē***   *áspetos*
     from.sky   PTC   PTC   break:PASS.AOR.3SG   endless:NOM
     aithḗr
     bright.air:NOM
     'And from heaven breaks open the infinite air.' (Hom. Il. 16.300; 8th c. BC)

Homeric Greek verbs with the morpheme *-thē-* are mainly reflexives, anticausatives, and, in rarer cases, passives. Grosse (1889) refers to only 30 examples with a purely passive interpretation (*ktathênai* acquire:PASS.AOR.INF) from the 129 examples that he notes (e.g., intransitive: *agerthênai* gather:PASS. AOR.INF). Cf. also relevant tables in the detailed study of the middle and passive in Homeric and Classical Greek by Allan (2003), as well as our discussion of Allan (2003) in Kulikov & Lavidas (2017).

Allan (2003) has shown that passive Aorists in both *-thē-* and *-ē-* have exactly the same interpretations: passive, spontaneous process, mental process, and

---

12   This is not an absolute claim because, in some cases, middle forms of the same verbs in the Present may indicate change-of-state. We thank the reviewer for the discussion of this issue.

TABLE 7.6   Meanings of root and thematic Aorists in Homer (Allan 2003)

a. **Root Aorists:**

Passive	*éktato* 'was killed'
Spontaneous process	*phthímēn* 'perished'
Mental process	*étlēn* 'endured, dared'
Body motion	*âlto* 'jumped'
Collective motion	*ksúmblēto* 'met with'
Speech act	*eûkto* 'boasted, prayed'
Indirect reflexive	*étheto* 'put sth for oneself'

b. **Thematic Aorists:**

Passive	*eskhómēn* 'was held'
Spontaneous process	*ōlómēn* 'perished'
Mental process	*elathómēn* 'forgot'
Body motion	*etrapómēn* 'turned'
Collective motion	*ēgrómetha* 'gathered'
Perception	*ēisthómēn* 'perceived'
Speech act	*ērómēn* 'asked'
Indirect reflexive	*ēgagómēn* 'led away for myself'

(collective and body) motion. According to Allan's results, the sigmatic middle Aorist is not attested with a spontaneous or a passive meaning.[13] Reflexives, however, can be marked with the sigmatic middle Aorist. See also Table 7.6, from Allan, which presents the meanings of root and thematic Aorists in Homer. With regard to Future, according to Allan, the passive Future has a generic/non-iterative/perfective meaning (as demonstrated mainly with Classical Greek examples by Allan).

As far as the distribution of active and mediopassive morphology in Homeric Greek is concerned, mediopassive morphology is considered the marked form in relation to active voice (for instance, by Bakker 1994: 24). The middle morphology is not identified with one construction (reflexive or anticausative). The middle morphology can be used in passive constructions (with the presence or absence of an agent-PP). Nor is the passive type identified with the passive construction. The passive type can be used productively in intransitive non-passive constructions – for instance with psych-verbs – in anticausative constructions. We suppose, therefore, that it is concerned with two different

---

13   We should note that Homeric and Classical Greek have three different morphological types of middle Aorist (root, thematic, and sigmatic) and two morphological types of passive Aorist (in *-ē-* and *-thē-*).

morphological forms of the same non-active category, which are used alternately. The mediopassive morphology is also frequently used in transitive constructions and adds the meaning that the result of the verb action concerns the subject: *títhemai nómon* (place:MP.PRS.1SG law:ACC) 'pass a law *in my own interests*'/ 'pass myself a law' – in contrast to the active *títhēmi* (place:ACT. PRS.1SG) 'pass a law'. In other words, the mediopassive morphology leads to absorption of the benefactive (or, it expresses indirect reflexivity, in traditional terminology). Moreover, the mediopassive verbs in reflexive constructions can take a direct object in the accusative in Homeric Greek, if the object-goal is directly related to the subject; see (8).[14]

(8)   *loúomai* 'wash'
      egkalúptomai 'veil/wrap up'
      peribállomai 'throw round or over oneself/put on'   } + NP-accusative
      peritíthemai 'place or put round/put on'

In Homeric Greek (see also above), an agent in PP or in the dative could appear in constructions with active verbs and an undergoer(theme)-argument as the subject. This concerns a typical instance of a (lexical) passive construction, but with verbs bearing active morphology (Jankuhn 1969). However, in Homeric Greek, the first signs of productive use of mediopassive morphology in passive constructions are also attested.

A corpus study clearly shows that the distribution of voice (active vs. middle vs. passive) *heavily depends on Tense/Aspect*; see Table 7.7. A chi-square was performed to assess the relationship between the voice morphology and the different tenses. The results of the Pearson chi-square analyses were statistically

---

14   The mediopassive voice morphology is also used in transitive constructions with deponents (for an analysis of the deponents in diachrony, cf. Lavidas & Papangeli 2006). We do not examine deponents in this study because these verbs do not change morphology (they always have non-active morphology) for purposes of transitivity alternations.

(1) *egṑ   Kleinían   hḗdion   mèn **theômai**   ḕ   tâlla*
      I:NOM Cleinias:ACC more.pleasantly PTC gaze:MP.PRS.1SG than the.other:ACC
      pánta
      all:ACC
      I would rather gaze at Cleinias than at all the other (beautiful objects in the world).'
      (X. Smp. 4.12; 5th–4th c. BC)

TABLE 7.7   Tense/Aspect and Voice in Homer: Percentage of active vs. mediopassive forms in the different tenses/aspects[a]

	Active	Middle (Mediopassive)	Passive
Present	85.16%	14.84%	
(3sg Indicative)	(1245/1462)	(217/1462)	
Aorist	77.07%	18.47%	4.46%
(3sg Indicative)	(4927/6393)	(1181/6393)	(285/6393)
		middle+passive=	
		22.93%	
Imperfect	75.77%	24.23%	
(3sg Indicative)	(3331/4396)	(1065/4396)	
Perfect	69.19%	30.81%	
(3sg Indicative)	(238/344)	(106/344)	
Pluperfect	60.70%	39.30%	
(3sg Indicative)	(227/374)	(147/374)	
Future	40.72%	59.28%	(0[b])
(3sg Indicative)	(191/469)	(278/469)	
Future Perfect	13.33%	86.67%	
(3sg Indicative)	(2/15)	(13/15)	

a   The verbs included in these corpus studies (Tables 7.7, 7.8 and 7.9) are all in the 3rd singular and indicative, to avoid effects in the results by person, number, and mood, or by the nature of participles and infinitives.
b   With regard to all Future forms, and not only 3sg Indicative, there is only one passive Future: *migésesthai* mix:FUT.PASS.INF.

significant for the comparison between the Perfect and all other tenses[15] and for the comparison between the Future and all other tenses.[16]

This means that the ratios of active vs. middle/passive morphology are not similar for all tenses/aspects, but are significantly correlated with the type of Aktionsart that is expressed by each of the tense/aspects. Accordingly,

---

15   vs. Present: χ^2=48.368, p<.001, with an effect size of φ=.164, which is a small effect size; vs. Imperfect: χ^2=7.442, p=.006, with an effect size of φ=.040, which is a small effect size; vs. Aorist: χ^2=11.338, p=.001, with an effect size of φ=.041, which is a small effect size; vs. Future: χ^2=64.497, p<.001, with an effect size of φ=.282, which is a medium effect size; vs. Pluperfect: χ^2=5.661, p=.017, with an effect size of φ=.089, which is a small effect size.
16   vs. Present: χ^2=367.741, p<.001, with an effect size of φ=.436, which is a large effect size; vs. Imperfect: χ^2=260.487, p<.001, with an effect size of φ=.231, which is a medium effect size; vs. Aorist: χ^2=304.467, p<.001, with an effect size of φ=.211, which is a medium effect size; vs. Future Perfect: χ^2=4.549, p=.033, with an effect size of φ=.097, which is a small effect size.

more Future forms (59.28%) appear with middle morphology than with active. The Perfect (and Pluperfect and mainly Future Perfect) show more types with middle morphology than the other tenses do (Present, Imperfect, Aorist) – even though the middle Perfect forms do not constitute the majority of Perfect forms.

Changes in Voice go in parallel with the development of Tense/Aspect in Greek: Koine Greek attests an entirely new system of Voice. This is represented in the differences in the distribution of Voice and Tense/Aspect in Classical (Table 7.8) and Koine Greek (Table 7.9), which can be compared with Homeric Greek (Table 7.7). The purpose of this corpus study is to present data and evidence on the change in the distribution of voice morphology among tenses/aspects. The significance of the study is evident in testing the quantitative representation of a possible relationship between tense/aspect and voice.

TABLE 7.8   Tense/Aspect and Voice in Classical Greek (Plato's works): Percentage of active vs. mediopassive forms in the different tenses/aspects

	Active	Middle (Mediopassive)	Passive
Imperfect	95.38%	4.62%	
(3sg Indicative)	(4769/5000)	(231/5000)	
Present	86.81%	13.19%	
(3sg Indicative)	(10883/12537)	(1654/12537)	
Perfect	77.22%	22.78%	
(3sg Indicative)	(1519/1967)	(448/1967)	
Pluperfect	65.87%	34.13%	
(3sg Indicative)	(83/126)	(43/126)	
Aorist	58.93%	20.73%	20.34%
(3sg Indicative)	(1069/1814)	(376/1814)	(369/1814)
		middle+passive=	
		41.07%	
Future	44.61%	51.04%	4.35%
(3sg Indicative)	(729/1634)	(834/1634)	(71/1634)
		middle+passive=	
		55.39%	
Future Perfect	9.09%	90.91%	
(3sg Indicative)	(2/22)	(20/22)	

Between Classical and Koine Greek, the tables demonstrate an increase of the active forms in the Present, Imperfect, and Pluperfect, but an increase of the mediopassive forms in the Aorist and Perfect. The results of the Pearson chi-square analyses were statistically significant for the comparison between the distribution of voice morphology in (a) the Present (Present in Classical Greek vs. Present in Koine Greek);[17] (b) the Imperfect (Imperfect in Classical Greek vs. Imperfect in Koine Greek);[18] (c) the Aorist (Aorist in Classical Greek vs. Aorist in Koine Greek);[19] and (d) the Perfect (Perfect in Classical Greek vs. Perfect in Koine Greek).[20] Voice morphology in the Future (and Future Perfect for Homeric and Classical Greek) in the different periods shows no statistically significant differences.

TABLE 7.9   Tense/Aspect and Voice in the New Testament (Koine Greek): Percentage of active vs. mediopassive forms in the different tenses/aspects

	Active	Middle (Mediopassive)	Passive
Imperfect	93.22%	6.78%	
(3sg Indicative)	(893/958)	(65/958)	
Present	90.99%	9.01%	
(3sg Indicative)	(2251/2474)	(223/2474)	
Pluperfect	85.71%	14.29%	
(3sg Indicative)	(36/42)	(6/42)	
Aorist	72.64%	12.57%	14.79%
(3sg Indicative)	(2525/3476)	(437/3476)	(514/3476)
		middle+passive=	
		27,36%	
Perfect	54.77%	45.23%	
(3sg Indicative)	(201/367)	(166/367)	
Future	48.66%	27.54%	23.80%
(3sg Indicative)	(364/748)	(206/748)	(178/748)
		51.34%	
Future Perfect			
(3sg Indicative)	(0)	(0)	

17   χ^2=32.986, p<.001, with an effect size of φ=.047, which is a small effect size.
18   χ^2=7.981, p=.005, with an effect size of φ=.037, which is a small effect size.
19   χ^2=102.886, p<.001, with an effect size of φ=.139, which is a small effect size.
20   χ^2=80.451, p<.001, with an effect size of φ=.186, which is a small effect size.

Tables 7.10a–c show the distribution of active, middle, and passive forms according to Tense/Aspect. Active forms are more frequently attested in the Aorist in Homeric Greek, but this picture changes in Classical and Koine (New Testament) Greek, where both Present and Aorist are very frequent with active morphology. With regard to the middle and passive morphology, forms in the Aorist present a decrease (whereas the frequency of the mediopassive in the Present increases).

The frequency of forms (regardless of voice morphology) in the Present increases after Homer, but the ratio between active and mediopassive Presents remains stable. The Perfect has a similar ratio of active and mediopassive in Homeric Greek, but the frequency of Perfect forms becomes higher for the mediopassive than active morphology in the following periods. No statistically significant differences are observed either for Pluperfect or for Future Perfect. A chi-square was performed to assess the relationship between the distribution of tenses in the active/mediopassive voice and different periods. The results of the Pearson chi-square analyses show that there is a change in the distribution of tenses and voice morphology, but that change is not statistically significant.[21]

TABLE 7.10A   Tense/Aspect and Voice in Homer: Percentage of use of the different tenses/ aspects with the different voice morphologies [m: middle / p: passive]

	Present (3sg Indicative)	Imperfect (3sg Indicative)	Aorist (3sg Indicative)	Future (3sg Indicative)	Perfect (3sg Indicative)	Pluperfect (3sg Indicative)	Future Perfect (3sg Indicative)	Total
Active	12.25% (1245)	32.78% (3331)	48.49% (4927)	1.88% (191)	2.34% (238)	2.23% (227)	0.02% (2)	10161 (100%)
Middle (Medio- passive)	7.22% (217)	35.42% (1065)	39.28% (1181)	9.25% (278)	3.53% (106)	4.89% (147)	0.43% (13)	3007 (100%)
	m+p[a] 6.59%	m+p 32.35%	m+p 35.87% (285) 100%	m+p 8.44%	m+p 3.22%	m+p 4.47%	m+p 0.39%	m+p 3292 (100%)
Passive								285 (100%)

a   m+p: if we add the percentage of middle/mediopassive forms (m) and passive forms (p).

---

21   For instance, Aorist in Homeric Greek vs. Aorist in Classical Greek: χ^2=2.008, p=.156; Aorist in Classical Greek vs. Aorist in Koine Greek: χ^2=3.018, p=.082; Aorist in Homeric Greek vs. Aorist in Koine Greek: χ^2=.356, p=.551.

TABLE 7.10B Tense/Aspect and Voice in Plato (Classical Greek): Percentage of use of the different tenses/aspects with the different voice morphologies

	Present (3sg Indicative)	Imperfect (3sg Indicative)	Aorist (3sg Indicative)	Future (3sg Indicative)	Perfect (3sg Indicative)	Pluperfect (3sg Indicative)	Future Perfect (3sg Indicative)	Total
Active	57.12% (10883)	25.03% (4769)	5.61% (1069)	3.83% (729)	7.97% (1519)	0.44% (83)	0.01% (2)	19054 (100%)
Middle (Mediopassive) m+p	45.87% (1654) / 40.88%	6.41% (231) / 5.71%	10.43% (376) / 9.29% (369)	23.13% (834) / 20.61% (71)	12.42% (448) / 11.07%	1.19% (43) / 1.06%	0.55% (20) / 0.49%	3606 (100%) m+p / 4046 (100%)
Passive			83.86%	16.14%				440 (100%)

TABLE 7.10C Tense/Aspect and Voice in New Testament (Koine Greek): Percentage of use of the different tenses/aspects with the different voice morphologies

	Present (3sg Indicative)	Imperfect (3sg Indicative)	Aorist (3sg Indicative)	Future (3sg Indicative)	Perfect (3sg Indicative)	Pluperfect (3sg Indicative)	Future Perfect (3sg Indicative)	Total
Active	35.90% (2251)	14.24% (893)	40.27% (2525)	5.81% (364)	3.21% (201)	0.57% (36)	0	6270 (100%)
Middle (Mediopassive) m+p	20.22% (223) / 12.42%	5.89% (65) / 3.62%	39.62% (437) / 24.35% (514)	18.68% (206) / 11.48% (178)	15.05% (166) / 9.25%	0.54% (6) / 0.33%	0	1103 (100%) m+p / 1795
Passive			74.28%	25.72%				692 (100%)

The following changes in Voice in Koine Greek can illustrate the development of the relationship between Voice and Tense/Aspect and the new status of this relationship (which appears in the quantitative data above). From the Koine Greek period, the marking of the autobenefactive interpretation by the mediopassive morphology is lost.[22] The active and mediopassive transitive forms are used, one next to the other, even in the same sentence, without an essential difference in the interpretation, as is shown in (9).

(9) ***timômen***   *tòm*   *múrion*   – ***timṓmetha***
     estimate:ACT.PRS.1PL ART.ACC numberless:ACC – estimate:MP.PRS.1PL
     tòn   *múrion*
     ART.ACC   numberless:ACC
     (In both cases, the verb means 'calculate/estimate') (Mayser 1926: 112)

With regard to the middle and passive morphology, which are distinguished only in the Aorist and Future, free alternation is still observed in Koine Greek, but the new element for this period is the extension of the passive morphology. For instance, the following deponents and intransitives have passive Futures and Aorists (instead of middle) in the Roman papyri (Chatzidakis 1892 [1975]: 193–200):

(10) *ēisthánthēn* feel:PASS.AOR.1SG (PMich. 486.7; 2nd c. AD)
      elupḗthē be.grieved:PASS.AOR.3SG (PMich. 497.15; 2nd c. AD)
      melēthêis care:PASS.AOR.2SG (PMich. 466.35; 2nd c. AD)

The new tendency for verbs in anticausative constructions in Koine Greek, and mainly in Early Byzantine – after the change in the system of voice marking is completed – is to be marked with active suffixes and not with mediopassive as in Homeric Greek. Hence, the active voice begins to be extended to anticausatives that participate in transitivity alternations. The change concerns only the anticausative morphology (Undergoer + *anoígetai* open:MP.PRS.3SG → Undergoer + *anoígei* open:ACT.PRS.3SG) without other syntactic changes in the existing alternation, as is shown in (11) below:

---

22   On the contrary, from Classical Greek onwards, the passive construction becomes more productive (more verbs can appear in a passive construction). Cf. Luraghi (2010: 70): "Passive became increasingly obligatory, and its extension proceeds from prototypically transitive verbs with accusative objects, to verbs with lower degrees of transitivity with non-accusative objects."

(11) Active transitive

    a. *eksaí[phnēs]* ***anoígō***   ***anugōi***   *toùs*
       suddenly   open:ACT.PRS.1SG open:ACT.PRS.1SG ART.ACC
       ophthalmoús mou
       eyes:ACC 1SG.GEN
       'And suddenly I open my eyes.' (UPZ 1.78 rp1)

    Active intransitive

    b. *hoútōs ouk* ***anoígei***   *tò*   *stóma*   *autoû*
       so   not open:ACT.PRS.3SG ART.NOM mouth:NOM 3SG.GEN
       'So his mouth does not open.' (NT, Act.Ap. 8.32; 1st c. AD)

We have demonstrated an interrelation and a parallelism in the development of Tense/Aspect and Voice in the history of the Greek language, which follows the original relationship (hypothesized for PIE) between Tense/Aspect and Voice. The lexical aspect (Aktionsart) is one of the categories that determine the developments in the verbal system of Homeric Greek and, in particular, is also relevant to several features of Voice. After verbal forms became linked to the (grammatical) aspect in Koine Greek, the encoding of the valency-reducing and valency-increasing derivations through voice endings becomes more regular in Greek verbs, creating a morphological opposition between transitives and intransitives. The active marks the transitive, whereas the mediopassive the intransitive: Transitives (with autobenefactive interpretation) are not marked with mediopassive morphology, whereas the new tendency is for active anticausatives to be marked with active morphology.

4   Conclusion

Starting with the hypothesis that the Perfect (Tense/Aspect) and the Middle (Voice) are historically related and that in (standard) Proto-Indo-European a number of *innovations* resulted in the introduction of some elements of the Perfect-Stative inflection into the Present system (cf. Kulikov & Lavidas 2013), we have examined the directionality of changes in the domain of Tense/Aspect and Voice in Greek.

    We have shown how the original relationship (hypothesized for Proto-Indo-European) between Tense/Aspect and Transitivity/Voice determines the direction of changes in Vedic and Greek. We argued that this (hypothesized) initial correlation between Transitivity/Voice and Tense/Aspect (intransitive perfect ~ transitive-causative present formations) is reflected both in Vedic and Greek in the form of some tendencies and relics. Taking as a basis for our study the

analysis of Vedic active Perfects that are used intransitively and syntactically belong with middle Presents, we have also demonstrated how this relationship depends on the new features acquired by the voice morphology as well as on the development of the categories Tense and Aspect. More specifically, we can conclude that historical evidence from Vedic and Greek gives us good reasons to believe that the original correlation between Tense/Aspect and Transitivity/Voice determines the direction of further changes in the corresponding domains of the morphological system of these languages: the emergence and expansion of the new productive markers of transitivity oppositions (passive/active and anticausative/causative) aiming to avoid collapse of the original system of oppositions by replacing the waning category (in our case, 'split causativity'). In other words, the linguistic material from the documented history of Greek, Vedic (and perhaps some other ancient Indo-European languages) furnishes important evidence for the non-arbitrary character (directionality) of certain changes in these domains of the linguistic system.

Furthermore, we have argued that in Homeric Greek, the morphological oppositions traditionally thought of as manifesting the category of Tense show certain relationships with Aktionsart, whereas the middle (mediopassive) voice, presumably originating in the Perfect/Stative part of the paradigm, still remains sensitive to Aktionsart characteristics. We have demonstrated that changes in Voice go parallel with Tense/Aspect developments in Greek: Koine Greek shows an entirely new system of Voice and Tense/Aspect. As verbal forms become linked to the (grammatical) aspect, the relic correlation between Transitivity/Voice and Tense/Aspect disappears, and the distribution of voices follows a common pattern (for all voices) that favors the active morphology, which is now exclusively related to transitives, in cases of transitive alternations, rather than the (aspectual) characteristics of the tenses. This, again, can serve as evidence for the non-random character of historical changes in the linguistic system of Koine Greek, triggered by the original, albeit disappearing, structure of this domain, which still preserves archaic traces of the original category Perfect/Stative that determines a plethora of new trends in the system of encoding of transitivity oppositions.

Altogether, diachronic evidence from the documented history of Greek and Vedic provides us with valuable data for the general theory of historical linguistics and for substantiation of the fundamental claim about the directionality (i.e. non-arbitrary character) of linguistic changes.

Acknowledgements

Abbreviations

The glosses follow Leipzig Glossing rules. Additionally, the following glosses have been adopted: MP – Mediopassive (non-active); PTC – particle.

References

Allan, Rutger. 2003. *The Middle Voice in Ancient Greek: A Study in Polysemy*. Amsterdam: Gieben.

Bader, Françoise. 1972. Parfait et moyen en grec. In: *Mélanges de linguistique et de philologie grecque offerts à Pierre Chantraine*, [ed. by Françoise Bader], 1–21. Paris: Klincksieck.

Bakker, Egbert J. 1994. Voice, Aspect and Aktionsart: Middle and Passive in Ancient Greek. *Voice: Form and Function*, ed. by Barbara Fox & Paul J. Hopper, 23–47. Amsterdam/ Philadelphia: John Benjamins.

Blass, Friedrich, Albert Debrunner & Friedrich Rehkopf. 1975 [1984]. *Grammatik des neutestamentlichen Griechisch*. Göttingen: Vandenhoeck & Ruprecht.

Chantraine, Pierre. 1927. *Histoire du parfait grec*. Paris: Champion.

Chantraine, Pierre. 1953. *Grammaire homérique. II: Syntaxe*. Paris: Klincksieck.

Chantraine, Pierre. 1961. *Morphologie historique du grec*. Paris: Klinksiek.

Chatzidakis, Georgios. N. 1892 [1975]. *Einleitung in die neugriechische Grammatik*. Leipzig: Breitkopf und Härtel.

Dahl, Eystein. 2010. *Time, Tense and Aspect in Early Vedic Grammar: Exploring Inflectional Semantics in the Rigveda*. Leiden: Brill.

Delbrück, Berthold. 1888. *Altindische Syntax*. Halle a. S.: Verlag der Buchhandlung des Waisenhauses.

Delbrück, Berthold. 1897. *Vergleichende Syntax der indogermanischen Sprachen. Zweiter Teil.* Strassburg: Trübner.

Di Giovine, Paolo. 1990–1996. *Studio sul perfetto indoeuropeo,* I–III. Roma: Il Calamo.

Gelderen, Elly van. 2004. *Grammaticalization as Economy*. Amsterdam/ Philadelphia: John Benjamins.

Gotō, Toshifumi. 1997. Überlegungen zum urindogermanischen «Stativ». *Berthold Delbrück y la sintaxis indoeuropea hoy: Actas del Coloquio de la Indogermanische Gesellschaft*, ed. by Emilio Crespo & José Luís, 165–192. Madrid: UAM & Wiesbaden: Reichert.

Grosse, Hermann. 1889. *Beiträge zur Syntax des griechischen Mediums und Passivums*. Dramburg: Druck von Th. Kämpf.

Horrocks, Geoffrey. 2007. The Language of Homer. *A History of Ancient Greek from the Beginnings to Late Antiquity*, ed. by Anastassios-Fivos Christidis, 475–487. Cambridge: Cambridge University Press.

Humbert, Jean. 1945. *Syntaxe Grecque*. Paris: Librairie Klincksieck.

Jankuhn, Harald. 1969. *Die passive Bedeutung medialer Formen untersucht an der Sprache Homers* (KZ; Ergänzungsheft 21). Göttingen: Vandenhoeck & Ruprecht.

Jasanoff, Jay. 1978. *Stative and Middle in Indo-European* (IBS 23). Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.

Kortlandt, Frederik H. H. 1979. Toward a Reconstruction of the Balto-Slavic Verbal System. *Lingua* 49: 51–70.

Kortlandt, Frederik H. H. 1981. 1st sg. Middle *-H$_2$. *Indogermanische Forschungen* 86: 123–136.

Kortlandt, Frederik H. H. 2009. *Baltica & Balto-Slavica*. Amsterdam: Rodopi.

Kortlandt, Frederik H. H. 2010. An Outline of Proto-Indo-European. *Studies in Germanic, Indo-European and Indo-Uralic*, ed. by Frederik H. H. Kortlandt, 37–45. Amsterdam: Rodopi.

Kulikov, Leonid. 1999. Split Causativity: Remarks on Correlations Between Transitivity, Aspect, and Tense. *Tense/Aspect, Transitivity and Causativity: Essays in Honour of Vladimir Nedjalkov*, ed. by Werner Abraham & Leonid Kulikov, 21–42. Amsterdam/ Philadelphia: John Benjamins.

Kulikov, Leonid. 2005. Reduplication in the Vedic Verb: Indo-European Inheritance, Analogy and Iconicity. *Studies on Reduplication*, ed. by Bernard Hurch, 431–454. Berlin: Mouton.

Kulikov, Leonid. 2006. Passive and Middle in Indo-European: Reconstructing the Early Vedic Passive Paradigm. *Passivization and Typology: Form and Function*, ed. by Werner Abraham & Larisa Leisiö, 62–81. Amsterdam/Philadelphia: John Benjamins.

Kulikov, Leonid & Nikolaos Lavidas. 2013. Reconstructing Passive and Voice in Proto-Indo-European. *Proto-Indo-European Syntax and its Development. (Journal of Historical Linguistics 3.1: Special issue*, ed. by Leonid Kulikov & Nikolaos Lavidas), 98–121.

Kulikov, Leonid & Lavidas, Nikolaos. 2017. Between Passive and Middle: Voice and Verbal Classes in Greek and Vedic. *Contrastive Studies in Verbal Valency*, ed. by L. Hellan, A. L. Malchukov & M. Cennamo, 298–325. Amsterdam: John Benjamins.

Kümmel, Martin. 1996. *Stativ und Passivaorist im Indoiranischen* (HS; Erganzungsheft 39). Göttingen: Vandenhoeck & Ruprecht.

Kümmel, Martin. 2000. *Das Perfekt im Indoiranischen*. Wiesbaden: Reichert.

Kuryłowicz, Jerzy. 1932. Les désinences moyennes de l'indo-européen et du hittite. *Bulletin de la Société de Linguistique de Paris* 30(1): 1–4.

Kuryłowicz, Jerzy. 1964. *The Inflectional Categories of Indo-European*. Heidelberg: C. Winter.

Latacz, Joachim (ed.). 2003. *Homers Ilias: Gesamtkommentar*. Munich/Leipzig: K. G. Saur.

Lavidas, Nikolaos. 2012. Passives in the History of Greek: Evidence for the Role of the Passive Suffix. *Folia Linguistica Historica* 33(1): 87–122.

Lavidas, Nikolaos & Dimitra Papangeli. 2006. Deponency in the Diachrony of Greek. *Deponency and Morphological Mismatches*, ed. by Matthew Baerman, Greville G. Corbett, Dunstan Brown & Andrew Hippisley, 97–126. Oxford: Oxford University Press.

Lavidas, Nikolaos, Georgia Fotiadou & Ianthi Maria Tsimpli. 2012. *Active vs. Nonactive Voice in the Greek Diachrony: Real or Apparent Optionality in the Use of Voice Morphology*; *Selected Papers of the 10th ICGL*, ed. by Zoe Gavriilidou, Angeliki Efthymiou, Evangelia Thomadaki & Penelope Kambakis-Vougiouklis, 390–400. Komotini: Democritus University of Thrace.

Luraghi, Silvia. 2010. The Extension of the Passive Construction in Ancient Greek. *Acta Linguistica Hafniensia* 42(1): 60–74.

Luraghi, Silvia, Anna Pompei & Stavros Skopeteas. 2005. *Ancient Greek*. Munich: Lincom Europa.

Mayser, Edwin. 1926. *Grammatik der griechischen Papyri aus der Ptolemäerzeit.* II. *Satzlehere* 1. *Analytischer Teil*. Berlin/ Leipzig: Walter de Gruyter.

Monro, David Binning. 1891. *A grammar of the Homeric dialect* (2nd edn.). Oxford.

Moser, Amalia. 1988. *The History of Perfect Periphrases in Greek*. PhD Dissertation, University of Cambridge.

Moser, Amalia. 2005. *Ἄποψη καὶ χρόνος στην ιστορία τῆς Ἑλληνικῆς* [Tense and Aspect in the History of Greek]. Parousia: National and Kapodistrian University of Athens.

Moser, Amalia. 2008. The Changing Relationship of Tense and Aspect in the History of Greek. *STUF-Sprachtypologie und Universalienforschung* 61(1): 5–18.

Moser, Amalia. 2014. From Aktionsart to Aspect: Grammaticalization and Subjectification in Greek. *Acta Linguistica Hafniensia* 46(1): 64–84.

Napoli, Maria. 2006. *Aspect and Actionality in Homeric Greek: A Contrastive Analysis*. Milano: Francoangeli.

Oettinger, Norbert. 1976. Der indogermanische Stativ. *Münchener Studien für Sprachwissenschaft* 34: 109–149.

Renou, Louis. 1925. *La valeur du parfait dans les hymnes védiques* (Collection linguistique 18). Paris: Champion.

Schwyzer, Eduard & Albert Debrunner. 1950. *Griechische Grammatik*: Zweiter Band; *Syntax und syntaktische Stilistik* (Handbuch der Altertumswissenschaft, 2. Abt., 1. Teil, 2. Bd.). Munich: C. H. Beck.

Sihler, Andrew L. 1995. *New Comparative Grammar of Greek and Latin*. Oxford: Oxford University Press.

Stang, Christian S. 1932. Perfektum und Medium. *Norsk Tidsskrift for Sprogvidenskap* 6: 29–39.

Velten, Hermann V. 1931. On the Origin of the Categories of Voice and Aspect. *Language* 7(4): 229–241.

Wackernagel, Jacob. 1904. Studien zum griechischen Perfectum. *Programm zur akademischen Preisverteilung*, Göttingen, 3–24 [Reprinted in J. Wackernagel, *Kleine Schriften*. Göttingen: Vandenhoek and Ruprecht, 1953, 1000–1021].

On Shared Structural Innovations: the Diachrony of Adverbial Subordination in Semitic

Na'ama Pat-El

Abstract

The Semitic languages share the same pattern for adverbial subordination, but they do not share cognate subordinators. Following widely accepted approaches to syntactic reconstruction, such as Harris & Campbell (1995), it is possible to reconstruct a proto construction for this family, even without cognate material. However, in this article I argue that adverbial subordination cannot be reconstructed to the proto language and the shared structure is a case of parallel development which was motivated by influence from a type of relative clause. I suggest that parallel development was triggered by the presence of a shared structural feature, which created similar pressures in different nodes and allowed for identical lines of development to take place, but nevertheless yielded distinct outcomes. The development of adverbial subordinators as outlined here shows that despite structural similarities in adverbial subordination among the Semitic languages, it is unlikely that this pattern is reconstructable to the proto language.

1 Introduction

Since the renewed interest in syntactic reconstruction, there has been a lively discussion about what it is that historical synctacticians reconstruct, and particularly what kind of linguistic material may be used for reconstruction.[1] The general consensus used to be that form-meaning minimal pairing is not available in reconstructing syntax. However, a number of scholars from a range of methodological approaches have argued that syntactic reconstruction, like the reconstruction of phonology and morphology, is based on strict

---

1 The paper was presented at the workshop on Syntactic Reconstruction in ICHL 20 (Osaka), and I thank participants there for their suggestions. I wish to thank my colleagues Pattie Epps and John Huehnergard, as well as two anonymous reviewers and the editors of this volume, especially Jóhanna Barðdal, for many helpful comments on earlier drafts.

correspondences; that is, syntagms need to be in a regular relationship, not merely showing superficial similarities (Harris & Campbell 1995; Campbell & Harris 2002; Harris 2008; Barðdal & Eythórsson, in several publications, most recently in this volume, Walkden 2014; see also the introduction to this volume for a review). Harris and Campbell and others further argue that what is actually being reconstructed are patterns, i.e., a hypothetical slot-based structure, without committing to lexical content. Thus, shared syntactic features are patterns, not overt forms. In other words, for syntactic reconstruction a higher level of abstraction is needed. Although when exemplifying the methodology, sentences with cognate material are used, Harris & Campbell (1995: 350–351) insist that unless a specific form, such as a verb type, conditions syntax, it is not necessary for sentences to share cognate material in order for a reconstruction to be possible (contra Lightfoot 2002). They suggest using safeguards to ensure comparability, which may include cognate material (though this does not need to be fulfilled), functional identity and morphological cognates (such as verbal morphemes, case endings etc.). The idea that with syntactic reconstruction, patterns are the cognate material has also been advocated in a number of other studies (e.g., Barðdal & Smitherman 2013).

In the following, I will examine a potential pitfall in this methodology through a case study from Semitic. The evidence presented here suggests that while formal identity between syntactic structures can point to a shared inheritance, that inheritance is not necessarily the underlying syntactic structure, but rather what might be called "poise" in the sense of Enfield (2003), namely the starting point of a change, or structural pressure to develop in a certain direction. This pressure may motivate similar paths of change in various languages which develop independently of each other. The independent evolution of similar patterns is referred to in the literature as 'drift' or 'parallel development'. While typically dismissed in studies of linguistic genealogy and reconstruction, I suggest that syntactic parallel development has value for syntactic reconstruction, by exposing underlying shared structures and more importantly, the possible mechanisms that operate in language.

Although the term "parallel development" has been used in linguistics at least since Meillet's 1918 article "convergence des développements linguistiques", the term is poorly defined and may be used to cover several rather loosely related phenomena (as noted already in Malkiel 1981). In many studies, parallel development may cover any similar feature in two languages. For example, Warnow et al. (2006) define parallel development as the development of a linguistic character in a particular set of languages, when "the state for that character arises more than once in the evolutionary history of that set of languages" (2006: 76). In other words, parallel development is when two

languages exhibit the same state, which does not appear in their ancestral origin. This should be distinguished from back-mutation, which is the reappearance of an ancestral state in a number of languages not through borrowing or inheritance. Crowley (1991: 180), who deals with morphosyntax, defines parallel development as "changes in languages that take place independently in separate languages after their separation from the protolanguage in such a way that the daughter languages end up converging structurally, or sharing features that were not originally present in the protolanguage". Recently, Keiser (2009) suggested parallel development is at play to explain similar (phonological) change in isolated speech communities. He defines it as "two or more events of internally-induced linguistic change rather than the spread of a single event" (p. 3), and argues that isolation between speech communities is a conditioning factor for this type of change to take place. It seems, however, that isolation is not a conditioning factor to change itself, but rather a helpful circumstance for the linguist who needs to distinguish between contact-induced change and internal change. Some linguists also assume that the rise of a feature through parallel development can occur independently only once, since it is unlikely, or very rare, for a complex feature to develop more than once in two related languages. (Minett & Wang 2003: 291, 322; Warnow et al. 2006: 80). Minett and Wang (2003) agree, however, that some features, primarily phonological features, are more likely than others to independently develop multiple times (similarly, Warnow et al. 2006: 77).

While the existence of the phenomenon is widely acknowledged, the causes of such a development in related languages are far less clear. Parallel development in various parts of the grammar may have different causes. Phonological parallel development is by far the most frequently reported. Phonological features are often binary and are likely to exhibit parallel development if the sound change is typologically common; these changes are often the result of human anatomy, rather than of the language as a system (Keiser 2009). Especially in the case of phonological changes, there is a clear possibility of chance similarities as a result of random variation (Lass 1997: 351, 380, Keiser 2009: 12). In morphology, features may also be binary and hence a change in either direction could be purely coincidental rather than a result of some structural trait. For example, English *weren't* spread as the common negative past tense of *be* in some unrelated varieties of spoken English, while in others only *wasn't* is used (Wolfram & Schilling-Estes 2003; for a similar case of spread in Semitic, see Hetzron 1976).

In the case of syntax, several issues are at play. The appearance of formally similar structures in related languages is less likely to be a result of an accidental similarity, given that syntactic developments tend to be more complex

and involve more steps than developments in other domains of the grammar. Similar structures in related languages may legitimately be seen as a reflection of inheritance, as Harris and Campbell have suggested. In this article I suggest that there are at least some cases where a reconstruction of an immediate ancestor based on cognate structures is unwarranted, despite structural similarity; what we should reconstruct is the point of origin of the change, rather than a fully formed proto-structure. My discussion assumes the following definition of syntactic parallel development: parallel development is the independent development of similar changes in languages of common genetic origin because of a trait in their common genetic material (cf. Rieseberg & Burke 2001; Wood, Burke & Rieseberg 2005). In contrast to shared innovation, which stems from a single line of development whose outcome is attested in later dialects, parallel development is multiple lines of development which took place in the dialects post-split from their ancestor independently of each other (that is, no borrowing is involved). I further argue that independent parallel development is more likely to recur in languages of a genetically cohesive family, which share the relevant structural features. Syntactic changes due to parallel development are restricted and guided by a shared structure, unlike parallel morphological and phonological developments which are frequently random. The concept of syntactic parallel development is essential to syntactic reconstruction because it has implications for the evaluation of genetic relatedness between languages within a family or a branch, as well as for the probability of a certain change to occur.

In order to provide evidence for my argument, I will review the repeated development of finite adverbial subordination in the Semitic language family. Evidence is taken from both classical and modern languages. In Section 2 I outline the syntax of relative clauses and adverbial subordination in Semitic, highlight the problems and suggest a reconstruction. In Section 3 I discuss some broader implications, drawing on the work of La Polla and Enfield, and Section 4 summarizes and concludes the article.

2   Subordination in Semitic

The Semitic languages are attested since the first half of the third millennium BCE. Some of its members, like Akkadian and Aramaic, have a documented history spanning 2,500 years. The family has a fairly stable grammatical profile. It has two genders (masculine and feminine), and three numbers (singular, dual, and plural) on nouns, adjectives, pronouns and verbs. Syntactically, the family shows a consistent head-dependent word order throughout: VSO;

N-Attribute, and prepositions. The family has a basic three case system, with nominative *-u*, accusative/adverbial *-a*, and genitive/prepositional *-i*. The first two cases mark subject and object in verbal sentences respectively or a predicate in a non-verbal sentence, while genitive is used to mark any nominal dependent of other nominals or prepositions. Nominals have a nasal post-case suffix, which marks *state*, namely, its presence marks the noun as having no dependents, and its absence – as a head with a dependent.

The subgrouping of Semitic below shows a basic split between East Semitic and West Semitic (Huehnergard & Pat-El 2019). East Semitic has no modern survivors. Modern Semitic languages include members of a number of West Semitic nodes: Arabic, Hebrew, Aramaic, Ethiopic, and Modern South Arabian; these are shown in italics in Figure 8.1 below. Except for Hebrew, all modern Semitic languages have a number of distinct dialects.

All the Semitic languages individually exhibit at least two types of subordinated finite sentences: relative and adverbial. The term *finite* refers to the fact that they contain full predication where a subject is either explicit or expressed through agreement pronominal affixes on the predicate. No participial constructions are discussed here. The Semitic relative is superficially known to many typologists, but its features have not been fully accurately described thus far; therefore, a brief overview is in order.

2.1 *The Semitic Relative Clause*

Every branch of Semitic attests to two strategies of relativization: via annexation (construct), i.e., a sentence dependent on a nominal head, or via a relative marker, while the nominal head, if overt, is in apposition to the relative marker.[2] The relative marker inflects for gender, number and case, and reflects the syntax of its overt or covert nominal head, with the exception of state: the relative marker is marked as the head of the relative while the appositional noun is independent. Traditionally the marker was described as a pronoun; this has recently been challenged (Huehnergard & Pat-El 2018). Whether the direct head is a relative marker or a noun, its case is determined by the syntax of the main sentence, not the relative clause. Both these markers are called 'heads' because they carry morphological markers of having dependents (in the sense of Nichols 1986). I will call these types of relatives *nominally-headed* and *marker-headed* relatives. Nominal heads will be glossed as cnst (=construct) in the examples below; the relative marker is always marked as head. Examples 1–4 below are of nominally-headed relatives.

---

2 Abbreviations used in this article: acc=accusative; cnst=construct; du=dual; gen=genitive; impf=imperfect; nom=nominative; obl=oblique; opt=optative marker; pass=passive; pf=perfect; pre=preterit; rel=relative marker; sub=subordinative.
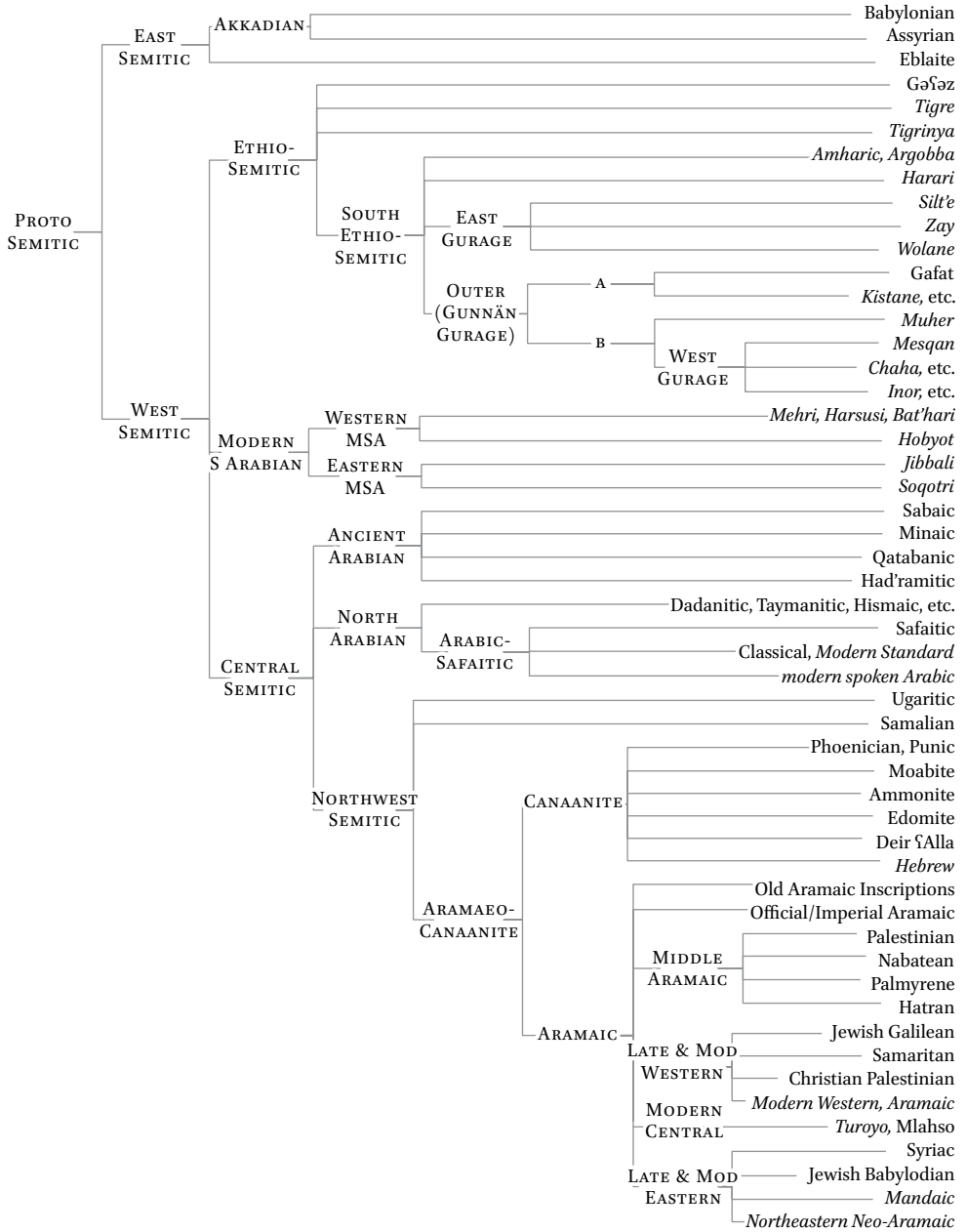
FIGURE 8.1   The internal classification of the Semitic languages

– nominally-headed relative

(1) *qišt-i*   *šarr-um*   *ana* *rēd-îm*   *iddin-u*
     gift-GEN.CNST king-NOM to   soldier-GEN give.PRET.3MS-SUB
     'The gift the king gave the soldier'
                                           (Babylonian Akkadian, Buccellati 1996: 489)

(2) *ʔilā*   *yawm-i*   *yubʕaϑūna*
     to   day.M-GEN.CNST   raise.IMPF.PASS.3MP
     'Until the day they are raised [from the dead]' (Classical Arabic, Qurān 7:14)

(3) *qiry-at*   *ḥānā*   *Dāwīd*
     a town-FS.CNST   camped   David
     'A city in which David encamped' (Biblical Hebrew, Isaiah 29:1)

(4) *ba-mawāʕəl-a* *yəkʷennanu* *masāfənt*
     in-days-CNST   rule   judges
     'In the days in which the judges rule' (Gəʕəz, Ruth 1:1)

Examples (1–2) above are of languages with an active case system, while examples (3–4) above are of languages with reduced or no case system, but where construct nouns are nevertheless morphologically marked. In examples (1–2), the head noun (Akkadian *qišt-i*, Arabic *yawm-i*) lacks the final post-case nasal morpheme which marks a noun as not having dependents. In example (3), the head noun, *qiry-at*, shows a feminine singular morpheme which only surfaces when a noun is a head (an independent form of the same noun would be *qiryā*). In example (4), the head noun *mawāʕəl-a* is marked with a morpheme (*-a*) which indicates a head noun.

Examples 5–8 below are of marker-headed relatives. In this type, the relative clause is dependent on a marker whose inflection is conditioned solely by the syntax of the main clause. The preceding appositional noun may not be overt, but is contextually recoverable. Some form of agreement marker, whether pronominal or part of the verbal agreement morphology, marks the function of the head noun in the relative clause. In the following examples, the relative marker is marked in bold font and the antecedent and agreement markers in the relative clause are underlined.

– marker-headed relatives

(5) **Šarru-kīn** *šar*   *māt-im*   **šu**   *Enlil*   *māḥir-am*
    Sargon   king.CNST   land-GEN   rel.NOM.MS   Enlil   rival-ACC
    lā   *iddin-u-**šum***
    NEG   give.PRET.3MS-SUB-to him
    'Sargon, king of the land, to whom Enlil has given no rival'
                                                    (Babylonian Akkadian, RIME 2.1.1.6.)

(6) *ϑumma* *ʔinna* **walad-ay**-*hi*   **llaðayni**   *qatal-ā-hu*
    then   ADV   child.M-DU.OBL-his   REL.M.DU.OBL   kill.PF-3M.DU-him
    fī Ninwā *harab-ā*[3]
    in Nineveh flee.PF-3M.DU
    'Then, his children, who killed him in Nineveh, fled' (Classical Arabic)

(7) *hinne* *ʔĕlōhê*-*nû* **ze**   *qiwwînû*   *l-ô*
    here   god.M-our   REL   hope.PF.1CP   to-him
    'Here is our god whom we trust.' (Biblical Hebrew, Isaiah 25:9)

(8) **nabiy-āt** *ʔəlla*   *tanabbay-**u***   *həyya*
    prophet-PL   REL.PL   prophesy.PF-3M.PL   here
    'The prophets who offer prophesies here'. (Classical Ethiopic)

Table 8.1 provides the reconstruction of the relative marker to Proto-Semitic. The inflection of the marker is identical to the inflection of adjectives in Semitic, i.e., it includes gender distinction, three cases in the singular but only two in the dual and plural: nominative and oblique. It is morphologically marked as a head (construct).

Both patterns are easily reconstructable to Proto-Semitic and although some attempts have been made to derive one from the other (Deutscher 2009), these attempts are not only highly speculative, but unnecessary. Both patterns are attested in every branch, either productively or as a relic, and their syntax is identical even in languages which no longer use the original Semitic relative marker, like the Canaanite branch (Pat-El 2012). Given the distribution of these relatives in Semitic, we should reconstruct both to Proto-Semitic as follows (the equal sign marks appositional relation; a somewhat similar claim,

---

3 The particle *ʔinna* conditions the accusative. In the dual, the accusative is not distinct from the genitive and is therefore glossed as 'oblique' in example (6).

TABLE 8.1   A reconstruction of the relative marker in Semitic
             (Huehnergard 2006: 112)

	masc.sg	fem.sg
nom.	*ðū*	*ðātu*
gen.	*ðī*	*ðāti*
acc.	*ðā*	*ðata*
	masc.dual	**fem.dual**
nom.	*ðawā*	*ð(aw)ātā*
obl.	*ðaway*	*ð(aw)ātay*
	masc.pl	**fem.pl**
nom.	*ðawū / ʔulū*	*ðawātu / ʔulātu*
obl.	*ðawī / ʔulī*	*ðawāti / ʔulāti*

on synchronic grounds with no diachronic implications, can be found in
Goldenberg 1995):[4]

$$N^{\text{case.cnst}} \qquad S$$
$$N{=}ðV^{\text{case.cnst}} \qquad S$$

Semantically and distributionally, these two strategies are identical (there is
no restrictive/non-restrictive distinction in Semitic); in terms of their internal
syntax, however, they differ. The marker-headed relative may contain either
a verbal or nominal predicate, and has an obligatory anaphoric pronominal
affix which represents the head noun in the relative clause (the so-called 're-
sumptive pronoun'; see underlined elements in examples 5–8 above). This
anaphoric pronoun may be missing only in cases where the antecedent is the
direct object in the relative clause (the phenomenon of direct object argu-
ments omission is known in main clauses as well; Khan 1988). However, the
nominally-headed relative clause may contain only verbal predication and
typically does not make use of anaphoric pronouns. In addition, nominally-
headed relatives tend to be very short, typically containing a single verb.

---

4 S stands for sentence. The equal sign marks appositional status here, that is the head noun
   and the relative marker are appositional to each other.

TABLE 8.2   Relative clause syntax in Semitic

	Marker head	Nominal head
Relation of nominal antecedent to RC	appositional	head
Use of resumption	obligatory	non-obligatory
Type of predication	nominal/verbal	verbal

These relative clauses differ in their relation to their nominal antecedent: the relative clause is dependent on the nominal antecedent of the nominally-headed relative, but appositional to the head of the marker-headed relative. As a result, the noun antecedent in nominally-headed relatives cannot be modified with any additional nominal attribute, e.g., adjectives, possessive suffixes, external possessive constructions, etc. Most languages abandoned or reduced the use of nominally-headed relatives at some point during their history, likely due to these distributional restrictions. See Table 8.2 for a summary of the differences between the relatives.

2.2   *The Semitic Adverbial Clause*

In this section I review the syntax and development of adverbial clauses in Semitic. I argue that they developed from one of the two types of relatives discussed above, the nominally-headed relative. I suggest that loss of inflection is a significant stimulus in this process. I also demonstrate that the innovation of new adverbial subordinators occurred multiple times in these languages.

Adverbial subordinate clauses in all the Semitic languages are marked with a subordinating particle followed by a full predication.

(9) *ašar   iqabbū   kasp-am   l-uddin*
     where   say.3MP.SUB   silver-ACC   OPT-give.IMPF.1CS
     'Let me deposit the silver where they designate'
                                                    (Akkadian, AbB 2 105:12–13)

(10) *ʔaḫṭaʔa-hū   sahm-ī   ḥīna   ramay-tu*
      miss.PF.3MS-him   arrow-my   when   shoot.PF.1CS
      'My arrow missed him when I shot [it]' (Classical Arabic)

(11) *kēn*   *tōʔbēd-ûn*   *ʕēqeb*   *lō*   *tišməʕ-ûn*
      thus   perish.IMPF-2MP   because   NEG   hear.IMPF-2MP
      bə-qôl   YHWH   *ʔĕlōhê-kem*
      in-voice   YHWH   god-your
      'You will thus perish because you have not adhered to YHWH, your God'
                                                           (Biblical Hebrew, Deuteronomy 8:20)

The source of these subordinators is not always clear; however, often the
morphological pattern is indicative of their origin. Almost all of these subor-
dinators are substantives, rather than adverbs or prepositions. For example,
in Hebrew, the particle *ṭerem* 'before' is from an unknown root with no other
attestations in the language, but it must have been a substantive because its
bi-syllabic pattern is only found with substantives in this language (Fox 2003:
107). Prepositions in Semitic primarily originate from substantives; still, except
for very few examples, none of the Semitic subordinators is attested as a prepo-
sition before it is attested as a subordinator. In some cases, subordinators are
not attested as prepositions at all (e.g., Hebrew *ṭerem* 'before', *pen* 'lest'; Arabic
ḥayṯu 'where'; Ugaritic *aṯr* 'where'), but in other cases prepositions and subor-
dinators are attested concurrently, none of which are obviously original (e.g.,
Akkadian *ēma* 'wherever/in wherever'; Hebrew *ʔaḥar* 'after conj./prep.').

Prepositions which can be reconstructed to Proto-Semitic in this function
are never attested as subordinators (Gai 1985: 125). Hence the development is
likely not a case of noun > preposition > adverbial subordinator, but rather
noun > adverbial subordinator. Such a development is quite different from
what is reported for other languages. In the languages of Europe, for exam-
ple, the most common sources for adverbial subordinators are adverbs, ad-
positions, complementizers, interrogatives and relativizers (Kortmann 1998:
216). Heine & Kuteva (2000) list a number of possible sources for subordina-
tors cross-linguistically, which are mostly postpositions and locative markers,
though in some languages, demonstratives and verbs are possible too.

Furthermore, subordinators regularly have a morphological form which is
identical to nominal heads ('construct' nouns). Not all languages mark nomi-
nal heads in the same way synchronically: Arabic has no post-case nasalization
('nunation') on nominal heads (cnst *ḥin-a-* vs. abs *ḥin-an*), Akkadian has no
post-case nasalization ('mimation') or case ending (cnst *ašar-* vs. abs *ašr-u-m*)
and Ethiopic has a special vocalic ending (cnst *ʔəsm-a* vs. abs *səm*). The se-
mantics of the subordinators may or may not be distinct from their original
meaning. For example, Hebrew *ʕēqeb* 'because' is likely derived from a noun
meaning 'consequence' (cf. Arabic *ʕaqib-un* 'consequence'); Arabic *ḥīna* 'when'
is a construct of a noun meaning 'incidence, duration of time'.

Adverbial subordinated clauses attested in the Semitic languages are syntactically very similar indeed, to the degree that they could easily be considered cognates. They share the following structural features:

a. These sentences are typically introduced by substantives (and not by interrogatives or prepositional phrases)

b. In languages where case is still viable, these particles carry a relic of case, typically accusative, for example, in Arabic: *ḥīn-a, ḥatt-ā, rait̠-a* and earlier *yawm-a,* but other cases, or none, are also attested

c. The morphology of the subordinator, if still apparent, is usually that of a nominal head ('construct' noun). For example, these forms lack final post-case nasalization in Akkadian and Arabic, and carry a special morpheme marking nominal heads in Ethiopic

d. The noun-turned-subordination marker is not represented syntactically in the finite sentence following it, e.g., via an anaphoric pronoun

What may be reconstructed on the basis of the correspondences in all branches is a common pattern, which following the principles outlined in Harris & Campbell (1995) and in Harris (2008), should be reconstructed as: N[case.cnst.] S. However, there are several problems with assuming that this is the likely reconstruction of adverbial subordination in Semitic: first, it is clearly a specific case of the reconstruction suggested above for nominally-headed relatives (N[case.cnst] S). Second, the value of N in this reconstruction is unreconstructable as there is no overlap between the sets of adverbial subordinators in the various languages, not even in closely related languages (Pat-El 2008: 61). Note the following selection for the variety of these forms in a number of languages. With the possible exception of particles built on *kī-* (Akkadian *kīma,* OSA *k,* Hebrew *kī*), none of the other markers listed here is etymologically related to other markers in the same column:

TABLE 8.3   A sample of subordinators in the Semitic languages

	before	Because	so that	When
Akkadian	*adi ... lā* (*adi*=until)	*kīma, aššum*	*kīma, aššum*	*kīma, ašar* (rare)
Arabic	*qabla, ʔamāma*	*liʔanna, bi-sabab*	*likay, ḥattā, biḥaytu*	*ḥīna*
Hebrew	*ṭerem*	*ʕeqeb, kī, yaʕan*	*ləmaʕan* neg. *pen*	*kī*
Ethiopic	*ʔəmqədma, ʔənbala*	*ʔəsma*	*kama* neg. *ʔənbala*	*soba*
OSA		*k, ywm*	*brt, k*	*lhm, ln, ywm, brt̠*

Almost no adverbial subordinator can be reconstructed as such further than the language in which it is attested. With no overlap in etymological content both an inheritance and a mutual borrowing are unlikely. Since there are no shared adverbial subordinators, even if the pattern appears to be shared, we cannot conclude that any of the attested subordinators in Proto-Semitic is original; in fact, they all seem to reflect a fairly shallow history. Since the reconstructed pattern is for all intents and purposes the nominally-headed relative, the adverbial function of the pattern itself cannot be said to be Proto-Semitic. The sole difference between nominally-headed relatives and adverbial clauses is that the head of nominally-headed relatives inflects according to its function in the main clause, while the adverbial subordinating marker does not inflect at all, or shows a fossilized inflection.

It is, therefore, more likely that adverbial subordination is not reconstructable to Proto-Semitic, but rather is a result of reanalysis of nominally-headed relatives which took place independently in each branch. The reanalysis of a nominally-headed relative was facilitated by ambiguity. As was mentioned above, in this type of relative, there are typically no anaphoric markers which represent the head noun in the relative clause; see example (12), where the antecedent of a marker-headed relative (*he-ʕārîm* **fp**) is represented in the relative via an anaphoric pronoun (*-hen* **3fp**). Such anaphoric checks are common in marker-headed relatives, and make the role of the head noun in the relative clause explicit. When the anaphoric pronoun is missing, the head noun is less likely to be understood as a head of a relative clause, because it is not represented in the relative clause (13):

(12)   ***he-ʕārîm***   *ʔăšer yāšab*   ***bā-hen***   *Lôṭ*
      DEF-cities.FP   REL   reside.PF.3MS   in-them.3FP   PN
      Nⁱ   PREP-PRONⁱ
      'The cities in which Lot resided.' (Biblical Hebrew, Genesis 19: 29)

(13)   *hôy ʔĂrîʔēl ʔĂrîʔēl qiryat*   *ḥānā*   *Dāwid* (**b-āh)
      Oh Ariel Ariel town.FS.CNST   camp.PF.3MS David (**in-her)
            Nⁱ   (**PREP-PRONⁱ)
      'Oh Ariel, Ariel, a city in which David encamped.'
                                                  (Biblical Hebrew, Isaiah 29: 1)

In addition, nouns in construct are proclitic and therefore unstressed, forming one intonational unit with their dependent, or the first element of their dependent if it is a sentence. As a result, in some languages, construct nouns, including head nouns in nominally-headed relatives, show vowel

shortening, sometimes syncope (Steiner 2012) and loss or fossilization of in-
flection (Huehnergard & Pat-El 2018). The lack of inflection, in particular case
inflection, on the head noun obscures the function of this noun in the main
clause. The lack of anaphoric resumption in the relative and the reduction or
loss of inflection on nominal heads of nominally-headed relatives leads to an
ambiguity regarding the relationship between the head and the subordinated
clause. Subsequently, these nominal heads were understood not as nouns with
a following relative clause, but rather as mere markers of subordination, fre-
quently semantically reduced; head nouns with specific semantics, e.g., tem-
poral, spatial or causal, were particularly susceptible to such reanalysis. The
example in (13) would then be understood as "Oh Ariel, Ariel, where David
encamped". In this case, the Hebrew noun *qiryā* 'city' did not become a sub-
ordinator, but in Akkadian, for example, nouns denoting 'place' (*ašar*) and
'house' (*bēt*), did. In example (14) below the noun *bēt* 'house' is semantically
reduced and simply means 'where, in which'. Even though *bēt* is marked as a
head of the relative, that is, it is in construct with no case and no boundedness
markers, it is synchronically not the head of a relative. Its sole function is mark-
ing a spatial adverbial clause. Since the nouns *bēt* and *ašar* are not attested as
adverbs in Akkadian, it is unlikely that the development is NOUN > ADVERB
> ADVERBIAL SUBORDINATOR (contra Givón 1974).

(14) *mātu   bēt   šarru   bēl-ī   iškun-īni ...*
     land   where   king   lord-my   place.PRET.3MS-me
    šarru   bēl-ī   ūdu
    king   lord-my   know.STAT.3MS
    'The king my lord knows the land where the king my lord placed me'
                    (Akkadian, Neo-Assyrian, SAA 16 127 013–14)

The role of head noun inflection, or lack thereof, in creating adverbial subor-
dinator particles is substantiated further in languages where the relative lost
its inflection. In these languages, the no-longer-inflected marker can also func-
tion as an adverbial subordinator.

    Most of the languages which allow relative markers to be used as adverbial
subordinators either lost nominally-headed relative as an option (Aramaic) or
mutated it (Arabic; Pat-El 2014). In example (15), the relative marker is a fossil-
ized noun *ʔăšer*, cognate with Akkadian *ašar* 'place / where', and can function,
albeit rarely, as an adverbial subordinator. This marker probably developed
from adverbial subordination, not from an adverb, as there is no evidence
that it was ever used as an adverb (Huehnergard 2006). Example (16), of a Late
Antique dialect of Hebrew, reflects a further reduction of this marker, resulting

in a proclitic *šeC-*, similarly with causal meaning. The Aramaic relative marker in example (17) has lost its inflection and prosodic independence fairly early, and can be used in this dialect and others as an adverbial subordinator. The relative marker in example (18) reflects an innovation within the historical record of Arabic, where the original relative marker, which was inflected for gender-number-case in some early dialects, was replaced by a combination of definite article and demonstrative, which subsequently lost its inflection (Stokes 2018).

(15) *nātan*   *ʔĕlōhîm śəkār-î*   *ʔăšer nātattî*   *šipḥāt-î*   *lə-ʔîš-î*
      give.PF.3MP god   wage-my REL   give.PF.1S servant.FS-my to-man-my
      'God paid me my due, because I gave my maidservant to my husband'
                                                            (Biblical Hebrew, Genesis 30: 18)

(16) *kol*   *ʔādām qôrē*   *kə-dark-ô*   *šen-neʔĕmar ...*
      every   man   read.PTCL.MS as-way-his   REL-say.PF.PASS.3MS
      'Every person reads according to his custom, because it is said [in the bible] ...' (Mishnaic Hebrew, Berakhot 1: 3)

(17) *hī*   *lābšā*   *mānīn ṣāʔīn d-lā*   *yittēn*   *barnāš*
      she   wear.PF.3FS   clothes dirty REL-NEG   give.IMPF.3MS   person
      ʕēnā-wy ʕəl-āh
      eyes-his on-her
      'She wears dirty clothes so that no-one will look at her'
                                                            (Jewish Palestinian Aramaic)

(18) *il-ḥaqq*   *ʕalayy ʔilli mā rabbet-ak*   *zayy n-nās*
      DEF-blame   on-me REL NEG teach.PF.1CS-you like   DEF-man
      'I am to blame because I didn't raise you properly' (Palestinian Arabic)[5]

The reanalysis of heads in nominally-headed relatives as adverbial subordinators was enabled by reduction of inflection, itself a phonological process motivated by the proclitic position of these heads (for more on the phenomenon of relative marker > adverbial subordinator in Arabic, see Diem 2007).

When languages stop using the nominally-headed relative productively or when such relatives are no longer transparent to speakers, the method of creating adverbial subordinating markers changes as well. Aramaic lost the

---

5 In other modern dialects of Arabic, some adverbial clauses, particularly temporal ones, include a head noun, typically indefinite, followed by a relative particle but missing a resumptive pronoun. Often the head nouns of these clauses are in construct (Brustad 2000:102–103).

ability to form nominally-headed relatives fairly early. In addition, its relative marker lost inflection and remained as an uninflected relative marker. In this branch, adverbial subordination is marked by both a marker derived from a noun or a prepositional phrase but also obligatorily followed by the relative marker (Pat-El 2008); e.g., *bātār dī* < *b-ʔātār dī* (in-place rel) 'after'.

3   Discussion

The repeated innovation of adverbial markers in Semitic can hypothetically have a number of explanations: (i) a result of contact with transference of matter; (ii) a result of contact with transference of pattern; (iii) a reflection of common inheritance followed by subsequent lexical replacement; or (iv) parallel development. An explanation based on contact is problematic for a number of reasons. While the syntax of the adverbial clause is similar across Semitic, there is no shared set of subordinators even between languages which were in prolonged and early contact (e.g., Akkadian and Aramaic). The lack of evidence makes tracing a common locus for the spread of this pattern impossible. Of course, it is possible that contact drove structural changes without transfer of cognate forms (e.g., Epps 2006, Matras & Sakel 2007); however, in this case, since the basic pattern is reconstructable to the proto-language, contact is an unlikely explanation. The development happened in every branch and its results are apparent in all attested languages regardless of the level of contact, or lack thereof, between them.

Lexical replacement could account for the different subordinators populating a similar structure. According to this scenario, speakers could identify links between subordinators and nouns with similar semantic value, and were able to exchange these items as part of a lexical renewal. The problem with such a hypothesis is that studies on the rate of lexical replacement suggest that even with lexical replacement, some cognate material, should be retrievable. Pagel et al. (2007) show that high-frequency words are replaced at a lower rate than low-frequency words. Their study suggests that the linguistic half-life of a word (the time in which there is 50% chance of replacement) varies between 750–10,000+ years depending on frequency rates. This study concentrated on nouns, but function words are more frequent in speech than common nouns, and are therefore expected to be replaced less frequently. We would, therefore, expect to find some cognates between languages whose split occurred relatively recently, like the Aramao-Canaanite subbranch; the fact that we do not find such material suggests that the explanation is likely not lexical replacement.

Independent development has been shown to account for similar structures with no cognate material in other language families. LaPolla (1994) reviews a number of cases in the Tibeto-Burman family ('anti-ergative' marking, ergative marking, direction marking, causative marking, person marking, and existential verbs), which not only show similar developmental paths, but also sometimes use similar morphological material in the evolution of these new categories. In each case even closely related languages show a development of the same category from different morphemes. While a reconstruction to a shared ancestor is unlikely, LaPolla concluded that independent parallel development reveals the motivation for the development, namely, the starting point, which can be traced to the proto language. LaPolla suggested that a set of semantic distinctions (agentivity/non-agentivity, animacy, and others) were essential to the organization of the protolanguage speakers' world view. In the current article, I suggested that the motivation may be even more concrete, as a set of fairly common and simple changes cause ambiguity in the interpretation of a shared pattern and lead to reanalysis.

A view on parallel development from a different angle is found in a number of publications by N. J. Enfield who developed the idea of 'typological poise', which he uses to explain how the existing grammatical structure encourages (or constraints) grammatical development in a certain direction (e.g., Enfield 2001, 2003). When two languages show similar structural compatibility to undergo a change in a specific direction, they are said to be 'poised'; such compatibility may be a result of a number of factors, such as areal or genetic relatedness. The term should be used as a measure for the likelihood that languages will evolve in a certain direction independently. Thus, if a language is 'poised' for certain grammatical developments, whether in its existing grammatical structure or semantic mapping, it will be more likely to realize certain changes. In the case of closely related languages with shared grammatical structure, a similar poise may be found in many or all members of a family. The evidence and analysis presented in this article support the hypothesis that language poise is a good predictor for independent development. The Semitic languages show a development which is relatively rare cross-linguistically (Heine & Kuteva 2000), but is a likely outcome internally, given these languages' poise.

Parallel development may create structures that could be erroneously attributed to inheritance, as they show distinct similarity and thus may be taken to be cognate patterns. However, such structures, while creating uniformity, also exhibit diversification. Therefore, unlike genetically shared features, which reflect real correspondence, parallel development is responsible for patterns that cannot be reconstructed to a common ancestor, but it can point to an underlying shared pressure, or stimulus of change. For the most part, parallel

development is assumed to be irrelevant for subgrouping (e.g., Nakhleh et al. 2005: 384); however, I suggest, similarly to LaPolla (1994), that there is value in studying the underlying stimuli which motivate parallel development. Although a chance factor cannot be dismissed, a detailed and specific repeated path of change with identified motivation in more than two languages within a genetically cohesive language branch is highly unlikely to be a random event (see also Keiser 2009: 12). Identifying the structural features likely to participate in a change ('poise') is needed for the evaluation of what types of changes are plausible (though not predictable). Although this was not directly addressed in the current article, reconstructing the underlying structure which stimulated the change, in this case, a type of relative clause, is possible based on instances of multiple instances of parallel development.

It is quite likely that instances of parallel development are more common than is currently documented in the linguistic literature. There are obvious complications in identifying such a situation, as in languages with a shallow historical record, it is almost impossible to distinguish parallel development from other causes of similarity. At this point I would like to cautiously suggest the following:

(1) languages with similar structures are susceptible to similar changes
(2) the results of parallel development cannot be used for subgrouping, but they expose stimuli of change, which given (1) are likely to be repeated (contra Nakhleh et al. 2005)
(3) understanding the motivation behind parallel development may help formulate better constraints on change within individual language
(4) the results of parallel development-induced change may reflect shared structural features not otherwise easily reconstructable

4   Summary

Given the inability to reconstruct a set of shared adverbial subordinating markers in Semitic, I have argued here that their development is likely the result of parallel development on the basis of nominally-headed relatives. The nominal head in such relatives is not typically represented in the relative clause via agreement markers, and so its syntactic function in the relative clause remains unspecified. In many languages, the case on such nouns was reduced due to their proclitic status. I suggested above that nominal heads in such relatives, especially heads with spatial, temporal and causal semantics, were susceptible to be reanalyzed as adverbial subordinators. The Semitic languages repeatedly

resorted to this common solution for the same reasons, using similar but not identical linguistic material. Such solutions are likely to be repeated as long as the relevant features are still part of the grammar.

I have argued based on the evidence presented here that syntactic reconstruction based on cognate patterns may conflate genuine inherited syntactic material with cases of parallel development, where a change has taken place independently in more than one node. I suggest that parallel development may be triggered by a shared structural feature, which created similar pressures in distinct nodes and allowed for identical lines of development to take place. While the incentive for the development may be reconstructed to the ancestral node, the result of the development cannot. The development of adverbial subordinators, as outlined here, shows that despite structural similarities in adverbial subordination among the Semitic languages, it is unlikely that this pattern is reconstructable to the proto language. Rather, I have suggested that the stimulus, a type of relative clause, is responsible for the development independently in each node. The case outlined here also shows that parallel development can occur multiple times in situations in which the cause for the development is still part of the synchronic grammar. The arguments in this article are not intended to deny that syntax can be reconstructed on the basis of cognate pattern, but rather to supply evidence that this is not always the correct analysis.

References

Barðdal, J. & Th. Smitherman. 2013. The Quest for Cognates: A Reconstruction of Oblique Subject Constructions in Proto-Indo-European. *Language Dynamics and Change* 3(1) 28–67.

Brustad, K. 2000. *The Syntax of Spoken Arabic*. Washington, D.C.: Georgetown University Press.

Campbell, L. & A. C. Harris. 2002. Syntactic Reconstruction and Demythologizing 'Myths and the Prehistory of Grammars'. *Journal of Linguistics* 38(3): 599–618.

Crowley, T. 1991. Parallel Development and Shared Innovation: Some Developments in Central Vanuatu Inflectional Morphology. *Oceanic Linguistics* 30(2): 179–222.

Deutscher, G. 2009. Nominalization and the Origin of Subordination. *Syntactic Complexity: Diachrony, Acquisition, Neuro-Cognition, Evolution*, ed. by T. Givón & M. Shibatani, 199–214. Amsterdam/Philadelphia, John Benjamins.

Diem, W. 2007. Arabic *alladī* as a Conjunction: An Old Problem and a New Approach. *Approaches to Arabic Linguistics*. ed. by E. Ditters & H. Motzki, 67–112. Leiden: Brill.

Dunn, M., S. C. Levinson, E. Lindström, G. Reesink & A. Terrill. 2008. Structural Phylogeny in Historical Linguistics: Methodological Explorations Applied in Island Melanesia. *Language* 84(4): 710–759.

Enfield, N. J. 2001. On Genetic and Areal Linguistics in Mainland South-East Asia: Parallel Polyfunctionality of 'Acquire'. *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*, ed. by A. Aikhenvald & R. M. W. Dixon, 255–290. Oxford: Oxford University Press.

Enfield, N. J. 2003. *Linguistic Epidemiology*. New York: Routledge Curzon.

Epps. P. 2006. Growing a Numeral System: The Historical Development of Numerals in an Amazonian Language Family. *Diachronica* 23(2): 259–288.

Fox, J. 2003. *Semitic Noun Patterns*. Winona Lake: Eisenbrauns.

Gai, A. 1985. Prepositions-Conjunctions in Akkadian and Ge'ez. *Welt des Orients* 16: 123–25.

Givón, T. 1974. Verb Complements and Relative Clauses: A Diachronic Case study in Biblical Hebrew. *Afroasiatic Linguistics* 1(4): 1–22.

Goldenberg, G. 1995. Attribution in the Semitic Languages. *Langues Orientales Anciennes: Philologie et Linguistique* 5(6): 1–20.

Harris, A. C. 2008. Reconstruction in Syntax: Reconstructions of Patterns. *Principles of Syntactic Reconstruction*, ed. by G. Ferraresi & M. Goldbach, 73–95. Amsterdam/Philadelphia: John Benjamins.

Harris, A. C. & L. Campbell. 1995. *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: Cambridge University Press.

Heine, B. & T. Kuteva. 2000. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

Hetzron, R. 1976. Two Principles of Genetic Reconstruction. *Lingua* 38(2): 89–108.

Huehnergard, J. 2006. On the Etymology of the Hebrew Relative *še-*. *Biblical Hebrew in Its Northwest Semitic Setting: Typological and Historical Perspectives*, ed. by A. Hurvitz & S. Fassberg, 103–126. Jerusalem.

Huehnergard, J. & N. Pat-El. 2018. The Origin of the Semitic Relative Marker. *Bulletin of the School of Oriental and African Studies* 81(2): 191–204.

Huehnergard, J. & N. Pat-El. 2019. Introduction to the Semitic Languages and their History. *The Semitic Languages*. ed. by J. Huehnergard & N. Pat-El, 1–21. Abingdon: Routledge.

Keiser, S. H. 2009. When 'Speech Islands' Aren't Islands: Parallel Independent Development, Drift and Minimal Levels of Contact for Diffusion. *Diachronica* 26(1): 1–35.

Khan, G. 1988. *Studies in Semitic Syntax*. Oxford: Oxford University Press.

Kortmann, B. 1998. The Evolution of Adverbial Subordinators in Europe. *Historical Linguistics 1997*. ed. by M. S. Schmidt, J. R. Austin & D. Stein, 213–228. Amsterdam/Philadelphia: John Benjamins.

LaPolla, R. J. 1994. Parallel Grammaticalizations in Tibeto-Burman Languages: Evidence of Sapir's 'Drift'. *Linguistics of the Tibeto-Burman Area* 17(1): 61–80.

Lass, R. 1997. *Historical Linguistics and Language Change*. Cambridge: Cambridge University Press.

Lightfoot, D. 2002. Myths and the Prehistory of Grammars. *Journal of Linguistics* 38(1): 113–136.

Malkiel, Y. 1981. Drift, Slope, and Slant: Background of, and Variations upon, a Sapirian Theme. *Language* 57(3): 535–570.

Matras, Y. & Sakel, J. 2007. Investigating the Mechanisms of Pattern Replication in Language Convergence. *Studies in Language* 31(4): 829–865.

Meillet, A. 1918. Convergence des développements linguistiques. *Revue philosophique* 85: 61–75.

Minett, J. W. & W. S.-Y. Wang. 2003. On Detecting Borrowing: Distance-Based and Character-Based Approaches. *Diachronica* 20(2): 289–330.

Nakhleh, L., D. Ringe & T. Warnow. 2005. Perfect Phylogenetic Networks: A New Methodology for Reconstructing the Evolutionary History of Natural Languages. *Language* 81(2): 382–420.

Nichols, J. 1986. Head-Marking and Dependent-Marking Grammar. *Language* 62(1): 56–119.

Pagel, M., Q. D. Atkinson & A. Meade. 2007. Frequency of Word-Use Predicts Rates of Lexical Evolution Throughout Indo-European History. *Nature* 449/7163: 717–720.

Pat-El, N. 2008. Historical Syntax of Aramaic: A Note on Subordination. *Aramaic in its Historical and Linguistic Setting*, ed. by H. Gzella & M. L. Folmer, 55–76. Wiesbaden: Harrassowitz.

Pat-El, N. 2012. The Syntax of *ʾăšer* and *šeC* Yet Again. *Language and Nature: Papers Presented to John Huehnergard on the Occasion of his 60th Birthday*, ed. by R. Hasselbach & N. Pat-El, 319–327. Chicago: The Oriental Institute.

Pat-El, N. 2014. The Morphosyntax of Nominal Antecedents in Semitic, and an Innovation in Arabic. *Proceedings of the Oslo-Austin Workshop in Semitic Linguistics*, ed. by L. Edzard & J. Huehnergard, 28–47. Wiesbaden: Harrassowitz.

Rieseberg, L. H. & J. M. Burke. 2001. The Biological Reality of Species: Gene Flow, Selection, and Collective Evolution. *Taxon* 50: 47–67.

Steiner, R. C. 2012. Vowel Syncope and Syllable Repair Processes in Proto-Semitic Construct Forms: A New Reconstruction Based on the Law of Diminishing Conditioning. *Language and Nature: Papers Presented to John Huehnergard on the Occasion of his 60th Birthday*, ed. by R. Hasselbach & N. Pat-El, 365–390. Chicago: The Oriental Institute.

Stokes, Ph. W. 2018. The Plural Demonstratives and Relatives Based on *ʾVl in Arabic and the Origin of Dialectal *illī*. *Re-engaging Comparative Semitic and Arabic Studies*, ed. by D. Birnstiel & N. Pat-El, 127–150. Wiesbaden: Harrassowitz.

Walkden, G. 2014. *Syntactic Reconstruction and Proto-Germanic*. Oxford: Oxford University Press.

Warnow, T., S. N. Evans, D. Ringe & L. Nakhleh. 2006. A Stochastic Model of Language Evolution that Incorporates Homoplasy and Borrowing. *Phylogenetic Methods and the Prehistory of Languages*, ed. by P. Forster & C. Renfrew, 75–87. Cambridge: McDonald Institute for Archaeological Research.

Wolfram, W. & N. Schilling-Estes. 2003. Parallel Development and Alternative Restructuring: the Case of *Weren't* Intensification. *Social Dialectology: In Honor of Peter Trudgill*, ed. by D. Britain & J. Cheshire, 131–153. Amsterdam/Philadelphia: John Benjamins.

Wood, T. E., J. M. Burke & L. H. Rieseberg. 2005. Parallel Genotypic Adaptation: When Evolution Repeats Itself. *Genetica* 123: 157–170.

Reconstructing Semantic Roles: Proto-Indo-European *-bʰi

Eugenio R. Luján and Ángel López Chala

Abstract

Traditional approaches to the reconstruction of Proto-Indo-European nominal morphosyntax have operated by first reconstructing the set of morphological cases for every declensional type, and then attempting to establish the meaning of the different cases, regardless of the specific ending that was used for each declensional type. However, more insight can be gained into the reconstruction of the nominal syntax of proto-languages by applying the concepts and methodologies developed in recent years in functional-typological approaches to language study. Under this approach, the aim of syntactic reconstruction in the nominal domain lies not in determining the meaning of a given case as a whole but rather in elucidating the semantic role(s) that a specific formative could be used for and, to the extent that this is possible, how those semantic roles relate to each other in historical terms. In this article we survey the semantic roles related to *-bʰi-endings in the old Indo-European languages. In the traditional reconstruction, *-bʰi has been considered the suffix expressing the Instrumental plural of the athematic declension. However, in the various branches of the family in which it is attested, *-bʰi-endings express a broad array of semantic roles. When charted on a diachronic semantic map of Instrument and related semantic roles, the *-bʰi-endings appear to cover neighbouring areas, and it becomes clear that they have followed well-known paths of semantic change. If we add the information about *-bʰi in the pronominal declension and its etymology, a neat grammaticalisation process is revealed. This results in a 'dynamic' reconstruction of the morphosyntax of the proto-language, which is more in accord with what we know about the actual processes of semantic change in grammatical markers and paradigmatisation of markers more generally.

1   Introduction: Reconstructing Semantic Roles

In this article we explore a new type of approach to the reconstruction of Proto-Indo-European (PIE) case syntax.[1] We apply the concepts and methodologies developed in recent years in functional-typological approaches to language to the reconstruction of the nominal syntax of a proto-language. Therefore, instead of trying to reconstruct the possible meanings of a given syntactic case in PIE by comparing the meanings associated with that syntactic case in the Old Indo-European languages, we will focus on a specific formative and will analyse in depth which semantic roles are expressed by means of that form in the various daughter languages.

The study of semantic roles already has a long tradition in linguistics, and semantic roles have been the object of intensive research in syntax in the last years (see Luraghi & Narrog 2014: 1–12, with further references). However, to the best of our knowledge, whether an approach based on the analysis of the semantic roles expressed by a specific form can help to refine syntactic reconstruction has not yet been explored. From this perspective, the aim of syntactic reconstruction in the domain of nominal syntax would not be to determine the meaning(s) of a case as a whole in a proto-language, that is, the meaning(s) and use(s) of the Accusative, the Dative, and so on. The goal would be, instead, to elucidate the semantic role(s) that a particular case ending could be used for, and, to the extent that this is possible, establish how those semantic roles relate to each other in historical terms, and what semantic changes can account for the different meanings across the languages of the family.

In contrast to more traditional approaches, some previous studies have already focused on the processes by which certain endings came to be used, aiming to recover their original functions and analysing how they entered the nominal declension. This was, for example, the case with Villar's (1981) approach to the reconstruction of the Dative and Locative endings *-ei/-i in Proto-Indo-European, elaborating on previous ideas that go back to W. D. Whitney and K. Brugmann. However, those studies were not framed in the same methodological perspective that we will be applying now.

In this article, we will take the PIE 'Instrumental' ending *-$b^h i$ as a case study. After reviewing the traditional approaches to the syntax of the Instrumental

---

case in PIE and pointing out some of their flaws, we will carry out a thorough revision of the data about *-*bʰi*-endings in the ancient Indo-European languages, identifying which semantic roles they can express in the languages in which they are attested. This will be the empirical basis for an integrated account of the data in terms of reconstruction and historical development.

Relying on the scholarly literature on the grammaticalisation of Instruments and related cases, we will try to understand which semantic changes must have taken place before the form came to express the variety of semantic roles that it is used for in the Old Indo-European languages. A helpful tool for that analysis will be semantic maps, which can be used to represent the linguistic relationship between forms and meanings or functions, as well as their closeness and similarity. We will chart our findings on the semantic maps proposed so far for the domain of the Instrument and related semantic roles to see whether a consistent pattern emerges for the data provided by the Old Indo-European languages. Finally, based on our findings, we will argue in favour of an etymology for the suffix *-*bʰi*.

The structure of the article is as follows: we first discuss certain problems in traditional approaches to the reconstruction of case syntax in Proto-Indo-European (Section 2) and review the proposals for the reconstruction of a PIE Instrumental case ending *-*bʰi* (Section 3). After this, we provide the results of a thorough analysis of the semantic roles associated with *-*bʰi*-endings in the various branches of the Indo-European family in which they occur in the nominal inflection: Old Indic (Section 3.1.1), Iranian (Section 3.1.2), Armenian (Section 3.1.3), Greek (Section 3.1.4), and Celtic (Section 3.1.5). We then bring together the data of all these branches (Section 3.2) and add the information about the use of *-bʰi-forms in the pronominal declension, which are significantly different (Section 3.3). On this basis we attempt a reconstruction of the original meaning of the PIE suffix *-*bʰi* and the paths that it must have followed before it came to be associated with the various semantic roles that it expressed in the Old Indo-European languages (Section 3.4). In the light of our research, we review the etymologies proposed so far for this marker and reconstruct the grammaticalisation process that it must have undergone (Section 3.5). The final part provides a summary of our findings and conclusions (Section 4).

2　　Problems with Traditional Approaches to the Reconstruction of Case Syntax

Traditional approaches to the reconstruction of Proto-Indo-European nominal morphosyntax have operated by first reconstructing the set of morphological cases for every declensional type and then by trying to establish the meaning

of the different cases, regardless of the specific form that was used for each declensional type (e.g. Fritz 2003: 261–274).

An Instrumental case has traditionally been reconstructed for all declensional types in Proto-Indo-European. According to Fortson (2010: 115–129), the endings would be as shown in Tables 9.1 and 9.2 for the athematic and thematic declensions.

Nevertheless, we find differences in the details of the reconstructed suffixes in the vast academic literature on this topic. We cannot review this literature here in full, so we will mention just a few examples of the different reconstructions proposed for the Instrumental and other oblique cases. We will not go into the details of the arguments in favour of reconstructing all of them, since we will argue below that a different kind of approach is needed.

TABLE 9.1   Reconstruction of the endings of the athematic declension in PIE according to Fortson (2010: 115–118)

	Singular	Plural
NOMINATIVE	*-s*	*-es*
ACCUSATIVE	*-m*	*-ns*
GENITIVE	*-(é)s*	*-ōm ?*
DATIVE	*-ei*	*-bh(i̯)os*
INSTRUMENTAL	*-(e)h₁*	*-bhi(-)*
LOCATIVE	*-i*	*-su*

TABLE 9.2   Reconstruction of the endings of the thematic declension in PIE according to Fortson (2010: 126–129)

	Singular	Plural
NOMINATIVE	*-os*	*-ōs*
VOCATIVE	*-e*	*-ōs*
ACCUSATIVE	*-om*	*-ons*
GENITIVE	*-os ?*	*-ōm*
ABLATIVE	*-ōt*	*-o(i)bh-*
DATIVE	*-ōi*	*-o(i)bh-*
INSTRUMENTAL	*-ō*	*-ōis*
LOCATIVE	*-oi*	*-oisu*

TABLE 9.3   Reconstruction of Instrumental and related endings according to different scholars

	Instr. sg.	Instr. pl	Dat. pl.	Abl. pl.
Szemerényi (1996: 155–192)	*-e/-o *-bhi/-mi	*-ōis *-bhis/-mis	*-bh(y)os/-mos	*-bh(y)os/-mos
Mallory – Adams (2006: 57–58)		*-ōis *-bhi	*-oibh(y)os/-omus *-mus	*-om *-bh(y)os
Beekes (2011: 186)		*-ōis *-bhi,	*-omus *-mus	*-ios

Szemerényi (1996: 155–192) reconstructs both *-e/-o* and *-bʰi/-mi* for the Instrumental singular, *-bʰis/-mis* and *-ōis* for the Instrumental plural and *-bʰ(y)os/-mos* for the Dative-Ablative plural. Like Fortson, Mallory and Adams (2006: 57–58) reconstruct Instr. Pl. *-bʰi*, Abl. Pl. *-bʰ(y)os* and Dat. Pl. *-mus* for the athematic declension, while for the thematic declension they propose Dat. Pl. *-oibʰ(y)os/-omus* and Inst. Pl. *-ōis*, but Abl. Pl. *-om* (same ending as the Gen. Pl.). Beekes (2011: 186) proposes Instr. Pl. *-bʰi*, Dat. Pl. *-mus* and Abl. Pl. *-ios* for the athematic declension, and Instr. Pl. *-ōis* and Dat. Pl. *-omus* for the thematic declension. On the basis of Indo-Iranian (Skt. Instr.-Dat.-Abl. *-bhyām*, Av. *-bya*, etc.), some scholars also reconstruct a PIE Instr. Du. *-bʰih₁* (see, e.g., Mallory and Adams 2006: 57). Some of the proposed reconstructions are summarised in Table 9.3.

In summary, relying on the evidence provided by the oblique endings of some ancient Indo-European languages, it could at most be argued that a *-bʰi*-ending must have been used in Proto-Indo-European, but what its precise array of meanings and functions was is far from clear.

Fritz (2003: 268) has summarised the values of the Instrumental case in Proto-Indo-European as follows:
– Instrumental of accompaniment
– Instrumental of means
– Instrumental of route
– Instrumental of constitution
– Instrumental of accompanying circumstances
– Instrumental of reason
– Instrumental of comparison

In traditional approaches, it has been customary to quote sentences from different ancient Indo-European languages when reconstructing the values of the Proto-Indo-European Instrumental, regardless of what the actual forms

are and what place they occupy inside the whole declensional system of the language. For instance, under the heading "instrumental of accompaniment" Fritz (2003: 268) provides the following examples:

LATIN:

(1) Postquam utrimque exitum est maxima copia (Plaut.*Amph*.219)
    after on.both.sides gone is great.ABL.SG quantity.ABL.SG
    'after they marched up **in great numbers** on both sides'

GREEK:

(2) enthá d' hikáneis nēí te kaì hetároisi (*Od*.11.160–161)
    here but arrive.2SG ship.DAT.SG and companion.DAT.PL
    'you arrive here **with the ship and the companions**'

VEDIC:

(3) víśvair ū́mebhir ā́ gahi (*RV* 5.51.1)
    all.INST.PL helpers.INSTR.PL here come.2SG.IMPV
    'come here **with all the helpers**'

The problem with this type of approach is that the actual position of the Instrumental case is radically different in the three languages. As can be seen in Table 9.4, in Latin the Instrumental and the Ablative have merged into the Ablative case and, as shown in Table 9.5, in Greek the only grammaticalised oblique case is Dative, while in Vedic (Table 9.6) we have a proper Instrumental case. In our view, it is too risky to assume that the Ablative in Latin and the Dative in Greek provide evidence at the same level for the reconstruction of the meaning and function of the Proto-Indo-European Instrumental as Vedic or other Indo-European languages that have a proper, devoted Instrumental case, which, if it did exist in PIE, must have had a different form from those found in Latin and Greek.

TABLE 9.4   Latin declension (athematic)

	Singular	**Plural**
NOMINATIVE	*consul*	*consules*
VOCATIVE	*consul*	*consules*
ACCUSATIVE	*consulem*	*consules*
GENITIVE	*consulis*	*consulum*
DATIVE	*consuli*	*consulibus*
ABLATIVE	*consule*	*consulibus*

TABLE 9.5   Greek declension (athematic)

	Singular	Plural
NOMINATIVE	*phúlaks*	*phúlakes*
VOCATIVE	*phúlaks*	*phúlakes*
ACCUSATIVE	*phúlaka*	*phúlakas*
GENITIVE	*phúlakos*	*phulákōn*
DATIVE	*phúlaki*	*phúlaksi*

TABLE 9.6   Vedic declension (athematic)

	Singular	Plural
NOMINATIVE	*marút*	*marútas*
VOCATIVE	*marút*	*marútas*
ACCUSATIVE	*marútam*	*marútas*
GENITIVE	*marútas*	*marútām*
ABLATIVE	*marútas*	*marúdbhyas*
DATIVE	*marúte*	*marúdbhyas*
INSTRUMENTAL	*marútā*	*marúdbhis*
LOCATIVE	*marúti*	*marútsu*

There are further problems concerning the reconstruction of an Instrumental case for PIE. To start with, it is doubtful whether a morphologically differentiated Instrumental case can be reconstructed for Proto-Indo-European. As Smitherman & Barðdal (2009: 268) have argued, relying on the evidence provided by the ancient Indo-European languages, it seems more reasonable to assume an "oblique, vaguely instrumental role to *-b^hi" without assigning it "a specific inflectional role in the Proto-Indo-European declension system".

Furthermore, the fact that the *-b^h- and *-*m*-endings of the oblique cases of the plural are absent from the Anatolian branch of the family strongly suggests that this is a late development in Proto-Indo-European (Fortson 2010: 118–119, among others). However, given that in Balto-Slavic and Germanic the oblique cases of the plural show endings in -*m*- instead of *-b^h-, it has been argued that in Proto-Indo-European the Dative-Ablative plural ending was *-*mos*, and the Instrumental plural ending was *-b^hi. The dialectal distribution of *-*m*- and *-b^h- endings would thus be the result of different processes of analogical

TABLE 9.7 Instrumental case markers in Sanskrit in the various declensions

	Athematic	Thematic	Feminine
SINGULAR	-ā	-ā/-ena	-ā/-ayā
PLURAL	-bhis	-ais/-ebhis	-ābhis
DUAL	-bhyām	-bhyām	-bhyām

levelling, resulting in *-b^hos/-b^hi(s) and *-mos/-mi (see Meier-Brügger 2003: 197 with further references). Finally, languages that have a proper Instrumental case display different endings for the Instrumental singular and plural in the various declensional types, as typified by Sanskrit (Table 9.7).

3   Case Study: Indo-European *-b^hi

As we have already noted, the ancient Indo-European languages can be classified into two groups according to the endings used for the formation of the oblique cases of the plural number (Instrumental, Dative, Ablative, and Locative): *-m-endings are found in Germanic and Balto-Slavic, while *-b^h-endings occur in Indo-Iranian, Armenian, Greek, Italic (including Latin), and Celtic. There are both *-mo-/-mi- and *-b^ho-/-b^hi endings, and they show different values, as shown in Table 9.8.

As we have discussed above, in the traditional reconstruction, *-b^hi has usually been considered the marker of the Instrumental plural of the athematic declension. The final -s of *-$b^h$is has been variously interpreted: plural marker -s or adverbial -s as in the adverbs of various ancient Indo-European languages such as Lat. bis 'twice', Gk. pýks 'with the fist', etc. Jasanoff (2009:

TABLE 9.8   -mo-/-mi- and *-b^ho-/- b^hi-endings (adapted from Meier-Brügger 2003: 197)

	Latin	Greek	Indo-Iranian	Proto-Germanic	Balto-Slavic[a]
ABLATIVE	-bus		*-bhi̯as	*-m-	*-mos
DATIVE	-bus		*-bhi̯as	*-m-	*-mos
INSTRUMENTAL		-p^hi	*-bhis	*-m-	*-mi(s)

a   According to Olander (2014: 267–286), however, the Dative plural ending of Proto-Balto-Slavic must have been *-mas and the Instrumental plural ending (except for o-stems), *-mīs.

138–144) has proposed that it comes from *-*bʰi* with the addition of what he considers to be the oldest Instrumental marker in Indo-European, *-*is*. In any case, morphologically, the ending can only have been originally *-*bʰi*, as proved by the Greek evidence – a loss of the final -*s* cannot be accounted for in Old Greek, so the Indo-Iranian Instrumental ending -*bhis* must be an innovation of this branch of the family.

The evidence for reconstructing a *-*bʰi*-ending is the following:

– Greek (Mycenaean and Homeric) -*phi*
– Armenian: Instrumental singular -*w*/-*b*/-*v*/-∅ (< *-*bʰi*), Instrumental plural -*wkʻ*/-*bkʻ*/-*vkʻ*/-*kʻ* (with plural marker -*kʻ*) (see further explanations in Schmitt (1981: 92)
– Old Indic (Vedic and Sanskrit): Instrumental plural -*bhis*, Dative-Ablative plural -*bhyas*, Instrumental-Dative-Ablative dual -*bhyām*
– Iranian: Avestan: Instrumental plural -*bĭš*, Dative-Ablative plural -*biiō*, Instrumental-Dative-Ablative dual -*biiā* (there is also -*βe* and an isolated instance of -*biiąm* [case uncertain], see Hoffmann & Forssmann 1996: 115); Old Persian Instrumental plural -*biš*, Instrumental-Dative-Ablative dual -*biyā*; etc.
– Celtic: Gaulish -*bi*/-*be*, Old Irish Dative -(*a*)*ib*, etc.

In principle, the Dative-Ablative plural ending -*fs* of the Sabellic languages might also come from PIE *-*bʰis*, but the correspondences with other Italic and Indo-European Western languages rather point to *-*bʰos* (Meiser 1998: 128–129). Shields (1974: 281–286) argued that* -*bʰi* was to be found in Tocharian Genitives in -*pi* (Toch. B -*epi*, Toch. A -*āp*/-*yāp*); see also Pinault (1989: 89).

The uses of the *-*bʰi*-endings, however, differ from language to language. Thus, in Vedic we have Instr. Pl. -*bhis*, Dat.-Abl. Pl. -*bhyas* and Instr.-Dat.-Abl. Dual -*bhyām*, the three of them serving for the expression of a wider range of semantic roles than those referred to by the name of the case. In Mycenaean Greek -*pi* (phonetically [-pʰi]) is not only Instrumental, but also Locative and it is even used with prepositions; further possibilities are found in Homeric Greek. In Gaulish, -*bi* also occurs and it is used with a Sociative value. In the following sections we will undertake a thorough review of those values.

In addition, *-*bʰei*, the *e*-grade variant of the same form (*-*bʰi*) appears in the Dative case of the personal pronouns of some Indo-European languages. We will return to this later, after having discussed the meanings and functions of the *-*bʰi*-endings in Section 3.1 below.

The interpretation of the -*bhyo*(-*s*) endings of the Indo-Iranian languages is uncertain. They might be related to the oblique ending *-*bos* in the Western Indo-European languages (Latin, Italic, Celtic, Lusitanian …), if we could somehow assume that these have irregularly lost their intervocalic -*y*-, but this is a

rather unlikely scenario. Furthermore, even if the regular ending in Sanskrit is *-bhyas*, there are instances of *-bhya* in Vedic, showing thus that *-s* is a later addition, as the Iranian data also suggests (cf. Avestan *maibiiā*, *ahmaibiiā*, etc.).

3.1 Semantic Roles of *-bʰi in the Old Indo-European Languages

We restrict our review to the adpositionless *-bʰi-forms now, but in Table 9.10 we also include information on whether the use of adpositions is possible or not in a given language. This is important, since, as Luraghi (2001a: 39) has warned, we should avoid calling syncretism the fact that the same grammatical case is used with different values if those values depend on the adposition that it goes with.

It is also important to distinguish between lexical semantics and grammatical semantics, for as Luraghi (2001a: 41) remarks for the Dative in Classical Greek, the specific lexemes that occur in a plain case (i.e. with no adposition) can be unambiguously interpreted as expressing a given semantic role according to its lexical features. We will see that this is the case with some of the semantic roles associated with *-bʰi-forms in certain Indo-European languages, and this will also help us to explain the different semantic and syntactic evolution in the nominal and pronominal declensions.

3.1.1 Old Indic

Old Indic is especially interesting for the development and extension of the *-bʰi-endings, given that there is internal evidence of the spread of its use. In the Vedic language, the Instrumental plural of the thematic declension can be either *-ais* or *-ebʰis*. The former has clear parallels in other Indo-European languages and must go back to PIE *-ōis, while the latter is clearly an innovation in Old Indic due to analogy with the athematic declension and its ending *-bhis*.

As with other Instrumental endings, the values of *-bhis* in Vedic are the following:[2]

INSTRUMENT

(4) śatáṃ   cákṣāṇo   akṣábhiḥ (*RV* 1.128.3)
     hundred   see.PTCP.NOM.SG   eye.INSTR.PL
     '(the god) seeing with his hundred eyes'

---

[2] We have conducted our own research on Vedic texts, but we have also taken into account previous work by Wenzel (1879) and Haudry (1977).

INTERMEDIARY

(5) yáḥ ...   dabhrébhiḥ ...   háṃsi   ū́yasaḥ (*RV* 1.31.6)
     REL.NOM.SG   little.INSTR.PL   destroy.2SG   more.numerous.ACC.PL
     'You, who destroy the more numerous (enemies) by means of few
     (friends).'

COMITATIVE

(6) devó   devébhir   ā́   gamat (*RV* 1.1.5)
     god.NOM.SG   god.INSTR.PL   to   come.SUBJ.3SG
     'May the god come with the gods.'

As Luraghi (2001a: 44) remarks, this value of the Instrumental is not as com-
mon in Vedic and *sam* 'together (with)' usually accompanies Instrumentals
with this meaning. Significantly, it has been observed that the Comitative
value of the plain Instrumental case in Vedic is mostly limited to plural ref-
erents, specifically count nouns, which, as Luraghi (2001b) observes, makes it
close to the meaning 'among'.

CAUSAL

(7) mā́ tvā   rudra   cukrudhāma   námobhir (*RV* 2.33.4)
     NEG 2SG.ACC Rudra.VOC   irritate.1PL.SUBJ   reverence.INSTR.PL
     'May we not irritate you with our reverences, Rudra.'

AGENT (with passives)

(8) ŕ̥ṣibhir   matíbhiḥ ...   hitám (*RV* 9.68.7)
     poet.INSTR.PL   mind.INSTR.PL   arranged
     'Arranged ... by the poets [Agent] with their minds.'

PERLATIVE

(9) antárikṣe   pathíbhiḥ   pátantam (*RV* 10.87.6)
     air.LOC.SG   path.INSTR.PL   flying.ACC.SG
     'flying along the paths in the air'

TIME (DURATION)

(10) pūrvíbhiḥ   hí   dadāśimá   śarádbhiḥ (*RV* 1.86.6)
     many.INSTR.PL   indeed   honour.PERF.1PL   autumn.INSTR.PL
     'We have worshipped for many years.'

LOCATIVE

(11) dhā́rābhir ójasā ā índrasya pītáye
stream.INSTR.PL strength.INSTR.SG PTCL Indra.GEN drink.INF.DAT
viśa (*RV* 9.65.14)
enter.IMPV
'Enter them (the jars) with strength in the streams of water, so that Indra can drink.'

Interestingly, all the examples of the Locative value provided by Haudry (1977: 101–103) are found with *-bhis*, not with the other endings of the Instrumental plural.

3.1.2 Iranian

3.1.2.1 *Old Persian*

In Old Persian *-bʰi*-endings typically accompany the prepositions *hadā* and *hacā*, with a Comitative and an Ablative value. Nevertheless, a few preposition-less occurrences of the *-bʰi*-endings are found in the corpus.

INSTRUMENTAL/MEANS

(12) vašnā : auramazdāhā : tya-maiy :
grace.INSTR.SG Ahuramazda.GEN.SG REL.NOM.SG-PRON.GEN.SG
kartam : imaibiš : akunavam : ūvnaraibiš :
do.PTCP.NOM.SG DEM.INSTR.PL do.1SG.PERF quality.INSTR.PL
'By Ahuramazda's grace, what I did, I did with those qualities.' (DNb 48)

There are also some examples of Dual instrumental which are found in natural pairs of the human body:

(13) yāumainiš : ahmiy : utā : dastaibiyā : utā : pādaibiyā :
skilled.NOM.SG be.PRS.1SG and hand.INSTR.PL and foot.INSTR.PL
'I am skilled with the hands and with the feet', i.e. 'I have developed my skills with the hands and with the feet.' (DNb 41)

COMITATIVE/ACCOMPANIMENT

We have found one example of the Comitative in the Achaemenid inscriptions, *viϑbiš* 'with the houses', but it is not a prototypical one given that this is an inanimate entity. This kind of inanimate Comitative is usually called Accompaniment (e.g. Luraghi 2003: 28):

(14) adam :   niyaçārayam :   kārahyā :   abicariš :
     1SG.NOM restore.1SG.PERF people.DAT.SG farmstead.ACC.PL
     gaiϑām-cā :   māniyam-cā :   viϑbiš-cā :
     livestock.ACC.PL-and meanial.ACC.SG-and house.INSTR.PL-and
     tyādiš :   Gaumāta : hya :   maguš :   adīnā:
     REL.ACC.PL NOM.SG   REL.NOM.SG magus.NOM.SG spoil.PERF.3SG
     'I restored to the people the farmstead, the livestock, the menial and
     (together with) the houses, of which Gaumāta the magus had spoiled
     them.' (after Schmitt's 1991: 53 edition and translation) (DB I 65)

3.1.2.2   *Old Avestan*
The uses of *-bʰi*-endings in Old Avestan are similar to those found in Vedic. A
list of values of the Instrumental can be found in Reichelt ([1909]1978: 232–239).

INSTRUMENT/MEAN
In Old Avestan the only examples of *-bʰi*-endings with the role of Instrument
are found with body parts:

(15) tə̄m   ahmākāiš   azdəbīš-cā
     PRON.ACC.SG POSS.INSTR.PL bone.INSTR.PL-and
     uštānāiš-cā   yazamaidē
     life.force.INSTR.PL-and venerate.PRS.1PL
     'we venerate him with our bones and souls' (Y.37.3)

In contrast, numerous examples of non-prototypical Instrumental, that is
Instrumental of means, can be found in the texts analysed:

(16) vohū   ϑβā   manaŋhā   vohū
     divine.INSTR.SG PRON.ACC.SG thought.INSTR.SG divine.INSTR.SG
     ϑβā   ašā   vaŋhuiiā̊   ϑβā
     PRON.ACC.SG harmony.INSTR.SG divine.INSTR.SG PRON.ACC.SG
     cistōiš   šiiaoϑənāiš-cā   vacə̄bīš-cā
     divine.understanding.GEN.SG ceremony.INSTR.PL-and word.INSTR.
     pairijasāmaidē
     PL-and   serve.1PL.PRS
     'with the divine Thought, with the divine Harmony, through the ritual
     acts and words of the divine understanding, we serve you' (Y. 36.4)

COMITATIVE

(17) kōi   drəguuō.dəbīš,   xrūrāiš
      PRON.NOM.PL supporters.of.deception.INSTR.PL   bloodthirsty.INSTR.PL
      rāmąm   dā̊ṇtē
      peace.ACC.PL give.3PL.INJ
      'Who will settle peace with the bloodthirsty supporters of deception?'
                                                                  (Y. 48.11)

PERLATIVE

In Old Avestan there are no examples of nouns with *-bʰi-endings with the prototypical role of Perlative. However, we have found one instance of a metaphorical Perlative:

(18) kauuā   vīštāspō   [...] nąsaṯ   vaŋhə̄uš
      sage.NOM.SG Vīštāspa.NOM.SG   reach.3SG.INJ divine.GEN.SG
      padəbīš   manaŋhō ...
      path.INSTR.PL thought.GEN.SG
      'By the paths of the divine Thought the sage Vīštāspa reaches (the idea that ...)' (Y. 51.16)

3.1.3 Classical Armenian

In Classical Armenian we find the ending *-bʰi for both the Instrumental singular and the plural (see Clackson 1994: 68–74 on different hypotheses about this).[3] Although other semantic roles such as Force, Cause or Quantity could be expressed by the instrumental case in the text of our corpus, the examples found are not undisputed. The nouns with this ending are associated with the following semantic roles:

INSTRUMENT

(19) cecein   z-glowx-n   ełegamb   (*Mk.* 15,19)
      hit.3PL.IMPF ACC-head.ACC.SG-ART.SG   cane.INSTR.SG
      'they hit his head with a cane'

COMITATIVE

Comitatives are usually expressed by means of a phrase combining the postposition *handerj* 'with' and the Instrumental case of the noun (Schmitt 1981: 92), but there are a few occurrences in which the plain Instrumental has this role by itself.

---

3 In addition to checking grammars and studies, such as A. Meillet (1936²), H. Jensen (1959), R. Schmitt (1981) or J. Matzinger (2005), we have analysed all the occurrences of the Instrumental in the Armenian translation of the New Testament (the four Gospels).

(20) ałałakein   amenayn   bazmowt'eamb-n   (*Lk*. 23, 18)
     shout.IMPF.3SG   all   mass.INSTR.SG-ART
     'they shouted together with all the people'

AGENT

The Instrumental case can be used for Agents in passives, as in the following example:

(21) kataresc'én   amenayn   grealk'-n
     fullfilAOR.PASS.3PL. everything write.NOM.PL.PTCP-ART
     margarēiwk'
     prophet.INSTR.PL
     'everything written by the prophets will be accomplished' (*Lk*. 18, 31)

Jensen (1959: 180) claimed that in Armenian the Agent was expressed by means of a preposition and the ablative case if the Agent noun refers to a human entity; in contrast, if the Agent noun refers to a non-animate entity, it would be expressed by the instrumental case, but this goes against the definition of the semantic role of Agent itself, because the entity that is assigned the semantic role of Agent must be animate.

MANNER

(22) k'ałc'rowt'eámb   lsēr   nma   (*Mk*. 6, 20)
     sweetness.INSTR.SG   listen.IMPF.3SG   DEM.DAT.SG
     'He (Herod) listened to him (John) with sweetness'

TIME

(23) xndrēr   t'e   ziard parapov   matnesc'ē
     seek.3SG.IMPF CONJ   how   moment.INSTR.SG   betray.AOR.SUBJ.3SG
     zna
     3SG.ACC
     'He sought how he might in the appropriate moment betray him'
                                                                     (*Mk*. 14, 11)

LOCATIVE

We have not found any occurrences of the plain Instrumental as a prototypical Locative in our corpus. Prepositional phrases with *aṙ* ('near'), *ənd* ('under'), *z-* ('above') and *šowrǰ* ('around') plus Instrumental are the usual way to convey that meaning in Classical Armenian (see Minassian 1976: 251–256). Nevertheless, there are a few examples of metaphorical Location:

(24) zi   héz   em   ew xonarh srtiw   (*Mt.* 11, 29)
     because mild be.PRS.1SG and humble heart.INSTR.SG
     'since I am mild and humble at heart'

3.1.4   Greek

The ending *-phi* occurs only in the oldest extant records of the Greek language. In Mycenaean it appears as *-pi* (= [pʰi]), usually with an Instrumental value. For the syntax of these forms in Mycenaean, see Bernabé & Luján (2006: 228–229) and Jiménez Delgado (2016: 93–98). A more detailed functional analysis of the Mycenaean *-pi*-ending, can be found in López Chala (2014: 113–124) and a comprehensive list of the occurrences of *-pi* is available in Francisco Aura's "Inverse index of the groups of transliterated syllabograms" in the web page of the *Diccionario Micénico* [Mycenaean Dictionary].[4]

However, the Instrumental use of *-pi* in Myceanean is not the most prototypical, as shown in PY Ta 714.3:

(25) *ta-ra-nu*   *a-ja-me-no*   *ku-wa-no*
     footstool.NOM.SG   inlaid.NOM.SG   lapislazuli(?).DAT.SG
     pa-ra-ku-we-qe   *ku-ru-so-qe*   *ku-ru-sa-pi-qe*
     emerald.DAT.SG-and gold.DAT.SG-and golden.INSTR.PL-and
     ko-no-ni-pi
     fringe.INSTR.PL
     'a footstool inlaid with lapislazuli(?) and emerald(?) and with golden fringes' (PY Ta 714.3)

Given its specific use in the extant Mycenaean tablets, Waanders (1997: 69–74) remarks that its meaning is not exactly Instrumental, but "ornamental" or "supplementary".

Interestingly, *-pi*-forms are frequently found with place-names. It has been disputed whether they should be interpreted as Locatives or Ablatives. According to the evidence found in the Linear B tablets this possibility was restricted to occurrences in which the noun suffixed by *-pi* was a place-name, as in the following example:

(26) *ma-ro-pi*   *pa-ro ro-ko*   *pa-ra-jo*   OVISᵐ 150 (PY Cn 40.8)
     Malon.INSTR from Lokhos.DAT.SG old.NOM.PL ram   150
     'in Malon, from Lokhos, old rams, 150'

---

4   http://bib.cervantesvirtual.com/portal/diccionariomicenico/contenido/inverso.jsp.

In one case (PY Un 1426.5), the ending *-pi* combines with a noun that may refer to a group of female workers (]*ki-ri-te-wi-ja-pi*), but the context of the tablet does not allow for a straightforward interpretation. It has been suggested that it has an Ablative function, but an agentive or comitative interpretation cannot be completely ruled out, either (see Hajnal 1995: 204–207 for a possible interpretation as a Comitative; however, in a recent article Bernabé 2014: 13–15 has argued that this term is an abstract, which would exclude that interpretation).

A form with *-pi* is also found with the preposition *o-pi* (= *opí*), which usually governs Datives referring to human beings. This only occurs with the phrase *o-pi … qe-to-ro-pi* (= *opì kʷetrópoppʰi*) '(watching) over the cattle' in PY Ae 108.a. In the last published tablet from Thebes there is a new occurrence of a prepositional phrase with a *-pi*-form: *pa-ro te-qa-jo qa-si-re-u-pi* 'from the Theban chiefs', in a quite interesting combination with a noun referring to humans.

In Homeric Greek, the instrumental ending *-phi* poses various morphological problems that will not be dealt with in this article (for an excellent account of the Homeric facts see Nieto 1987, where different theories about the syntactic values of *-phi* can be found). From a syntactic point of view, it has been argued (e.g. Lejeune 1956: 208) that *-phi* is used as both Genitive and Dative, but this assertion must be understood in the sense that it shares some of the values of those morphological cases. Specifically, it shares the Ablative value with the Genitive of Homeric Greek; however, it should be noted that the Ablative value of the *-phi*-forms in Homer is almost exclusively restricted to prepositional phrases – only one occurrence of an Ablative value without preposition is attested (Nieto 1987: 297):

ABLATIVE

(27) naûphin   amunómenoi   metà   Boiōtôn
      ship.INSTR.PL protect.PTCP.PRES.NOM.PL with   Beotians.GEN
      emákhonto
      fought.3PL
      'They fought together with the Beotians to protect (the ships by keeping
      the enemies) from the ships.' (*Il.* 13.700)

LOCATIVE
The Locative value of the *-phi*-case in Homer is well attested:

(28) dóru   makrón,   ho   hoi   klisíēphi
      spear.ACC.SG big.ACC.SG REL.ACC.SG 3SG.DAT tent.INSTR
      léleipto
      leave.3SG.MID.PLPF
      'the big spear that he had left in his tent' (*Il.*13.168)

INSTRUMENT

(29) autàr Odysseùs   kheír'   epimassámenos
     but   Odysseus.NOM.SG   hand.DAT.SG feel.for.PTCP.NOM.SG
     pháryngos   lábe   deksiterêphi
     throat.GEN.SG took.3SG   right.INSTR
     'Odysseus felt for her throat and took it with his right hand' (*Od.* 19.480)

DATIVE

The only clear instance is *Iliad* 2.363, for the rest of the alleged occurrences appear in combination with the verb *pepoithénai* 'persuade'.

(30) hōs phrḗtrē   phrḗtrēphin arḗgēi (*Il.*2.363)
     so   clan.NOM.SG   clan.INSTR   succour.PRS.SUBJ.3SG
     'so that the clan succours the clan'

SOURCE

(31) ... hoppóte naûphin   aphormētheíen   Achaioí (*Il.*2.794)
     when   ship.INSTR.PL   depart.IMPF.OPT.3PL   Achaeans. NOM.PL
     '... when the Achaeans would depart from their ships'

It could also be argued that in certain occurrences the *-phi*-endings are associated with the roles of Manner, Cause, and Comparison in Homeric Greek, but the examples are rare and uncertain.

Outside Mycenaean and Homer, the ending *-phi* only occurs in certain poets (Hesiod, Alcman and Ibycus) as a Homerism, in an inscription from Cyrene (*karophi*), and in three glosses in Hesychius' lexicon (*eurésphi*, *Ídēphi*, *passalóphi*), who characterises them as belonging to the Beotian dialect.[5] The fact that this dialect must have had this ending can be supported by the adjective *epipatrophion*, occurring on an inscription from Tanagra, which presupposes the phrase *epi patrophi (see Nieto 1987: 274 with further references).

3.1.5   Celtic

In continental Celtic *-*bʰi*-endings are only attested in Gaulish (not in Celtiberian). We will analyse the evidence provided by this language in full and will then refer briefly to the evidence provided by Irish.

---

5 *SEG* XX 756. The meaning is not clear; different interpretations of this inscription can be found in Morpurgo-Davies (1969: 49).

3.1.5.1　*Gaulish*

In Gaulish the ending *-bi* appears in the inscription from Alise-Sainte-Reine (RIG L-13), which reads as follows (a complete analysis of the meaning and uses of this ending in Gaulish can be found in López Chala 2015: 13–21):

(32)　Martialis ·　　　Dannotali /　　　　ieuru ·
　　　Martial.NOM.SG　Dannotalos.GEN.SG　consecrate.PRT.3SG
　　　Vcuete ·　　　　sosin /　　　celicnon　　　　etic /　gobedbi ·
　　　Vcuetis.DAT.SG　DEM.ACC.SG　building(?).ACC.SG　CONJ　smith.INSTR.PL
　　　dugiiont-io /　　　　Vcuetin /　　　in [...]　Alisiia
　　　revere.3PL.PRS-REL　Vcuetis.ACC.SG　in　　　Alisia.LOC.SG
　　　'Martial, son of Dannotalos, consecrated this building (?) to Vcuetis
　　　and (he did this) together with the smiths that revere Vcuetis at Alisia'.

The word *gobedbi* has been identified since the early years of the 20th century as the 'Dative' plural of the noun meaning 'smith' based on parallels in insular Celtic, but its syntactic analysis as a Comitative was first proposed by Lejeune (1979) and it is generally accepted by the research community (Lambert 1994: 99–100). Meid (1992: 29–30), however, believed that *gobedbi* and *Vcuete* were Beneficiaries, while, according to Schriver (1997: 182) *gobedbi* would be the Agent of this sentence.

　　The Comitative value of this ending recurs in the inscription from Néris-Les-Bains (RIG L-6):

(33)　Bratonos /　　　　Nanton{t}in(os) /　Epađatexto/rigi ·
　　　Bratonos.NOM.SG　PATR.NOM.SG　　Epadatectorix.DAT.SG
　　　leucutio / suiorebe ·　　　　　　logi/toi
　　　sacred.forest(?).ACC.SG sister.INSTR.PL　establish.PRT.3SG
　　　'Bratonos, son of Nantonos, established a sacred forest (?) for Epadatec-
　　　torix together with his sisters'.

Suiorebe, with the phonetic evolution of final *-i* into *-e*, shows the ending *-bi* in the word meaning 'sister', so it could probably be interpreted as a Comitative. Possibly, the ending *-bi* must also be identified in the sequence *eiabi*, the Dative plural of the anaphoric pronoun in the Larzac lead inscription (Lambert 1994: 57). In this case, the syntax of the inscription allows for its interpretation as a Comitative 'with them (the women)'.

　　The ending is now attested on the Châteaubleau tile, the text of which is far from clear, and its interpretation is, as a result, rather controversial.[6] The

---

6　On the Châteaubleau tile see Lambert (1998–2000) and Schrijver (1998–2000).

word *anmanbe* appears twice on this inscription. It is possibly the Gaulish word for 'name' with a *-be*-ending, thus being an Instrumental plural with the semantic role of Means. We provide one of the occurrences below, and suggest a very tentative interpretation, based on Lambert's (1998–2000) and Schrijver's (1998–2000) proposals.

(34) *neI*   *anmanbe*   *gniIou*   (*RIG* L-93)
     NEG   name.INST.PL   know.PRS.1SG
     'and I do not know (her) by her names'

The ending *-bi* is also to be identified in the words *mesamobi* 'worst' (cf. OIr. *messam*) and *gandobe* 'rare (?) (cf. OIr. *gand* 'rare'), both in an inscription on a dish from Lezoux (*RIG* L-66) (Lambert 1994: 146–147). The context, however, is too fragmentary to propose a syntactic analysis. Most probably, though, *-bi* has an Instrumental value there.

3.1.5.2 *Old Irish*

The plural Dative case of Old Irish comes from **-bʰi(s)*, as demonstrated by the palatal character of the final *-b* of this case, as in the examples shown in Table 9.9.

Traditionally, the Dative plural ending of Old Irish has been reconstructed as **-bʰis* with final *-s* based on the alleged correspondence with Skt. *-bhis*, but, as Thurneysen (1946: 182) remarked, the Irish data do not allow for deciding whether that final *-s* was lost or never existed. As discussed above, the Gaulish data seem to favour the view that the ending was **-bʰi* and, as a matter of fact, the last part of the Ardmore Ogham inscription (Macal. no. 208) reads: *Dolatibigaisgob …* It probably includes a *-bi* Dative.[7]

TABLE 9.9   Dative plural and dual in Old Irish (examples)

NOM. SG.	*fer* 'man'	*túath* 'tribe, people'	*cathir* 'town'
DAT. PL.	*fer(a)ib*	*túath(a)ib*	*cathrach(a)ib*
DAT. DU.	*fer(a)ib*	*túath(a)ib*	*cathrach(a)ib*

---

7 Against Thurneysen (1946: 182) himself, who, despite acknowledging that there were no reasons for reconstructing the *-s* in **-bʰi*, still thought that it was more probable.

As for the values of the so-called 'Dative' of the Old Irish, it is important to stress that it is a prepositional case with no inherent syntactic meaning of its own, but its meaning completely depends on the preposition that governs it. In fact, in his textbook, Stifter (2006: 37–38) has chosen to call it the 'prepositional case' instead of Dative. He states:

> The prepositional case is called *dative* in all traditional grammars. In my view this is an unfortunate designation, since the Old Irish prepositional is a syncretism case going back functionally and formally to four different Common Celtic and Indo-European cases: dative, instrumental, ablative, and locative. The datival function (that is, the function as the indirect object of the verb) of this case is nowhere prominent in Old Irish.

Stifter (2006: 38) notes that in poetry and legal texts, which can be considered representative of an "archaic" language, independent occurrences of this case are found, practically with all readings. In most instances, however, its value is Instrumental, even if, in some of them, this may be a non-prototypical Instrument. This is the case with *segdaib súiliub seallglausaib* 'with shining blue-grey eyes', which appears in the description of a woman in one of the poems included in *The Exile of the Sons of Usnech* (Windisch 1880: 69).[8] In fact, this use is to a great extent similar to what we found in Mycenaean (see § 3.1.4). On the uses of the Dative in Old Irish, see also Müller (1999: 177–178), who found occasional occurrences of prepositionless Datives as Instrumentals in her corpus, almost exclusively in legal texts. She also pointed out that she found three examples of the prepositionless dative denoting Manner rather than Instrument. However, *-ib* has sometimes a Sociative meaning, as in the phrase *cléirchib tuathaib* 'with clergymen and laymen' (Lambert 1994: 62).

3.2   *General Overview of the Data*

Table 9.10 summarises the semantic roles associated with nouns marked by *-bʰi*-endings in the relevant ancient Indo-European languages.

---

8   We thank an anonymous reviewer for pointing us to this example.

TABLE 9.10 Semantic roles associated *-bʰi-endings in the nominal declension and in pronouns (other than personal)

		Instrument	Locative	Sociative	Agent (with passives)	Ablative	Means	Manner	With adposition
Greek	Mycenaean	x	x (only with place-names)	(x)					x (with *o-pi* and *pa-ro*)
	Homer	x	x			(x)			x
Celtic	Old Irish	x (most frequent use of independent 'Dative' in archaic language)		x					x
	Gaulish	(x)		x (only with human beings)					
Vedic		x	x	x	x		x	x	x
Old Persian		x		x (non prototypical)			x		x
Avestan		x	x (Perlative)	x			x		
Armenian		x		x	x		x	x	x

3.3 *-bʰi *in the Pronominal Declension*

A variant of *-bʰi* also appears in the Dative case with certain personal pronouns, such as 2nd Sg. Skt. *tubhyam* or Latin *tibī*. The long -ī has usually been analysed as resulting from the *e*-grade (*-bʰei*), however, Jasanoff (2009: 140) analyses *-bʰei* as a reshaping of *-bʰi* with the Dative singular ending *-ei.*[9]

---

9 See Mendoza (1998: 37–38) for an analysis of these forms. Szemerényi (1996: 218) analysed *t(w)-ebhi* (> *tibī*) as containing the postposition *ebhi* 'to', which, with either *e*- or *o*-grade, results in Skt. *abhi*, OPers. *abiy*, and OCS *obŭ*.

The attested forms are the following:

- Sanskrit 2nd sg. Dative *tubhya*(*m*), 2nd pl. Dative *yuṣmábhya*(*m*), 2nd pl. Instrumental *yuṣmā́bhis*, 1st pl. Dative *asmábhya*(*m*), 1st pl. Instrumental *asmā́bhis*
- Old Avestan: 1st sg. Dative *maibiiā*, 1st pl. Dative *ahmaibiiā*, 2nd sg. Dative *taibiiā*, 2nd pl. Dative *yūšmaibiiā* and *xšamaibiiā*
- Classical Armenian: 1st sg. Instrumental *inew*, 1st pl. Instrumental *mewkʻ*, 2nd sg. Instrumental *kʻew*, 2nd pl. Instrumental *jewkʻ*; Reflexive pronoun sg. Instrumental *iwrew*, *iwreaw*, *iwreamb*, pl. Instrumental *iwreambkʻ*
- Slavic: OCS 2nd sg. Dative-Locative *tebě*, Reflexive pronoun Dative-Locative *sebě*
- Baltic: OPr. 2nd sg. Dative *tebbei*, Reflexive pronoun Dative *sebbei*
- Latin: 2nd sg. Dative *tibī*, 2nd pl. Dative-Ablative *uōbīs*, 1st pl. Dative-Ablative *nōbīs*, Reflexive pronoun Dative *sibī*
- Oscan 2nd sg. Dative **tfei**, Reflexive pronoun Dative **síbei**; Umbrian *tefe*

The fact that Latin has *nobīscum* 'with us' (Dat. *nobīs* + *cum* 'with') and *uobīscum* 'with you' (Dat. *uobīs* + *cum* 'with') as special forms for the Sociative cannot be taken as evidence that there was an evolution from Instrument to Comitative, as argued by Stolz (1998), because with 1st and 2nd persons an Instrumental value can hardly be expected. Instead, they must be accounted for along the same lines as Luraghi (2001a) argued for the Greek prepositions *sýn* and *metá*: when the Comitative markers come to be used as Instruments, they need to be further characterised to keep them distinct.

Table 9.11 provides a summary of the values of the *-bʰi* foms in the pronouns. For our purposes, we use labels corresponding to at least two levels of case,

TABLE 9.11 Meanings associated to personal pronouns with **-bʰ(e)i*-endings

	Ablative	Dative	Locative	Instrumental	?	Comitative
Latin *tibi*	NO (*te*)	x	NO (*te*)	NO (*te*)		NO (*tecum*)
Latin *nobis/uobis*	x	x				NO: *nobiscum/uobiscum*
OCS		x	x			
Vedic	NO (*tvat*)	x	NO (*tvayi*)	NO (*tvayā*)		NO (*tvayā*, usually with *sam* or other)
Old Avestan		x				
Classical Armenian					x	

including morphological case and semantic function. However, the standard grammars of the Old Indo-European languages do not provide consistent information about the semantic roles associated with the various pronominal cases. For discussion of this problem and its consequences for typological cross-linguistic comparison when trying to draw semantic maps, see Narrog (2010: 237–238).

3.4 Semantic Roles of the *-bʰi-endings

We now aim at providing an interpretation of the data involving the *-bʰi-endings that we have found in the ancient Indo-European languages. However, before discussing those data, it will be convenient to quickly revise the evidence generally accepted for the grammaticalisation of Instruments. Like Narrog (2010: 236), we can assume with the scholarly literature on grammaticalisation (e.g. Heine 2003: 579, Hopper & Traugott 2003: 125) that, concerning the relationship between synchronic multifunctionality and diachronic meaning extension, a morpheme is expected to acquire a new function (B) while retaining the older one (A) and only at a later stage will the older meaning eventually be given up. Schematically, the expected semantic evolution is: A > A+B (>B). That is, for our purposes, it is expected that the various meanings associated with *-bʰi-endings in the Old Indo-European languages can be linked with each other in this way, through a series of semantic changes.

Concerning the specific evidence generally assumed for the grammaticalisation of Agents and Instruments, the information provided in the *World Lexicon of Grammaticalization* (Heine & Kuteva 2002) is summarised in Table 9.12.[10]

TABLE 9.12 Paths of grammaticalisation of Agents and Instruments (based on the data of Heine & Kuteva 2002)

Source	Target
COMITATIVE >	INSTRUMENT
INSTRUMENT >	ERGATIVE
	MANNER
COMITATIVE >	AGENT*
	INSTRUMENT
	MANNER*
	TEMPORAL*

---

10 For the concepts of "source" and "target" of grammaticalisation and how they can be framed in the general theory of grammaticalisation, see Heine & Kuteva (2002: 6).

Bearing this background information in mind, we can now return to the analysis of the meanings of the *-*bʰi*-endings and the interpretations proposed so far. Burrow (1973: 239) and Villar (1974: 325–326) argued that the *-*m*- and *-*bʰ*- must have originally had a wider and vaguer use, which was later restricted in the individual languages when they were grammaticalised as cases of the nominal inflection. This view, however, seems to go against the evidence of the grammaticalisation processes that we know now. Nevertheless, Villar must be right when he states that *-*bʰ*-endings were originally not marked for number. He analyses *-*bhyas* as having Gen.-Abl. *-as* < *-*os*, *-bhyām* as having the dual ending *-*ā* < *-*ō*, and *-bhis* as showing the plural marker *-s*.

The evidence provided by the ancient Indo-European languages combined with what we know about semantic extension and change in the domain of the Instrumental-Comitative supports the view that the *-*bʰi*-forms were originally used as Comitatives (see Stolz, Stroh & Urdze 2006: 363–363 for unidirectionality in the evolution from Comitative to Instrument). This has been, in fact, the traditional assumption in Indo-European linguistics, as Luraghi (2001b) points out. This goes back to Delbrück (1867: 50), Wenzel (1879) and also Hübschmann (1875) for the Iranic Instrumental case.

In the nominal declension, these forms evolved into Instruments. We can hypothesise that this happened via its uses as non-prototypical Comitatives, as suggested by the evidence of the Old Indo-European languages themselves and paralleled by better documented changes (Luraghi 2001b). Once these forms evolved into Instruments they followed the different paths of semantic change and extension that we have summarised in Figure 9.1, according to the data that we analysed in the previous sections.

Following a tendency for Comitative to require more morphological marking than Instrument (Stolz 1998, Luraghi 2001b), plain *-*bʰi*-forms only rarely appear with this value in the ancient Indo-European languages. They have been replaced in such uses by combinations of Instrumental plus adposition, as clearly shown by Latin *nobiscum* and *uobiscum*. The only exception seems to be Gaulish, in which, in contrast to the other Old Indo-European languages, the Comitative value of the *-bi*-forms is clearly attested; however, given the fragmentary preservation of this language we cannot be sure that Comitative was actually the most frequent meaning of the ending in that language. Some instances of the Comitative value of the plain Dative can be found in Old Irish, but the Instrumental value is more frequent in "archaic texts", as stated above.

In contrast, in the pronominal inflection, given that the personal and reflexive pronouns intrinsically refer to human beings (and other animate entities

such as gods) this change would not be expected. This is why *-*bʰei*-forms do not appear as Instruments in the pronominal declension unless there has been an analogical extension from the nominal declension (as in 1st and 2nd plural personal pronouns in Sanskrit). The *-*bʰei*-forms in the pronominal declension have evolved, instead, from Comitatives into Datives, which is also a semantic role prototypically assigned to human beings.

We have mapped all the information that we have been able to recover about the -*bʰi*-endings from the analysis of the ancient Indo-European languages in the semantic map in Figure 9.1. Semantic maps have been developed in recent years as an important methodology for the analysis of the multifunctionality of grammatical morphemes. As stated by Haspelmath (2003: 213), a semantic map has the advantage that it "does not imply a commitment to a particular choice among monosemic and polysemic analyses". From a cognitive perspective, assuming that grammatical morphemes, like any other meaningful element in language, have a structured polysemy (cf. Geeraerts 1997), the aim of syntactic reconstruction would be to recover the syntactic and semantic path followed by a given ending until it reached its attested meaning in the older languages.

The map in Figure 9.1 has been adapted from Narrog's (2010) article, in which he reviews the evidence offered in previous studies concerning the semantic extension of the grammatical markers for Instrument and related semantic
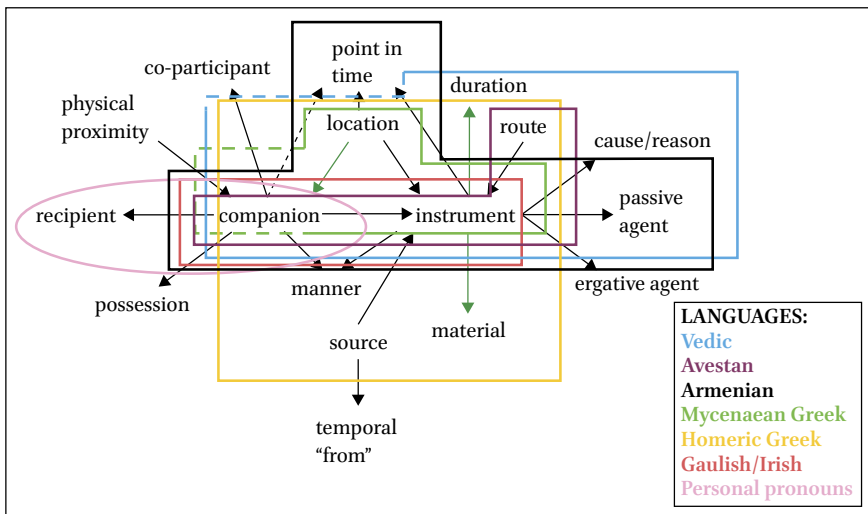
FIGURE 9.1   Semantic map of the *bʰi-endings in the ancient Indo-European languages (based on Narrog's 2010 map of the comitative-instrumental area)

roles (cf. also Narrog & Ito 2007). Narrog rightly points out that when the diachronic dimension is added, "classical" semantic maps allow for possibilities of representation that other types of map do not. We have charted onto his semantic map the actual extension of the *-b^hi-endings in the ancient Indo-European languages. For a general view about semantic maps and their use in the diachronic analysis of polysemy see Malchukov & Narrog (2008).

3.5　　　*On the Etymology of* *-b^{h}i

Further insights into the semantic and syntactic evolution of *-b^hi-forms can be obtained if we look into the etymology of these forms. This ending has traditionally been related to the preposition *bi found in the Germanic languages (Haudry 1982: 24–25; Schmitt-Brandt 1998: 216; Fortson 2010: 118–119, among others): Goth. *bi*, OHG *bī*, etc. This preposition appears as a complex preposition in other languages, such as Gk. *amphí*, OHG *umbi*, OCS *obi*, etc. It also appears in adverbs such as Lat. *ubi* 'where?', *ibi* 'there', Gk. νόσφι 'far away' or Hitt. *kuwapi* 'ever'.

According to Berenguer (2000: 394–404) and Jasanoff (2009: 199) the following adverbs are reconstructable for Proto-Indo-European:[11]

- *h_1e/o-b^{h}í* 'thereabouts, by that way' > 'to, near, across around': Ved. *abhí*, OCS *obĭ*, Goth. *bi*
- *h_2(e)mb^{h}í* 'sidewise' > 'around, on both sides of' (reconstructed as *$h_2$nt-b^{h}i*, cf. *h_2(e)nt-* 'front, side' by Jasanoff 1976): Gk. *amp^{h}í*, Lat. *am(b)-*, Gaulish *ambi-*, OIr. *imm*, OHG *umbi*, etc.
- *k^wo-b^{h}í* 'where, when': Hitt. *kuwapi*, Lat. *alic(ubi)*

This evidence would strongly suggest that *-b^hi was some kind of adverb or postposition meaning 'near' or 'next to'; that is, it conveyed the meaning 'physical proximity' that is charted in Figures 9.2 and 9.3 below. If so, this presupposes that the Sociative value that can be assumed as the starting point for the later development of the Instrumental value in the nominal declension and its usages as a Dative in the prononimal declension would have originated in turn from that local meaning 'near' or 'next to'. This seems to be in accordance with what we know about the semantic evolution of markers with local value. We have highlighted the probable paths of change in the nominal declension and in the pronouns of the ancient Indo-European languages in Figures 9.2 and 9.3 below.

11　　Hewson & Bubenik (2006: 345) suggest that in Indo-Iranian we may have a merger of *$h_3$eb^{h}i* 'to(wards)' and *h_2ṃb^{h}í* 'about, around'.
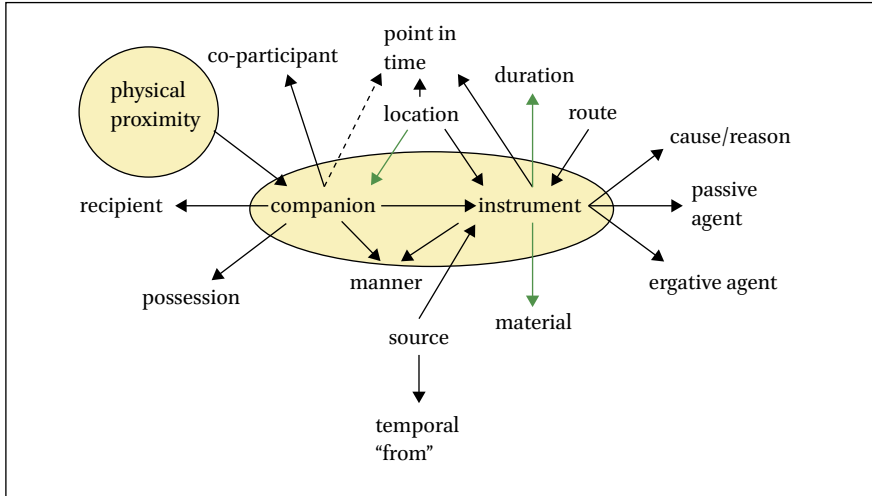
FIGURE 9.2 Path of change of *-bʰi-endings in nominal inflection (based on Narrog's 2010 map of the comitative-instrumental area)
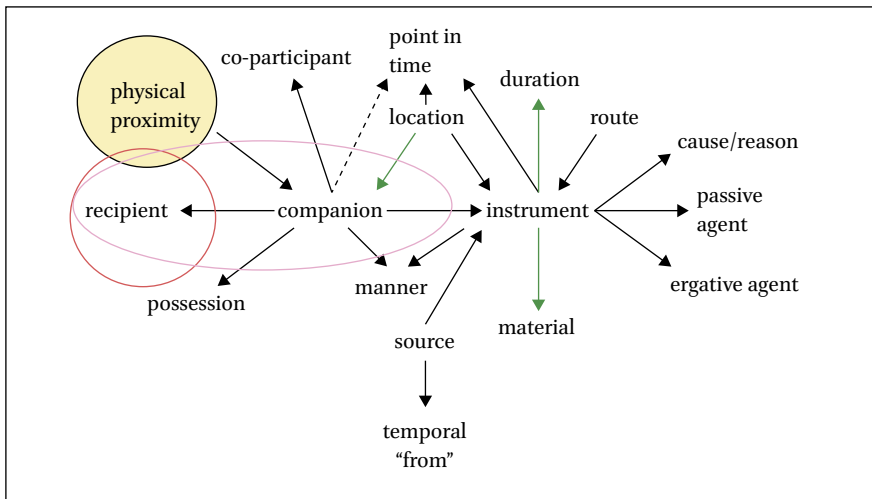
FIGURE 9.3 Path of change of *-bʰi-endings in pronominal inflection (based on Narrog's 2010 map of the comitative-instrumental area)

4 Concluding Remarks

Our aim has been to explore a new approach to reconstructing the nominal morphosyntax of a proto-language. Instead of trying first to reconstruct the

set of cases that may go back to the proto-language and then to assign them to an array of syntactic functions, regardless of the specific markers that are used for each of them, we have taken a different path. We have analysed in detail the semantic roles denoted by the words marked by a specific ending in the oldest phases of the languages of the Indo-European language family (Section 3.1). Our aim has been to comprehend how those meanings may relate to each other and what semantic changes must be reconstructed in order to account for the whole constellation of meanings.

We have focused on the 'Instrumental' *-$b^h i$-marker in the Indo-European language family. In the various branches of the family in which this ending occurs (Indo-Iranian, Greek, Armenian, and Celtic), *-$b^h i$-endings express a whole array of semantic roles, which encompass not only Instrument, Means, Intermediary or Comitative and Accompaniment, but also Agent, Cause and Manner, as well as local semantic roles, such as Locative and Perlative, and Time (Sections 3.1 and 3.2).

The range of meanings that *-$b^h i$-endings are associated with in the Old Indo-European languages is, however, far from chaotic. When the meanings occurring in the different branches are charted in a diachronic semantic map of Instrument and related semantic roles (Figure 9.1), they appear to cover neighbouring areas and it becomes clear that they have followed well-known lines of semantic change. If we then apply recent knowledge about directionality in semantic change, a clear pattern emerges: when the semantic maps for the different branches are brought together (Section 3.4 and Figure 9.1), all the meanings displayed by *-$b^h i$-endings can be diachronically derived from the core area of the Comitative and Instrument roles, even if some branches may have gone further in the evolution along a given line (e.g., passive Agent, Duration or Manner). The meanings associated with the endings are better accounted for if we analyse them in an integrated way.

Furthermore, when the information about the semantic roles expressed by *-$b^h i$-endings in the pronominal declension is added (Section 3.3), an even more interesting pattern surfaces: pronominal *-$b^h i$-endings are quite systematically related to the Dative case (i.e. the case that typically expresses the Recipient) in the Old Indo-European languages. The link to the Instrumental value in the nominal declension lies in the Comitative. The different semantic evolution in the nominal and the pronominal declensions can be explained on the basis that 1st and 2nd person pronouns refer to humans and a change from Comitative > Instrument cannot, therefore, be expected, since the semantic role of Instrument is associated with non-human, inanimate entities. In contrast, a change from Comitative > Recipient is in line with the fact that both semantic roles are typically associated with animates, specifically humans. This

shows that the semantic and lexical features associated with the base to which the case ending is added are relevant to account for the divergent evolution of the meanings of the same ending in the nominal and pronominal inflections.

If we then integrate the etymological information about the ending *-b^hi, which is related to adverbs and adpositions in various Indo-European languages, a clear grammaticalisation process is unveiled (Section 3.5). It starts with a postposition meaning 'near' or 'next to' and results in a grammatical case associated with a Comitative value and, later on, either to Recipient (pronominal declension) or Instrument (nominal declension) (Figures 9.2 and 9.3). The change Comitative > Instrument is further supported by the fact that in various Old Indo-European languages Comitatives cannot be expressed by the plain *-b^hi-endings and need additional marking (prepositions), a tendency that has been pointed out in previous studies: when a Comitative marker comes to also express Instrument, the old marker is usually not enough to encode the Comitative.

In conclusion, the perspective that we have adopted in this article allows for understanding the semantic changes undergone by the *-b^hi-marker and reconstructing the paths of change, from the oldest reconstructable stages to the actual attested meanings in the earliest phases of the languages. This results in a 'dynamic' reconstruction of the morphosyntax of the proto-language. This in turn is more in accord with what we know about the actual processes of semantic change in grammatical markers and paradigmatisation of markers, as opposed to a more traditional, 'static' reconstruction. In those older approaches, even when the morphological reconstruction is not straightforward, morphology and syntax operate at two separate levels and in two successive operational phases, with syntax only entering the stage once the morphological level has been reconstructed. In contrast, our dynamic approach allows for a more detailed and realistic reconstruction.

Abbreviations

ABL	ablative
ACC	accusative
Amph.	*Amphitruo*
AOR	aorist
ART	article
Av.	Avestan
CONJ	conjunction
DAT	dative

DB	Inscription of Darius I from Bisutun
DEM	demonstrative
DNb	Inscription of Darius in Naqsh-i-Rustam (text b)
DU	dual
GEN	genitive
Gk.	Greek
Goth.	Gothic
Hitt.	Hittite
Il.	*Iliad*
IMPF	imperfect
IMPV	imperative
INF	infinitive
INJ	injunctive
INSTR	instrumental
Lat.	Latin
Lk.	Gospel of Luke
LOC	locative
MID	middle
Mk.	Gospel of Mark
Mt.	Gospel of Matthew
NOM	nominative
NEG	negation
OCS	Old Church Slavonic
Od.	*Odyssey*
OHG	Old High German
OIr.	Old Irish
OPers.	Old Persian
OPT	optative
PASS	passive
PATR	patronymic
PERF	perfect
PIE	Proto-Indo-European
PL	plural
Plaut.	Plautus
PLPF	pluperfect
POSS	possessive
PRT	preterite
PRON	pronoun
PTCP	participle
PTCL	particle

PRS	present
PY	Pylos
REL	relative
RIG	*Recueil des Inscriptions Gauloises*
RV	*Ṛgveda*
SEG	*Supplementum Epigraphicum Graecum*
SG	singular
Skt.	Sanskrit
SUBJ	subjunctive
Tokh.	Tokharian
Ved.	Vedic
VOC	vocative
Y.	Yasna

References

Beekes, Robert Stephen Paul. 2011. *Comparative Indo-European Linguistics: An Introduction* (2nd edition, revised and corrected by Michel de Vaan). Amsterdam/ Philadelphia: John Benjamins.

Berenguer, José Antonio. 2000. *Estudio sobre las partículas indoeuropeas con base consonántica y laringal* (Manuales y Anejos de Emerita XLII). Madrid: C.S.I.C.

Bernabé, Alberto. 2014. Los términos micénicos en *-e-wi-ja*, *Mycenaean Studies in Honour of Francisco Aura Jorro,* ed. by Alberto Bernabé & Eugenio R. Luján, 1–19. Louvain-la-Neuve: Peeters.

Bernabé, Alberto & Eugenio R. Luján. 2006. *Introducción al griego micénico: Gramática, selección de textos y glosario*. Zaragoza: Depto. de Ciencias de la Antigüedad de la Universidad de Zaragoza.

Burrow, Thomas. 1973. *The Sanskrit Language*. London: Faber & Faber.

Clackson, James. 1994. *The Linguistic Relationship Between Armenian and Greek*. Oxford/Cambridge MA: Blackwell.

Delbrück, Berthold. 1867. *Ablativ localis instrumentalis im Altindischen, Lateinischen, Griechischen und Deutschen: Ein Beitrag zur vergleichenden Syntax der Indogermanischen Sprachen*. Berlin: Dümmler.

Fortson, Benjamin W. 2010. *Indo-European Language and Culture. An introduction* (2nd edition). Malden/Oxford: Wiley-Blackwell.

Fritz, Matthias. 2003. Proto-Indo-European Syntax. *Indo-European Linguistics*, ed. by Michael Meier-Brügger, 238–276. Berlin/New York: Walter de Gruyter.

Geeraerts, Dirk. 1997. *Diachronic Prototype Semantics*. Oxford: Clarendon.

Hajnal, Ivo. 1995. *Studien zum mykenischen Kasussystem*. Berlin/New York: de Gruyter.

Haspelmath, Martin. 2003. The Geometry of Grammatical Meanings: Semantic Maps and Cross-Linguistic Comparison. *The New Psychology of Language*, ed. by Michael Tomasello, 211–242. Mahwah: Erlbaum.

Haudry, Jean. 1977. *L'emploi des cas en védique*. Lyon: L'Hermès.

Haudry, Jean. 1982. *Préhistoire de la flexion nominale indo-européenne*. Lyon: Institut d'études indo-européennes de l'Université Jean Moulin (Lyon III).

Heine, Bernd. 2003. Grammaticalization. *The Handbook of Historical Linguistics*, ed. by Brian D. Joseph & Richard D. Janda, 575–601. Oxford: Blackwell.

Heine, Bernd & Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

Hewson, John & Vit Bubenik. 2006. *From Case to Adposition: The Development of Configurational Syntax in Indo-European Languages*. Amsterdam/Philadelphia: John Benjamins.

Hoffmann, Karl & Bernhard Forssman. 1996. *Avestische Laut- und Flexionslehre*. Innsbruck: IBS.

Hopper, Paul J. & Elizabeth Closs Traugott. 2003. *Grammaticalization*. Cambridge: Cambridge University Press.

Hübschmann, Heinrich. 1875. *Zur Casuslehre*. München: Theodor Ackermann.

Jasanoff, Jay. 1976. Gr. *ámphō*, lat. *ambo* et le mot indo-européen pour 'l'un et l'autre', *Bulletin de la Société de Linguistique* 71: 125–131.

Jasanoff, Jay. 2009. **-bhi*, **-bhis*, **-ōis*: Following the Trail of the PIE Instrumental Plural. *Internal Reconstruction in Indo-European: Methods, Results, and Problems* (*Section Papers from the XVI International Conference on Historical Linguistics, University of Copenhagen, 11th–15 August, 2003*), ed. by Jens Emelgård Rasmussen & Thomas Olander, 137–149. Copenhagen: Museum Tusculanum Press.

Jensen, Hans. 1959. *Altarmenische Grammatik*. Heidelberg: Carl Winter.

Jiménez Delgado, José Miguel. 2016. *Sintaxis del griego micénico*. Sevilla: Universidad de Sevilla.

Lambert, Pierre-Yves. 1994. *La langue gauloise*. Paris: Errance.

Lambert, Pierre-Yves. 1998–2000. La tuile gauloise de Châteaubleau (Seine-et-Marne), *Études Celtiques* 34: 57–115.

Lejeune, Michel. 1956. La désinence -φι en mycénien, *Bulletin de la Sociéte de Linguistique* 54: 187–218.

Lejeune, Michel. 1979. Notes d'étymologie gauloise: – III: Retour sur ΡΟΚΛΟΙΣΙΑΒΟ; – IV. Le pilier des Nautes Parisiaque et la 3ª personne du pluriel en gaulois, *Études Celtiques* 16: 101–116.

López Chala, Ángel. 2014. La desinencia -*pi* (-φι) en griego micénico: un análisis funcional, *Ardua Cernebant Iuuenes* (*Estudios Clásicos*, Anejo 2), ed. by Violeta Gomis, Alberto Pardal & Jesús de la Villa, 113–124. Madrid: Sociedad Española de Estudios Clásicos.

López Chala, Ángel. 2015. La desinencia *-*bhi* en galo: un análisis funcional, *E Barbatulis Puellisque. Actas del Segundo Congreso Ganimedes*, ed. by Mireia Movellán & Rodrigo Verano, 13–21. Sevilla: Universidad de Sevilla.

Luraghi, Silvia. 2001a. Syncretism and the Classification of Semantic Roles. *Sprachtypologie und Universalienforschung* 54(1): 35–51.

Luraghi, Silvia. 2001b. Some Remarks on Instrument, Comitative, and Agent in Indo-European. *Sprachtypologie und Universalienforschun* 54(4): 385–401.

Luraghi, Silvia. 2003. *On the Meaning of Prepositions and Cases: The Expression of Semantic Roles in Ancient Greek*. Amsterdam/Philadelphia: John Benjamins.

Luraghi, Silvia & Heiko Narrog. 2014. Perspectives on Semantic Roles: An Introduction. *Perspectives on Semantic Roles,* ed. by Silvia Luraghi & Heiko Narrog, 1–21. Amsterdam/Philadelphia: John Benjamins.

Malchukov, Andrej & Heiko Narrog. 2008. Case Polysemy. *The Oxford Handbook of Case* ed. by Andrej Malchukov & Andrew Spencer, 518–534. Oxford: Oxford University Press.

Mallory, James & Douglas Q. Adams. 2006. *The Oxford Introduction to Proto-Indo-European and the Proto-Indo-European World*, Oxford: Oxford University Press.

Matzinger, Joachim. 2005. *Untersuchungen zum altarmenischen Nomen: Die Flexion des Substantivs*, Dettelbach: J. H. Roll.

Meid, Wolfgang. 1992. *Gaulish Inscriptions: Their Interpretation in the Light of Archeological Evidence and their Value as a Source of Linguistic and Sociological Information*. Budapest: Archaeolingua.

Meier-Brügger, Michael. 2003. *Indo-European Linguistics*, Berlin/New York: Walter de Gruyter.

Meillet, Antoine. 1936. *Esquisse d´une grammaire comparée de l´arménien classique* (2nd edition), Vienna: Imprimérie des PP. Mekhitaristes.

Meiser, Gerhard. 1998. *Historische Laut- und Formenlehre der lateinischen Sprache*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Mendoza, Julia. 1998. Los pronombres personales, los posesivos y el reflexivo. *Manual de Lingüística Indoeuropea*, vol. 3, F. R. Adrados, A. Bernabé & J. Mendoza, 27–72. Madrid: Ediciones Clásicas.

Minassian, Martiros. 1976. *Manuel pratique d'arménien ancien*, Paris/Vaduz: Fondation des Frères Ghoukassiantz/Libraire Klincksieck.

Morpurgo-Davies, Anna. 1969. Epigraphical -φι. *Glotta* 47: 46–54.

Müller, Nicole. 1999. *Agents in Early Welsh and Early Irish*. Oxford: Oxford University Press.

Narrog, Heiko. 2010. A Diachronic Dimension in Maps of Case Functions. *Semantic Maps: Methods and Applications* ed. by Michael Cysouw, Martin Haspelmath & Andrej Malchukov. *Linguistic Discovery* 8(1): 233–254.

Narrog, Heiko & Shinya Ito. 2007. Reconstructing Semantic Maps: The Comitative-Instrumental area. *STUF–Sprachtypologie und Universalienforschung* 60(4): 273–292.

Nieto Hernández, Purificación. 1987. Un problema de la lengua homérica: la desinencia -φι, *Emerita* 55: 273–306.

Olander, Thomas. 2014. *Proto-Slavic Inflectional Morphology: A Comparative Handbook*. Leiden/Boston: Brill.

Pinault, Georges-Jean. 1989. Introduction au tokharien. *LALIES* 7: 3–224.

Reichelt, Hans. 1909 [reprinted 1978]. *Awestisches Elementarbuch*. Heidelberg: Carl Winter.

Schmitt, Rüdiger. 1981. *Grammatik des Klassisch-Armenischen mit sprachvergleichenden Erläuterungen*. Innsbruck: IBS.

Schmitt, Rüdiger. 1991. *The Bisitun Inscriptions of Darius the Great: Old Persian Texts*. London: School of Oriental and African Studies.

Schmitt-Brandt, Robert. 1998. *Einführung in die Indogermanistik*. Tübingen/Basel: A. Francke Verlag.

Schrijver, Peter. 1997. *Studies in the History of Celtic Pronouns and Particles* (Maynooth Studies in Celtic Linguistics 11). Maynooth: Department of Old Irish.

Schrijver, Peter. 1998–2000. The Châteaubleau Tile as Link between Latin and French and between Gaulish and Brittonic. *Études Celtiques* 34: 135–142.

Shields, Kenneth. 1974. Evidence of I.E. **-bhi* in Tocharian. *Folia Linguistica Historica* 11: 281–286.

Smitherman, Thomas & Jóhanna Barðdal. 2009. Typological Changes in the Evolution of Indo-European Syntax? *Diachronica* 26(2): 253–273.

Stifter, David. 2006. *Sengoídelc: Old Irish for Beginners.* Syracuse: Syracuse University Press.

Stolz, Thomas. 1998. Komitative sind mehr als Instrumentale sind mehr als Komitative: Merkmahlhaftigkeit und Markiertheit in der Typologie der mit-Relationen. *Sammelband des II. Internationalen Symposions zur Natürlichkeitstheorie, 23–25. Mai 1996*, ed. by Karmen Teržan-Kopecky, 83–100. Maribor: Pädagogische Fakultät.

Stolz, Thomas, Cornelia Stroh & Aina Urdze. 2006. *On Comitatives and Related Categories: A Typological Study with Special Focus on the Languages of Europe.* Berlin: Mouton de Gruyter.

Szemerényi, Oswald. 1996. *Introduction to Indo-European Linguistics*. Oxford: Oxford University Press.

Thurneysen, Rudolf. 1946. *A Grammar of Old Irish: Revised and Enlarged Edition*. Dublin: Dublin Institute for Advanced Studies.

Villar, Francisco. 1974. *Origen de la flexión nominal indoeuropea*. Madrid: C.S.I.C.

Villar, Francisco. 1981. *Dativo y locativo en el singular de la flexión nominal indoeuropea*. Salamanca: Universidad de Salamanca.

Waanders, Frederik M. J. 1997. *Studies in Local Case Relations in Mycenaean Greek*. Amsterdam: J. C. Gieben.

Wenzel, Heinrich. 1879. *Über den Instrumentalis im Rigveda*. Tübingen: Laupp.

Windisch, Ernst. 1880. *Irische Texte*. Leipzig: Hirzel.

Index