Unity and diversity in grammaticalization scenarios

Edited by

Walter Bisang

Andrej Malchukov

Studies in Diversity Linguistics 99

language
science
press

Studies in Diversity Linguistics

Chief Editor: Martin Haspelmath
Consulting Editors: Fernando Zúñiga, Peter Arkadiev, Ruth Singer, Pilar Valenzuela

In this series:

1. Handschuh, Corinna. A typology of marked-S languages.

2. Rießler, Michael. Adjective attribution.

3. Klamer, Marian (ed.). The Alor-Pantar languages: History and typology.

4. Berghäll, Liisa. A grammar of Mauwake (Papua New Guinea).

5. Wilbur, Joshua. A grammar of Pite Saami.

6. Dahl, Östen. Grammaticalization in the North: Noun phrase morphosyntax in Scandinavian vernaculars.

7. Schackow, Diana. A grammar of Yakkha.

8. Liljegren, Henrik. A grammar of Palula.

9. Shimelman, Aviva. A grammar of Yauyos Quechua.

10. Rudin, Catherine & Bryan James Gordon (eds.). Advances in the study of Siouan languages and linguistics.

11. Kluge, Angela. A grammar of Papuan Malay.

12. Kieviet, Paulus. A grammar of Rapa Nui.

13. Michaud, Alexis. Tone in Yongning Na: Lexical tones and morphotonology.

Unity and diversity in grammaticalization scenarios

Edited by

Walter Bisang

Andrej Malchukov

Freie Universität   Berlin

Contents

Walter Bisang & Andrej Malchukov

Preface

Walter Bisang

Andrej Malchukov
Johannes Gutenberg-Universität Mainz

The present volume originated from the symposium on "Areal patterns of grammaticalization and cross-linguistic variation in grammaticalization scenarios" held on 12-14 March 2015 at Johannes Gutenberg-Universität Mainz. The main purpose of the conference was to bring together leading experts on grammaticalization, combining expertise in grammaticalization theory with expertise in particular language families, in order to explore cross-linguistic variation in grammaticalization scenarios. The participants together with the organizers of the conference (Walter Bisang & Andrej Malchukov) aim at a systematic study of grammaticalization scenarios as well as research on their areal variation, all of this leading to a planned Comparative Handbook of Grammaticalization Scenarios and an accompanying database. Additionally, certain papers which address some of the main questions raised by the organizers of the conference have been invited to the present volume.

Grammaticalization studies and grammaticalization theory have been one of the most successful research paradigms introduced in late 20th century linguistics. The milestone of grammaticalization research includes such works as Lehmann (1995[1982]) on "Thoughts on Grammaticalization", Heine, Claudi & Hünnemeyer (1991) on "Grammaticalization: A Conceptual Framework", Bybee, Perkins & Pagliuca (1994) on "The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World", Heine & Kuteva (2002) on "World Lexicon of Grammaticalization" and Hopper & Traugott (2003) on "Grammaticalization",to name a few. Even critiques of grammaticalization theory (see e.g., Newmeyer 1998, Campbell & Janda 2001; also see Lehmann 2004 for a critical response) did not

stop this research, which numbers in thousands of publications (see the monumental "The Oxford Handbook of Grammaticalization" by Narrog & Heine 2011 for the state of the art in research on grammaticalization).

Yet, in spite of its obvious successes, some aspects remain controversial and are in need of further study. One aspect concerns areal variation in grammaticalization scenarios. Contrary to the alleged universality of grammaticalization processes and paths, grammaticalization shows areal variation, as was most emphatically pointed out by Bisang with particularly telling examples from Southeast Asian languages (Bisang 1996; 2004; 2011; Bisang 2015; also see Ansaldo & Lim 2004). Even though there are many grammaticalization paths in these languages, most of them characteristically diverge from such processes by the absence of the co-evolution of meaning and form as it is generally taken for granted in the literature. Thus, the semantic development of a lexical item into a marker of a grammatical category (e.g., verbs meaning 'give' > benefactive markers) is not necessarily accompanied by phonetic reduction and morphologization (there are phonological properties that operate against the development of bound forms, see Ansaldo & Lim 2004). This lack of form-meaning coevolution in grammaticalization processes in Southeast Asian languages is just one manifestation of areal variation in grammaticalization scenarios which has been underestimated in the literature. Another one is the higher relevance of pragmatic inference as it is manifested in the lack of obligatoriness and in the multifunctionality of grammaticalized markers. Second, the universality of grammaticalization processes has yet to be reconciled with a wide-spread belief that these processes are construction-specific. Given that the constructions in questions are language-specific, it is an open question how one should account for cross-linguistic patterns of grammaticalization. While the construction-specific nature of grammaticalization has been long acknowledged in the literature (Bybee, Perkins & Pagliuca 1994), this aspect came to the fore with the advent of Construction Grammar approaches to grammaticalization (Gisborne & Patten 2011, Traugott & Trousdale 2013). Both aspects noted above raise the issue of how to reconcile universal and language-particular aspects of grammaticalization phenomena. The contributions to this volume address this issue in one way or another.

Perhaps the paper by **Bernd Heine**, **Tania Kuteva** and **Heiko Narrog** on "Back again to the future: How to account for directionality in grammatical change?" addresses this question heads on. Drawing on the material from Khoisan languages but also on comparative data from Germanic, the authors trace the development of future markers. They note that though originally we are dealing with different source constructions including motion verbs, all of them result in a fu-

ture meaning. The answer which the authors give to the puzzle stated above, is that constructional details within or across languages do not preclude universality. They suggest that universality should be formulated in functional (semantic, cognitive) terms as a semantic relation between the source and the target concepts (here the relation between directed spatial movement and the meaning of future). This is a very interesting solution to the problem, even if it is formulated in a rather absolute manner. After all, it is clear that in other cases constructional details would matter, as in the case of 'give'-verbs that develop into benefactive markers in constructions with a verbal host or into a dative marker in constructions with a nominal host. It remains to be seen if the notion of 'host' (from Himmelmann 2004) is sufficient to explain all the divergent paths of grammaticalization. Another solution to the puzzle, which is not at variance with the solution suggested in this paper, is that the constructional details are obliterated as grammaticalization proceeds—a process that instigates convergence between different paths.

The chapter by **Guillaume Jacques** on "The origin of comitative adverbs in Japhug" studies an interesting scenario in a Tibetan language where a proprietive denominal form develops into a comitative marker. The path where a denominal proprietive verb (with a meaning 'having N') in its nonfinite form is reanalyzed as a comitative form of a noun (that is: 'one having branches' > 'with branches') has not been specifically recorded in the literature even though the development from a possessive to a comitative function at a more general level is well documented (for example, for serial verbs). This then provides an additional example of the importance of functional aspects for the explanation of the commonalities of grammaticalization paths, as suggested by the above paper of Heine et al.

Denis Creissels's paper on "Copulas originating from 'see / look' verbs in Mande languages" proposes a new grammaticalization path involving the routinization of an ostensive use of the imperative of 'see' or 'look'. This pathway is not documented in the literature (Heine & Kuteva 2002). Creissels documents this scenario across Mande languages and additionally notes some parallel phenomena in Arabic varieties and in French. French *voici*/*voilà* constitutes a well-known example of how the imperative of verbs with the meaning of 'see' is grammaticalized into an ostensive predicator. In some Arabic varieties the development seems to be mediated through the stage of a modal/discourse particle: The grammaticalization path SEE/LOOK (imperative) > MODAL/DISCURSIVE PARTICLE > COPULA is unusual since it goes partly against intersubjectification as one would expect in the development from copula to discourse particle.

Larry Hyman in his paper on "Multiple argument marking in Bantoid: from

Walter Bisang & Andrej Malchukov

syntheticity to analyticity" shows how to account for the adoption of alternative grammaticalization strategies when a language develops from high syntheticity (agglutination) towards analyticity. The challenge is how to account for the pathway from the inherited head-marking verb structures of Proto-Bantoid to the more analytical structures found in many of the daughter languages. After a careful examination of the data involving valency-changing morphology ("valency extensions") of more conservative Bantu languages and their analytic counterparts in more innovative Bantoid languages, the author raises the question of an ultimate explanation for the move to more analytic structures. Since conventional scenarios, appealing to "erosion" as a byproduct of natural sound change, or else to language contact (in the line of McWhorter's pidginization scenario) seem to be inapplicable here, Hyman suggests that morphology was lost as a result of maximal-size "templatic" constraints on stems. The idea that the shift to analyticity is due to constraints on the number of syllables is highly interesting, but it also raises the question of the factor that ultimately conditioned the templatic constraints. More generally it shows that little is known about the paths of attrition (phonological reduction), no matter whether it is due to erosion or to templatic constraints.

Annie Montaut in her paper on "Grammaticalization of participles and gerunds in Indo-Aryan: preterite, future, infinitive" discusses developments of non-finite forms to finite markers in Indo-Aryan languages. One such path from passive past participle to past tense is well-known, as it is famously responsible for the rise of ergativity. The author however notes that similar developments are documented in the evolution of the passive future/obligative participle and the infinitive into future tense markers in different branches of in Indo-Aryan. Interestingly, the latter developments have not resulted in ergativity. One of the factors accounting for this difference is competition with other forms. As the author shows the resilience of the old future in Indo Aryan languages inhibited the development of gerunds into future markers in some languages of the Western branch. Another factor is analogical influence from other patterns, among them the responsibility of dative subject sentences for the realignment of the gerund construction in Western Indo-Aryan. Thus, competition with other forms and analogical influence can go a long way in explaining variation in grammaticalization paths as well as the alignment of individual verbal forms. This issue is also highlighted by a comparison of Indo-Aryan with Romance. From such a broader perspective, one cannot exclude the existence of general functional constraints in this domain as well (cf. Malchukov & de Hoop 2011 on the TAM-hierarchy for ergativity splits).

Christian Lehmann, one of the founding fathers of grammaticalization research, discusses the topic of "Grammaticalization of tense/aspect/mood marking in Yucatec Mayan". He shows that the formation of preverbal TAM markers is due to the convergence of different constructions, including adverbial modification, complementation based on aspectual or modal verbs, the motion cum purpose construction and the verb-focus construction. Yet, to cite the author, "although the four constructions are clearly distinct, they share a clause-initial position which becomes the melting-pot for the aspectual and modal formatives recruited from different sources". The author's notion of a "melting pot" seems similar to the concept of "attractor position" in the approach of Bisang (1992), even though the latter term has been applied to the typologically rather different languages of East and mainland Southeast Asia. More generally, Lehmann's (and Bisang's) scenario is again in line with the hypothesis that formal reduction is the ultimate explanation of convergence in grammaticalization paths.

Johannes Helmbrecht discusses the grammaticalization of demonstratives in Hoocąk and other Siouan languages. As noted by Helmbrecht, while the evolution of demonstratives into anaphoric pronouns and finally to third person pronouns is well documented, the origin of demonstratives themselves is not well studied. On the basis of comparative Siouan data Helmbrecht shows that the two bound deictic forms *-re* and *-ga* are systematically combined with the three positional verbs *nąk* 'sit', *ąk* 'lie' and *jee* 'stand' in order to form a new paradigm of demonstratives. The verbal origin of these new demonstrative markers can explain why they classify the head noun according to its spatial position (neutral, horizontal, vertical). Other Siouan languages show variation on this theme, but they all have a classificatory demonstrative as an output structure even though the source constructions involving a positional verb are not identical. This situation provides again good evidence for convergent paths.

The final paper by **Björn Wiemer & Ilja Seržant** on "Diachrony and typology of Slavic aspect: What does morphology tell us?" discusses the evolution of aspect in Slavic languages. It is a paper which combines typological and historical approaches trying to trace the origin of the Slavic aspectual system and explain why similar developments have not been attested in other European languages. As a tentative explanation for the renewal of the perfective/imperfective opposition, which in a way continues an older distinction between aorist and imperfect in Proto Indo-European, the authors implicate the substrate influence from Uralic/Altaic in Slavic. While this explanation is tentative it gains credibility, since similar areal explanations have been proposed for other grammatical subsystems. Thus, the preservation of a rich case system in Slavic has been attributed

Walter Bisang & Andrej Malchukov

to an Uralic/Altaic substrate (see Kulikov 2009 for discussion and references).

All the papers presented in this volume provide valuable contributions to the documentation of grammaticalization paths. The authors propose novel grammaticalizations paths not reported in the literature (Helmbrecht, Lehmann, Jacques, Creissels), they offer explanations for the universality and the parametrization of grammaticalization scenarios (Heine et al. from a more general theoretical perspective based on data on the emergence of future markers, Lehmann on TAM marking in Yucatec Maya and Montaut on alignment systems in Indo-Aryan), they provide in-depth analyses of neglected aspects of grammaticalization (Hyman on paths of phonetic attrition), and they explore the role of areal factors and language contact as an explanatory factor of grammaticalization processes (Wiemer & Seržant).

As far as the question of resolving the tension between the construction-specific nature of grammaticalization and the universality of its paths is concerned, there are two answers emerging. One answer, clearly articulated in the article by Heine et al., proposes that universal paths should be formulated in functional/conceptual terms, while the details of the input constructions differ. In addition, several contributions point out that constructional differences become partially opaque in the processes of reduction associated with grammaticalization (most clearly illustrated by Lehmann and Helmbrecht). Hence, we would like to suggest that the convergent trajectory of grammaticalization paths can be partially explained by form-related grammaticalization processes of reduction which blur distinctive properties of individual constructions. We see this as a promising perspective to reconcile the differences between the universal approach and the perspective of Construction Grammar. Future research will show the relative impact of these two explanatory factors for different grammaticalization scenarios, but it is expected that both play a role in later, more systematic explanations.

Walter Bisang & Andrej Malchukov
Mainz, March 2016

Acknowledgements

ation in grammaticalization processes and areal patterns of grammaticalization" (Bi 591/12-1).

References

Ansaldo, U. & L. Lim. 2004. Phonetic absence as syntactic prominence. Grammaticalization in isolating tonal languages. In O. Fischer, M. Norde & H. Perridon (eds.), *Up and down the cline—the nature of grammaticalization*, 345–362. Amsterdam & Philadelphia: Benjamins.

Bisang, W. 1992. *Das Verb im Chinesischen, Hmong, Vietnamesischen, Thai und Khmer. Vergleichende Grammatik im Rahmen der Verbserialisierung, der Grammatikalisierung und der Attraktorpositionen*. Tübingen: Gunter Narr.

Bisang, Walter. 2004. Grammaticalization without coevolution of form and meaning: The case of tense-aspect-modality in East and mainland Southeast Asia. In Walter Bisang, Nikolaus P. Himmelmmann & Björn Wiemer (eds.), *What makes grammaticalization? A look from its fringes and its components*, 109–138. Berlin: Mouton de Gruyter.

Bisang, Walter. 2011. Grammaticalization and typology. In Heiko Narrog & Bernd Heine (eds.), *The Oxford handbook of grammaticalization*, 105–117. Oxford: Oxford University Press.

Bisang, W. 2015. Problems with primary vs. Secondary grammaticalization: The case of East and mainland Southeast Asian languages. *Language Sciences* 47. 132–147.

Bisang, Walter. 1996. Areal typology and grammaticalization: Processes of grammaticalization based on nouns and verbs in East and mainland South East Asian languages. *Studies in Language* 3(20). 519–597.

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.

Campbell, L. & R. Janda. 2001. Introduction: Conceptions of grammaticalization and their problems. *Language Sciences* 23. 93–112.

Gisborne, Nikolas & Amanda Patten. 2011. Construction: Grammar and grammaticalization. In H. Narrog & B. Heine (eds.), *The Oxford handbook of grammaticalization*, 92–104. Oxford: Oxford 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.

Walter Bisang & Andrej Malchukov

Himmelmann, N. P. 2004. Lexicalization and grammaticalization: Opposite or orthogonal? In W. Bisang, N. P. Himmelmann & B. Wiemer (eds.), *What makes grammaticalization? A look from its fringes and its components*, 21–42. Berlin: Mouton de Gruyter.

Hopper, Paul J. & Elizabeth C. Traugott. 2003. *Grammaticalization.* Cambridge: Cambridge University Press.

Kulikov, Leonid. 2009. Evolution of case systems. In A. Malchukov & A. Spencer (eds.), *The Oxford handbook of case*, 439–458. Oxford: Oxford University Press.

Lehmann, Christian. 1995[1982]. *Thoughts on grammaticalization.* Munich: LINCOM EUROPA.

Lehmann, Volkmar. 2004. Grammatication via extending derivation. In Walter Bisang, Nikolaus P. Himmelmann & Björn Wiemer (eds.), *What makes grammaticalization? (A look from its fringes and its components)*, 169–186. Berlin, New York: De Gruyter Mouton.

Malchukov, A. L. & H. de Hoop. 2011. Tense, aspect and mood-based differential object marking. *Lingua* 121. 35–47.

Narrog, H. & B. Heine. 2011. *The Oxford handbook of grammaticalization.* Oxford: Oxford University Press.

Newmeyer, Frederick J. 1998. *Language form and language function.* Cambridge, MA: MIT Press.

Traugott, E. Closs & G. Trousdale. 2013. *Constructionalization and constructional change.* Oxford: Oxford University Press.

Chapter 2

Back again to the future: How to account for directionality in grammatical change?

Bernd Heine

University of Cologne

Tania Kuteva

University of Düsseldorf

Heiko Narrog

Tohoku University

Grammaticalization is commonly understood as a regular and essentially directional process. This generalization appears to be agreed upon in some form or other across many different schools of linguistics, even if it has not gone unchallenged. But there are different views on *what* exactly is regular. Taking the development from movement-based verbs to future tenses as an example, the present paper argues that neither contextual features nor inferential mechanisms, analogy, or constructional form seem to provide a sufficient basis for explaining the evolution of grammatical categories. The paper is based on the one hand on findings made in !Xun, a Southwest African language of the Kx'a family, formerly classified as "Northern Khoisan", and on the other hand on a comparison of this language with observations made in the Germanic languages English, Dutch, and Swedish.

1 Introduction

Grammaticalization is widely defined as a regular and a directional process. This generalization appears to be agreed upon in some form or other across many different schools of linguistics (but see also, e.g., Newmeyer 1998, Norde 2009, and

Bernd Heine, Tania Kuteva & Heiko Narrog

the contributions in *Language Sciences* 23), and for many it is unidirectionality what grammaticalization is about.

There are, however, different views on *what* exactly is regular. Taking the grammaticalization from movement-based verbs to future tenses as an example, the present paper will argue that neither contextual features nor inferential mechanisms, analogy, or constructional form seem to provide a sufficient basis for explaining the directionality in the evolution of grammatical categories. The paper is based on the one hand on findings made in !Xun, a Southwest African language of the Kx'a family, formerly classified as "Northern Khoisan" (Heine & Honken 2010), and on the other hand on a comparison of this language with observations made in the Germanic languages English, Dutch, and Swedish.

The paper is organized as follows. §2 deals with the grammaticalization of a range of future tense categories in the "Khoisan" language !Xun. In §3, the observations made in !Xun are related to findings made on the reconstruction of similar future tenses in three Germanic languages. The implications of this comparison are discussed in §4, and some conclusions are drawn in §5.

There is by now a plethora of definitions of grammaticalization. For the purposes of this paper we will define it as the development from lexical to grammatical forms and from grammatical to even more grammatical forms. And since the development of grammatical forms is not independent of the constructions to which they belong, the study of grammaticalization is also concerned with constructions and with even larger discourse segments (Heine & Kuteva 2002: 2). In accordance with this definition, grammatical developments that do not conform to the definition, such as cases of degrammaticalization, degrammation, desinflectionalization, or debonding (Norde 2009; Norde & Beijering 2014), are not strictly within the scope of grammaticalization theory (see also Ramat 2015: 330).

2 Future tenses in !Xun dialects

The !Xun language, also called Ju, is a traditional hunter-gatherer language of southwestern Africa. The language, classified by Greenberg (1963) as forming the Northern branch of the "Khoisan" family, has recently been re-classified as forming one of the two branches of Kx'a (Heine & Honken 2010), the other branch of this isolate consisting of the ǂ'Amkoe language of Southern Botswana, consisting of the varieties ǂHoan, N!aqriaxe and Sasi (Güldemann 2014).

!Xun is spoken by traditional hunter-gatherers in Namibia, Angola, and Botswana (Heine & König 2015). It is a highly context-dependent language, showing fairly

substantial analytic-isolating morphology; there is only a small pool of items having exclusively grammatical functions (Heine & König 2005). Typological characteristics include the presence of a noun class system with four classes, distinguished in pronominal agreement but not on the noun, and contiguous serial verb constructions. The basic word order is SVO, although there is a minor SOV order, and a modifier-head construction in nominal possession. Sentences in two of its eleven dialects (E3 and W2), though not in others, are divided into two information units separated by a topic marker, where the topical constituent precedes and the non-topical one follows the marker. Phonological features include four click types and four distinct tone levels. The language is divided into eleven dialects, listed in Table 1.

Table 1: A classification of !Xun dialects

Branch	Cluster	Dialect (reference form)
1 Northwestern (NW-!Xun)	1.1 Northern	N1
		N2
	1.2 Western	W1
		W2
		W3
	1.3 Kavango	K
2 Central (C-!Xun)	2.1 Gaub	C1
	2.2 Neitsas	C2
3 Southeastern (SE-!Xun)	3.1 Ju ǀ 'hoan	E1
	3.2 Dikundu	E2
	3.3 ǂx'āō-ǁ'àèn	E3

In his grammar of E1, the best documented !Xun dialect, Dickens (2005: 25) notes: "In Ju ǀ'hoan, the circumstances in which a sentence is spoken often determine its tense, and the verb itself, unlike its English equivalent, is never inflected for time." The only forms that he finds in the dialect to express tense or aspect are the auxiliaries *kȍh* (*koh* in his writing) for past tense and *kú* for the imperfective, and even these auxiliaries are used only optionally. This does not seem to apply to the other dialects (see Heine & König 2015). As Table 2 shows, we found dedicated future tenses in eight of the eleven dialects, and only in two dialects there is none, namely in C2 and E1; for the K dialect there is no information.

Bernd Heine, Tania Kuteva & Heiko Narrog

Table 2: Future tense markers in the !Xun dialects
No information exists on the K dialect of Table 1. Listed in Table 2
are only *dedicated* future tense categories, that is, categories whose
primary function it is to express future tense.

N1	N2	W1	W2	W3	C1	C2	E1	E2	E3
ú, ò- tā	o, ò- tā	g‖è-ā	oā	ōā	o, oga	-	-	ú:	g‖è

There are a number of similarities in the structure of the future tense markers listed in Table 2. First, the markers are throughout placed between the subject and the verb and, second, they are free rather than bound forms. But the markers also differ from one another, in that there are a number of different, or partly different forms.

There are no historical records of the language, but internal reconstruction work by Heine & König (2015) suggests that no conventionalized future tense form or construction can be traced back to Proto-!Xun, the hypothetical ancestor of the dialects. But there are two verbs, namely *ú 'go' *g‖è 'come', which can. The only reasonable hypothesis is that these verbs were there earlier than the future tense markers and that the former must have been involved in the historical development from the former to the latter. On this analysis, at least eight of the eleven dialects of the language appear in fact to have developed movement-based future tenses. Four dialects transparently used the verb *ú 'go', developing what following Dahl (2000) we call a de-andative future. Two other dialects apparently used the verb *g‖è 'come', creating a de-venitive future in Dahl's terminology; we will return to this below.

But the constructions were not the same in the dialects. While all involved a sequence of two verbs, V_1 and V_2, three different constructions can be distinguished on the basis of their morphosyntactic behavior, which we will refer to with the terms in (1).

(1)   Morphosyntactic types of future categories
   a.   Complement-based
   b.   Serializing
   c.   Particle-based

In *complement-based* futures, the future marker consists of a movement verb (V_1) meaning 'go' or 'come' plus the transitive suffix -*ā* (glossed 'T'). This suffix,

which turns, e.g., intransitive verbs into transitive ones, serves to add a comple-ment to the valency of the verb.[1] Such a complement can be a noun phrase (cf. (3), an adverbial phrase, or a complement verb, as in (2), and the second verb (V_2) behaves structurally like a complement of V_1.[2] Thus, the meaning of (2) can structurally be rendered as '(S)he doesn't go to the coming', where the movement verb (V_1), *ú* 'go', is ambiguous in that it has future tense as its second reading (un-less indicated otherwise. The examples presented below are taken from Heine & König 2015).

(2) N1 dialect (Southeastern Angola)
     yà ǀōā ú-   á tcí.
     N1 NEG go/FUT- T come
     'He will not come.'

In *serializing* futures, the two verbs V_1 and V_2 are simply juxtaposed, cf. (3) and (4), as they are in the serial verb construction of the language (König 2010; cf. Bisang 1998; 2010)

(3) E2 dialect (Northeastern Namibia)
     m̄i ú: gè- à Tàmzó.
     1.SG go/FUT stay- T Tamzo
     'I am going to stay in Tamzo.'

(4) E3 dialect (Eastern Namibia, western Botswana)
     mí m̄ (kú) gǀè   kx'āè kā.
     1.SG TOP PROG come/FUT get   N4
     'I'll have it.'

In *particle-based* futures, the future marker consists of an element that is seem-ingly etymologically opaque. Examples are provided by the markers *ò-tā* in (5), *oga* in (6),[3] and *óá* in (7).

---

[1] The suffix, glossed 'T', is called the "transitive suffix" by Dickens (2005: 37-38).

[2] Note that verbs in !Xun can typically be used in nominal slots, whereas nouns cannot be used as verbs.

[3] The only data available on the C1 dialect stem from Vedder (1910-1911), who has no consistent tone markings and frequently confounds voiceless and voiced consonants. Thus, *oga* presum-ably is phonetically [*oka*]. Furthermore, he gives the meaning of *gǀyee* as 'go', which is most likely a mistake and we have tentatively changed it to 'come' on the basis of strong evidence from the other ten dialects.

Bernd Heine, Tania Kuteva & Heiko Narrog

(5) N1 dialect (Southern Angola)

m̄   txòm, à   ò-tā ǁé   […].
1.SG uncle 2.SG FUT die.SG

'My uncle, you are going to die […].' (The tale of the lion and the jackal; Heine & König 2015)

(6) C1 dialect (North-central Namibia)

na   tí   oga gǀyee.
1.SG ICPL FUT come

'I'll come.' (Vedder 1910-1911: 20)

(7) W2 dialect (Northern Namibia)

hä̀ má nǁä̀   n   óà   gǀè.
N1 TOP later FUT come

'He'll come later.' (Own data)

On the basis of the dialect comparisons carried out by Heine & König (2015) it is possible to reconstruct these three particles. First note that there is general vowel lowering in the dialects whereby *u* tends to be lowered to *o* when there is a non-high vowel in the following syllable, hence *u > o*. The particles *ò-tā* and *oga* can be reconstructed back, respectively, to the sequences *ú tà* and *ú kà*, both meaning 'go and' (see §2.1). Finally, the particle *óá* can be reconstructed to the combination *ú-ā*, that is, 'go' plus the transitive suffix introducing a complement. Table 3 lists the various future tense markers and their reconstructed forms.

Table 3: Future tense markers in the !Xun dialects and corresponding reconstructed forms (cf. Heine & König 2015)

N1	N2	W1	W2	W3	C1	C2	E1	E2	E3
ú, ò-tā	o, ò-tā	gǀè-ā	óá	ōā	o, oga	-	-	ú:	gǀè
*ú, *ú tà	*ú, *ú tà	*gǀè-ā	*ú-ā	*ú-ā	*ú, *ú kà			*ú	*gǀè

2.1 Accounting for the future tenses

We observed in (1) that the future tense constructions in the !Xun dialects appear to be built on three different constructions which we referred to, respectively, as

the complement-based, the serializing, and the particle-based types. Now, there are three main constructions in the dialects used to connect two verbs or verb phrases, illustrated in (8) with examples from the W2 dialect. In the complementation construction of (8)a, V_1 is the main verb and V_2 is introduced as its complement. If V_1 is an intransitive verb, as *ú* 'go' in (8)a is, it takes the transitive suffix *-ā*, otherwise there is no formal marking. In verb serialization no formal marking is needed, as (8)b shows: V_1 and V_2 are simply juxtaposed and any complement that V_1 may have follows V_2.

Coordination, by contrast, uses either of the additive conjunctions *tà (*tè in the Southeastern dialects) and *kà 'and', as we saw already in §2.1, cf. (8)c. The functions of these conjunctions are not exactly the same: Whereas the former conjoins separate events, the latter typically conjoins events that are conceived as wholes (Heine & König 2015: 320).

(8) W2 dialect (own data)

    a. *hä má kè ú- á m̄ ..*   [Complementation]
        3.SG TOP PAST go- T eat
        'He went to eat.'

    b. *hä má kè ú m̄ ..*   [Verb serialization]
        3.SG TOP PAST go eat
        'He ate while going.'

    c. *hä má kè ú kā m̄.*[4]   [Coordination]
        3.SG TOP PAST go and eat
        'He went and ate.'

The three constructions illustrated in (8) do not all express the same meaning but are available to speakers as different options to connect verbs or verb phrases. And all the future tense constructions discussed in §2.1 can be traced back to them.

Thus, example (2) above is suggestive of complementation, and so is example (7) from W2, where the future tense marker *óá* can be reconstructed back to a combination of *ú 'go' plus the transitive suffix *-ā. (3) and (4), on the other hand, are instances of the verb serialization construction, consisting of two verbs following one another without any formal linkage. In fact, both are ambiguous between a serial lexical and a grammatical interpretation: Thus, *ú: gè* in (3) can

---

[4] Instead of *kā* 'and', a much more common coordinating conjunction is *tà* and its equivalents in other dialects.

Bernd Heine, Tania Kuteva & Heiko Narrog

mean either 'go (and) stay' or 'will stay' and, similarly, *g/è kx'āè* in (4) can be translated variously as 'come (and) get' or 'will get'. The collocations *ò-tāllé* 'will die' in (5) and *oga g/yee* 'will come' in (6), by contrast, can be reconstructed back, respectively, to the coordination construction of Proto-!Xun (*ú tāllé* 'go and die', *ú kā g/è* 'go and come', respectively).

To conclude, there appear to have been three different highly schematic constructions involving altogether six partially schematic constructions that developed in the same direction towards future tense constructions, namely [*g/è + V], [*g/è-ā + V], [*ú + V], [*ú-ā + V], [*ú kà + V], and [*ú tà + V] (see Table 3). Note further that in some of the dialects (N1, N2 and C1) there are two different source constructions leading to the same target, namely a future tense construction.

To be sure, these constructions could be argued to have involved a general schema [$V_1 + V_2$], but their morphosyntax was different, both on a schematic and a more substantive level. The question then is: How is this situation in the !Xun dialects to be explained, that is, what was responsible for this diversity in source constructions? Shared genetic origin is unlikely to account for this situation, with one possible exception: The markers *ò-tā* in N1 and N2 and *óá* and *ōā* in W2 and W3, respectively, may each be due to a shared ancestor within the respective dialect group. But overall, these future constructions cannot be traced back to one common construction in the proto-language.

It would seem that there is only reasonable answer to this question, namely with reference to the meaning of the source and the target constructions. What they all have in common is that there was a verb expressing deictic movement and belonging to the basic vocabulary in the sense of Swadesh (1952), and that in present-day !Xun there is a construction whose main function it is to express future tense. The result was, in the terminology of Dahl (2000), either a *de-andative* or a *de-venitive* future depending on whether the movement verb was 'go (to)' or 'come (to)'.

It goes without saying that the overall process is more complex. For example, the source construction may also give rise to other target constructions, and future tense may only be one of the functions expressed. But in accordance with the definition of grammaticalization used here (see §1), our interest is exclusively with this one pathway of change, ignoring the wealth of possible alternative constructional histories.

On this view, which is in accordance with the framework of Heine, Claudi & Hünnemeyer (1991), there is some fixed semantic relation between source concepts for 'go' and 'come' and the grammatical target concept of future tense in specific contexts. What this seems to entail is the following hypothesis:

(9)   Compared to semantic features, other factors that are likely to be involved are of secondary import in the development from lexical to grammatical material. An explanation of this development must therefore be over and above meaning-based.

Note, however, that !Xun is a language for which no historical records are available, thus making detailed diachronic reconstruction impossible and a falsification of the hypothesis difficult. We will now test the hypothesis in (9) with data from Germanic movement-based future tenses, for which arguably the best descriptions are available.

3 Future tenses in English, Dutch and Swedish

The account presented in this section is by no means meant to do justice to the grammaticalization of the three Germanic future tenses based on movement verbs; rather, our interest is restricted to testing the hypothesis in (9). The account is based on the collostructional, distinctive collexeme analysis by Hilpert (2008). Unlike what we observed in !Xun, the constructional format to be found in all three languages is essentially the same (but see §4), involving what we referred to in §2.1 as the complementation construction: The movement verb (V_1) of the source construction is the main verb and its complement contains a non-finite verb (V_2), turning via grammaticalization into the new main verb; hence, the constructional change underlying all grammaticalizations to be discussed can be rendered as leading from (10)a to (10)b.

(10)   a.   [main verb V_1 - non-finite complement verb V_2]
        b.   [future tense auxiliary - main verb]

Following Hilpert (2008), our main concern is with the constructional context of the tense categories.

3.1 The de-andative English *be going to*-future

The first example concerns the evolution of the English *be going to*-future, a de-andative future in the terminology of Dahl (2000). The grammaticalization of this evolution has been extensively studied (see Hopper & Traugott 2003; Mair 2004; Hilpert 2008 and the references therein; see also Disney 2009). It seems to be well established that the construction was fully grammaticalized in the Early Modern English period by the end of the 17th century or the mid 18th century,

and that a drastic increase in its text frequency first occurred in the 19th and early 20th centuries. Note that according to Mair (Mair 2004: 129; Mair 2011: 244–245), the increase of frequency is the outcome, not the driving force of the *be going to*-future.

In his corpus-based collostructional study, Hilpert (2008: 118-121) analyzes the following three stages of this de-andative future: 1710-1780 (let us call it period 1), 1780-1850 (period 2), and 1850-1920 (period 3). During period 1, the construction strongly harmonized with telic and dynamic verbs, and all distinctive collexemes select for animate, intentional subject referents.

During period 2, it is still telic and dynamic verbs that the construction harmonizes with, most elements being compatible with an intentional reading, *be* and *have* now are among the most frequently used complement verbs. However, there are now also inanimate subjects that exclude an intentional interpretation but rather signal imminent future events, like in (11):

(11) English (between 1770 and 1820; Hilpert 2008: 120)
In the true sleepy tone of a Scottish matron when ten o'clock is going to strike.

During period 3, there appear to be hardly any lexical restrictions. The verb *happen* now belongs to the ten most frequent complement verbs, and unintentional complement verbs are fully acceptable. Hilpert (2008: 121) concludes that "the occurrence of spontaneous, non-intended events is only encoded by *be going to* in later stages of its development".

3.2 The de-andative Dutch *gaan*-future

On the basis of the data available, Hilpert (2008: 113) classifies the history of this de-andative future into three periods of time: centuries 16-17, 18-19, and 20, let us refer to them as periods 1, 2, and 3, respectively.

During period 1, Hilpert found all distinctive collexemes of this period to share an "atelic aspectual character". The collexemes encode events involving intentional movement of an animate agent. The events expressed commonly involved literal and intentional motion, associated with atelic situation types.

During period 2, most of the distinctive collexemes have the "telic aspectual contour of accomplishment verbs". There are on the one hand also intentional actions of human agents, but on the other hand also unintended processes such as *sterven* 'die'. The constructional meaning "is now broadening to accommodate events that are not connected to the intentions of human agents".

In period 3, the new verbal complements (distinctive collexemes) are again mostly atelic. The future meaning of *gaan* is fully conventionalized, combining also with verbs denoting involuntary human activities. And now, *gaan* can also combine with inanimate subjects, as in (12):

(12) Dutch (Hilpert 2008: 117)
      Wat gaat er   dan gebeuren, Sander?
      what goes there then happen   Sander

      'What is going to happen then, Sander?'

3.3 The de-venitive Swedish *komma att* future

De-venitive futures concern source constructions involving 'come (to)' as the matrix verb. The following is a sketch of the grammaticalization of the Swedish *komma att*-future construction based on the collostructional, distinctive collexeme analysis by Hilpert (2008: 125-131). Hilpert distinguishes three diachronic stages in the development of the construction, we will refer to them as period 1 (centuries 16-18), period 2 (century 19), and period 3 (century 20).

In the earliest documented records of period 1, the most distinctive verbs describe non-agentive human activities and involuntary reactions. Verbs, such as *förakta* 'despise', *sova* 'sleep', and *rodna* 'blush', and *höra* 'hear' describe activities carried out unintentionally, but have animate subject referents, e.g., *sova* 'sleep', *höra* 'hear' (Hilpert 2008: 128).

It is only in period 2 that typically intentional activities can felicitously combine with the matrix verb *komma att*, such verbs being, e.g., *klara* 'manage' or *skicka* 'send', and the frequency of animate subject referents increases, but in this period there are also examples of future events that are beyond the control of the subject referent, thus expressing predictions about future events.

In period 3, a common pattern consists in the use of atelic and stative verbs, and the *komma att*-construction "can express a plain sense of prediction", but also "timeless generic truths that are epistemic rather than modal" (Hilpert 2008: 130).

3.4 The futures compared

The following is not meant to be an evaluation of different linguistic models, nor does it aim at a comprehensive treatment of this subject (for which see Börjars & Vincent 2011); rather, it is restricted to the following questions:

Bernd Heine, Tania Kuteva & Heiko Narrog

(13) a. Does the framework account for the regularities of change in the development of future tense categories?

   b. Does the framework propose a reasonable explanation for unidirectionality?

Both English and Dutch have a de-andative future, historically derived from an auxiliary construction involving a verb for 'go (to)', but the evolution of the two futures was clearly different. Hilpert (2008: 122) summarizes the differences thus: "Converse preferences for perfectivity, transitivity, and agentivity can be shown to permeate their respective developments. A historical perspective on the shifting collocational preferences of the two constructions reveals that *be going to* had a special affinity towards speech act verbs, while with *gaan*, movement verbs had a special role". Central to the development of English *be going to* were in fact perfective speech act verbs.

In its early stages, Dutch *gaan* commonly occurred with typically imperfective movement verbs, and it expressed intentional movement. In later usage, the construction accommodates verbs without the meaning of movement and intentionality. This contrasting genesis is to quite some extent reflected in the present situation. English *be going to* attracts verbs that are transitive, punctual, and highly agentive (Hilpert 2008: 121-122). Dutch *gaan*, by contrast, attracts verbal complements that are intransitive, temporally extended, and non-agentive; intention is not (i.e., no longer) a part of its constructional semantics.

Table **??** deals with some lines of semantic development in the movement ('go'-) verbs, while Table 5 summarizes the corresponding developments in the verbal complements of the two future tenses.[5] As these data suggest, there is no difference in the former but dramatic differences in the latter developments; we will return to this issue below.

4 What is directional in the evolution of future tenses?

That grammaticalization is essentially (though not entirely) unidirectional, or that there is asymmetry between what is and what is not directional (Börjars & Vincent 2011), is a generalization that appears to be agreed upon in some form or other across different schools of linguistics (but see also e.g. Newmeyer 1998; Norde 2009, and the contributions in *Language Sciences* 23), and for many, it is unidirectionality that grammaticalization is about.

---

[5] The information on the Swedish *komma att*-construction is incomplete and therefore not listed in these tables.

Table 4: Major semantic developments of the matrix (motion) verbs in two de-andative future tenses of Germanic languages (based on Hilpert 2008: 116-123)

	English *be going to*	Dutch *gaan*
Early usage	Movement, intention	Movement, intention
Present usage, earlier phase	- Movement, +/-Intention	- Movement, +/- Intention
Present usage, later phase	- Intention	- Intention

Table 5: Major semantic developments of the verbal complements in two de-andative future tenses of Germanic languages (based on Hilpert 2008: 116-123)

	English *be going to*	Dutch *gaan*
Early usage	Common with perfective speech act verbs	Common with imperfective movement verbs
Present usage, earlier phase	transitive, punctual, and highly agentive verbs	intransitive, temporally extended, and non-agentive verbs

The evolution of de-andative and de-venitive futures has been described as one that is in accordance with the unidirectionality hypothesis. No case has so far been reported where a future tense gave rise to a lexical verb meaning 'go' or 'come' while the opposite development is well documented ever since it was first discussed in detail by Bybee and associates (Bybee, Pagliuca & Perkins 1991; 1994). But in the constructional history of such categories there are many linguistic, pragmatic, and sociolinguistic factors involved. The question then is: What is it in this history that is in fact directional?

In §4.1 we will look at some factors that have been argued to show directionality in grammaticalization but do not seem to be uncontroversial. In §4.2 then we will endeavor to isolate phenomena that, at least on the basis of the data discussed in Sections §2 and §3, appear to go in one direction. In addition we will then look into the question of how to account for directionality.

Bernd Heine, Tania Kuteva & Heiko Narrog

4.1 What is not directional?

A number of factors and theoretical concepts have been invoked to account for the kinds of grammaticalizations discussed in §3, yet which on closer look raise some questions. We will now look at them in turn.

4.1.1 Constructions

One of the theoretical concepts that has more recently been discussed in detail concerns the morphosyntactic format of the constructions involved in grammaticalization: Does the grammaticalization of future tenses require a specific constructional format to take place?

It would seem that the answer is in the negative. We noticed that in the dialects of !Xun it was not one type of construction that was responsible for the rise of future tenses but rather three. This is different in the case of the three Germanic futures dealt with in §3. But even here there appear to be striking differences between the languages examined, as Table 6 shows. Whereas the English *be going to-* and the Swedish *komma att*-constructions introduce the verbal complement by means of a preposition, there is no preposition in the Dutch construction. And whereas English requires the verb to be constructed in the progressive aspect, this is not a requirement in many other languages.

Table 6: The constructional form of source constructions for movement-based futures

Language	'Go' as the matrix verb	Use in progressive aspect	Prepositional complement
E. *be going to*	+	+	+
D. *gaan*	+	-	-
S. *komma att*	-	-	+

Furthermore, in a number of other languages there are construction types that differ dramatically from the ones to be found in the languages examined here. For example, rather than an infinitival or other non-finite complement verb there is a finite verb that serves as the complement of the movement verb, as the following example from the Pipil language of Guatemala shows (Campbell 1987).

14

(14) Pipil (Aztecan, Uto-Aztecan; Campbell 1987: 268)

ti-	**yu**- t	ti-	yawi- t	ti-	pa:xa:lua- t	ne:pa ka ku:htan.
we-	go- PL	we-	go- PL	we-	walk- PL	there in woods

'We are going to go take a walk there in (the) woods.'

To conclude, which morphosyntactic form a construction takes does not seem to be a factor that determines directionality.

4.1.2 Context

Another structural feature concerns the context frame: Does the "same" grammaticalization process occurring in different languages involve the same kind of context?

The example of the English and Dutch de-andative futures, or of the Swedish de-venitive *komma att*-future suggest that the answer is again in the negative: As we saw in Table 5 of §3.4, the de-andative futures of English and Dutch drew on highly contrasting kinds of contexts (that is, complement verbs). Nevertheless, the end product was essentially the same, namely the schematic grammatical function 'future'. Context change can even show a reversal of directionality. For example, Dutch *gaan* occurred at the first stage (period 1, century 16-17th century) in the context of atelic complement verbs, changing to telic verbs in period 2 (18-19th century). In period 3 (20th century) finally, there was another move back to atelic verbs (Hilpert 2008: 116-7).

Thus, there does not appear to be clear evidence that directionality is necessarily determined by the nature of the contextual features involved.[6]

4.1.3 Inferential mechanism

Much the same appears to apply to a number of semantic features associated with grammaticalization: The analysis of movement-based futures suggests that not all semantic changes in the development of movement-based futures are unidirectional.

One of them concerns the inferential mechanism involved. According to one position surfacing implicitly or explicitly in the relevant literature – one that

---

[6] Andrej Malchukov (p.c.) rightly asks whether grammaticalization does not always involve context expansion; note that context extension constitutes one of the four parameters of grammaticalization in the framework of Heine & Kuteva (2002: 2). According to that framework, context extension is a necessary but not a sufficient condition for grammaticalization to take place; what is required in addition is at least also desemanticization.

can be traced back to Bybee, Perkins & Pagliuca (1994b), it is the nature of the inferential pathway leading from source to target concept that is crucial in grammaticalization, rather than a "macro-shift" from source to target. This pathway is said to be not only responsible for regularities in grammatical change but also for directionality (Bybee, Perkins & Pagliuca 1994b: 268).

Depending on which aspect of the pathway one has in mind, this position must remain controversial. Take the example of the English and Dutch de-andative futures that we presented in §3. They are suggestive of an inferential pathway leading from physical motion via intentional action to prediction (i.e., future tense), as sketched in (15)a. But this is not the only pathway that has been identified. There is an alternative pathway for de-venitive futures (involving verbs meaning 'come (to)') that does *not* involve intention, leading from directed motion via the aspectual notion inchoative to prediction (Dahl 2000: 322; Hilpert 2008: 126). Thus, in addition to (15)a, there is also (15)b.[7]

(15) Inferential mechanisms in the development of motion-based Germanic future tenses (Hilpert 2008: 126, 183)

    a. Directed motion > intention > future tense (English, Dutch)

    b. Directed motion > inchoative > future tense (Swedish)

To conclude, there does not appear to be a regular inferential mechanism leading from motion to prediction; rather, there may be different pathways involved. More specifically, intentionality does not appear to be crucial for movement-based future tenses to arise.

4.1.4 Intentionality

More specifically, intentionality is a concept that has been invoked in a number of grammaticalization studies to account for regular grammatical change, most of all for changes leading to future tense markers. For Bybee, Perkins & Pagliuca (1994b: 254), "all futures go through a stage of functioning to express the intention, first the speaker, and later the agent of the main verb", and this hypothesis was adopted by Heine (1995; see also Ultan 1978).

It would seem, however, that this hypothesis has to be abandoned on the basis of observations such as the following from movement-based future tenses. These observations suggest not only that intentionality is not necessarily involved in

---

[7] For volition-based future tenses, Bybee, Perkins & Pagliuca (1994b: 256) propose the following pathway:DESIRE > WILLINGNESS > INTENTION > PREDICTION.

movement-based futures, as we just saw. On the contrary, it can also be at variance with the unidirectionality hypothesis. In the earliest documented records of period 1 of the Swedish *komma att*-future, the most distinctive verbs describe involuntary reactions: Verbs such as *förakta* 'despise', *sova* 'sleep', and *rodna* 'blush', and *höra* 'hear' select animate subject referents but describe activities carried out unintentionally. It is only at a second stage, in period 2, that typically intentional activities can felicitously combine with the matrix verb *komma att* (Hilpert 2008: 128, 131).

In the development of the de-andative futures of English and Dutch, by contrast, there was an opposite directionality from intentional participants to loss of intentionality as a distinctive feature. Thus, in the English *be going to*-future, all distinctive collexemes selected animate, intentional subject referents in the earliest period 1 (1710-1780). During period 2 (1780-1850) there are now also inanimate subjects that exclude an intentional interpretation and in period 3 (1850-1920), unintentional complement verbs are now fully acceptable (Hilpert 2008: 121).

Much the same development from intentional to unintentional events can be observed in the Dutch *gaan*-future. During period 1 (16-17th centuries), the collexemes encode events involving intentional movement of an animate agent: The events expressed commonly involve literal and intentional motion. During period 2 (18-19th centuries), there are on the one hand also intentional actions of human agents, but on the other hand also typically unintended processes such as *sterven* 'die'. The constructional meaning "is now broadening to accommodate events that are not connected to the intentions of human agents" (Hilpert 2008: 116). In period 3 (20th century), *gaan* can also combine with inanimate subjects, incapable of intentional actions.

Intentionality is closely related to agentivity and, in fact, what has been said about the former also applies in some way or other to the latter. For example, it has been argued that in some pathways of grammaticalization, concepts for willful, agentive participants are transferred to also denote inanimate concepts and a body of evidence has been presented for this hypothesis (Heine, Claudi & Hünnemeyer 1991; Heine 1997). As the data in §3 suggest, however, this not a requirement for the development of movement-based future tenses: In the earliest documented records of period 1 of the Swedish *komma att*-future, the most distinctive verbs describe non-agentive human activities and involuntary reactions.

In sum, neither of the concepts intentionality and agentivity necessarily behaves directionally: There can be a change from intentional to unintentional activities (cf. the English and Dutch de-andative futures) but also from unintentional to intentional activities (cf. the Swedish de-venitive future). And changes

Bernd Heine, Tania Kuteva & Heiko Narrog

do not necessarily lead from agentive to non-agentive subjects.

4.1.5 Telicity

And much the same as intentionality concerns telicity and the aspectual contours of verbs or events. The Dutch *gaan*-future was associated with atelic verbs in the 16th and 17th centuries: "all distinctive collexemes of this period share an atelic aspectual character" (Hilpert 2008: 116). This situation changed substantially in the 18th and 19th centuries, when most of the distinctive collexemes had the telic aspectual contour of accomplishment verbs. Finally, in the 20th century, the distinctive collexemes are again mostly atelic (Hilpert 2008: 117). Thus, there appears to be a bidirectional development from atelic to telic on the one hand from telic to atelic verbal events on the other.

Assuming that these are not idiosyncratic, exceptional examples, they show that not all semantic changes in grammaticalization are directional.

4.1.6 Analogy

In a recent study, Fischer (2013) proposed an explanatory account for the English *be going to*-future tense in terms of analogy. She hypothesizes that it was similarity, or structural analogy on the morpho-syntactic level that played a central role in the development of this tense construction.[8] There was a change in *going* from lexical verb to auxiliary and the spread of infinitives from expressing concrete movement to also expressing mental activities, and next also to subjects that were inanimate or empty rather than animate and agentive. The role played by analogy was that, once there is an auxiliary construction that could behave like an [AUX - V] pattern it "will attract constructions (with different kinds of infinitives/subjects that are in use after other, (functionally) similar [AUX - V] patterns, such as *shall/will* + infinitive" (Fischer 2013: 522).

Fischer (2013) appears to favor a perspective according to which analogy is less about what speakers do than about what they do *not* do. She argues that in analogy one "treats something like something else because one does not spot any difference, so it is a negative force rather than a positive one" (Fischer 2013: 519).

Analogy has been invoked in quite different frameworks dealing with grammaticalization, including generative ones (e.g., Kiparsky 2012) and functional

---

[8] Olga Fischer (p.c.) emphasizes that analogy in the sense of the term used by her includes in the same way the meaning, pragmatics, and the form of the construction concerned.

ones (e.g., Hopper & Traugott 2003: 39-40). For the latter, analogy effects (linguistic or sociolinguistic) rule spread rather than "rule change" - in other words, analogy presupposes "reanalysis" in grammaticalization. For example, the grammaticalization of the Old English noun *had* 'person, condition, rank' into a derivative morpheme representing an abstract state (e.g. *biscophad* 'bishophood' is said to have involved two instances of reanalysis: (a) compounding followed by (b) semantic and morphological change. Thus, the development from nominal to derivative morpheme was due to "reanalysis." Analogy subsequently had the effect that the derivative morpheme no longer required association with a word referring to a person but rather could be extended to new contexts, giving rise to Modern English expressions such as *falsehood.*

Analogy is a ubiquitous mechanism, influencing all kinds of grammatical processes, including the present as well as others described in detail by Fischer (2013). But if taken as the main factor to account for the development then this raises questions. With reference to directionality, this raises questions such as the following:

(a) Why is there a development from lexical to [AUX-V] pattern – why should there not be a development in the opposite direction? In other words, what accounts for the grammaticalization from lexical verb to auxiliary (AUX)?

(b) Is there reason to rule out the possibility that analogy may not also work in the opposite direction, namely leading from the pattern [AUX-V] to another pattern [main verb-infinitival complement] – a pattern where English would have offered a plethora of models?

(c) Finally, and most importantly, why should analogy be directional – would there be any more general motivation? Fischer (2013: 521) proposes processing errors as playing an important role in analogical processes. The question then is why the same kind of grammaticalization from a pattern [main verb 'go' + non-finite verbal complement] to [AUX-V], to be observed in many languages across the globe, should have involved the same process, considering that not all of these languages disposed of a pattern such as English [shall/will + infinitive].

To conclude, analogy is an important factor in all kinds of grammatical change, but it does not seem to account for the kind of directionality to be observed crosslinguistically in the grammaticalization from a lexical verb of goal-oriented physical motion to future tense marker. Accordingly, rather than as a unidirectional process, Fischer (2000: 153) views grammaticalization as "a more or less accidental concurrence" that "may lead one way as well as another." Note that her interest appears to be not with crosslinguistic typological generalizations but primarily with understanding the history of English and other Germanic languages.

Bernd Heine, Tania Kuteva & Heiko Narrog

Thus, analogy in the way proposed by Fischer can be an important trigger but does not seem to be responsible for the directionality to be observed in grammaticalization (but see also Kiparsky 2012 for a different concept of analogy).

4.1.7 Frequency

Frequency of use as an explanatory notion is invoked most of all in usage-based approaches (e.g., Bybee 2011; Torres Cacoullos & Walker 2011: 225). For Bybee and associates, high frequency of use of linguistic phenomena appears to be criterial for grammaticalization to happen (Bybee 2003; 2006): "Thus as long as frequency is on the rise, changes will move in a consistent direction" (Bybee 2011: 77).

While frequency is in fact an important factor, it would seem that more evidence is needed to establish that frequency by itself can immediately be causally responsible for the presence of new functional categories.[9] Furthermore, one wishes to know what accounts for increased frequency, that is, *why* do interlocutors use certain linguistic expressions more frequently than others – in other words, frequency may tell us little about why people use their languages the way they do.

Is frequency really responsible for directionality – e.g., to the effect that the more frequently a linguistic expression is used the more it will be grammaticalized? It would seem that this question cannot be clearly answered in the affirmative. First, there is linguistic material that is used highly frequently but does not appear to be grammaticalized. This is suggested on the one hand by frequency counts of lexical items, some of which occur highly frequently in texts but may show little effects of grammaticalization. Second, that there is no one-to-one relationship between frequency and grammaticalization can be seen in developments where some grammatical element experiences a decrease in its frequency of use but no corresponding decrease in its grammaticalization. And third, there are some research findings suggesting that the contribution of frequency to grammaticalization is not entirely uncontroversial (Hoffmann 2004; 2005; Brems 2007; Mair 2011; Hilpert 2013: 10). As we saw in §3.1, the dramatic increase in text frequency that the English *be going to*-construction experienced in the early 20th century is shown by Mair (2004: 129) to be the outcome rather than the driving

---

[9] We are concerned here *exclusively* with frequency in the rise of a new functional category. The situation is different in subsequent developments of such a category. Note further that a distinction must be made between frequency of the element that provides the source of grammaticalization and that of later uses of this element, as well as between type and token frequency (cf. Mair 2011: 244).

force of grammaticalization.

On account of such observations one may hesitate to hold frequency of use responsible for directionality in grammaticalization.

The catalog of factors discussed above is far from exhaustive. What it suggests, however, is that many of the hypotheses that have been volunteered must be taken with care. The question then is what is it ultimately that makes grammaticalization an essentially unidirectional process? This is the subject of the next section.

4.2 What is directional?

We saw in the preceding section that a number of the factors that characterize the history of the future tense categories in the Germanic languages surveyed do not seem to be directly responsible for the directionality to be observed in the grammaticalization of these categories. Such changes are either not directional in that they may go in both directions of a chain of grammaticalization or else their contribution to the process is not entirely clear.

It would seem that there is essentially only one factor that can be identified both in the !Xun dialects and in the Germanic futures, as proposed in our discussion of §3.4, namely the shift from lexical (or less grammatical) to grammatical meaning, entailing a gradual transformation of lexical as grammatical morphosyntax. The latter process has received considerable scholarly attention (e.g., Lehmann 1982; Heine, Claudi & Hünnemeyer 1991; Bybee, Perkins & Pagliuca 1994b, and subsequent works), being described as one of structural (morphosyntactic and morphophonological) reduction; we will return to this issue below.

What all cases examined in this paper in fact share is that there appears to be a fixed semantic relation between source concepts for 'go' and 'come' and the grammatical target concept of future tense in the languages concerned, in accordance with our hypothesis in (9). This relationship implies a "macro-shift" of the kind discussed in this paper. Such a shift can, but need not, take place in virtually any language, and it can be arrested at any point in history, that is, it may be, and not seldom is incomplete – in other words, the grammaticalization process need not take its full course. In the latter case there is only a weakly grammaticalized future tense.

To be sure, in the case of the !Xun dialects there may also have been some kind of drift effect in the sense of Sapir (1921) that contributed to the fact that in eight of the ten documented dialects a movement-based future tense arose. But this does not account for the hundreds of other languages in Africa and elsewhere where a similar development took place. And, as far as the information available suggests,

clearly the most perspicuous common denominator of all these developments is the source-target relationship between deictic movement verbs in combination with another verb and the grammatical function of future tense. As we saw in §2, this combination need not be one of a matrix verb and its complement, it can as well be one of coordination or verb serialization, or even of clause subordination.

Considering that this development tends to require an extended period of time, possibly involving various intermediate stages and constructional changes, this generalization raises the question of what the underlying causal factors are that can be held responsible for this relationship. We have no clear answer to this question, which is in need of much further research work. But there are a few suggestions made in works on this subject matter which may be of help in such work. According to Hilpert (2008: 109) there is an implicature inherent in the meaning of directed movement whereby the content of the verbal complement of 'go (to)' and 'come (to)' implies a situation in time later than reference time, thereby enabling a "presupposition of a future event". In a similar fashion, Bybee, Perkins & Pagliuca (1994b: 268) suggest that the temporal meaning that comes to dominate the semantics of the construction "is already present as an inference from the spatial meaning. When one moves along a path toward a goal in space, one also moves in time." This is, for example, also in accordance with what is possibly one of the earliest uses of the English *be going to*-future in (16), which provides a possible context for a future interpretation.

(16) English
    *ther passed a theef byfore Alexandre that **was goying to be hanged** whiche saide …*

    'a thief who was going to be hanged passed before Alexander and said …
    (1477, Mubasshir ibn Fatik, Abu al-Wafa'; *Dictes or sayengis of the philosophhres* [LION: EEBO]; from Traugott 2012)

Note further that according to Traugott & Dasher (2002: 84), in the early stages of the English *be going to*-future "the change is primarily abstraction (spatial > temporal)", as in the following example:

(17) *Witwoud: Gad, I have forgot what **I was going to say** to you.*
    (1699; Traugott & Dasher 2002: 84)

The interpretation proposed here is in accordance with that described in detail in Heine, Claudi & Hünnemeyer (1991) and Heine (1997), where the implicature or inference is captured in terms of a metaphorical transfer (SPACE > TIME) within

metonymic-metaphorical model proposed there (Heine, Claudi & Hünnemeyer 1991: 70, 113).[10]

What this interpretation argues for is that grammaticalization processes such as the ones described in this paper are ultimately due to the cognitive-communicative strategies that interlocutors recruit in order to create their discourse contributions. And one major strategy is to use concrete, referential and clearly delineated expressions to also convey more abstract, non-referential and/or less clearly delineated meanings. In doing so they constantly propose new discourse options, and some of these new options may be used regularly and give rise to new patterns of grammar. On this understanding there is not really "coevolution of form and meaning" (Bybee, Perkins & Pagliuca 1994b: 4); rather, the evolution of the former is caused by and, hence, is preceded in time by that of the latter. Accordingly, the directionality to be observed in structural reduction is derivative of the semantic changes to be observed in grammaticalization processes of the kind examined in this paper.[11]

The !Xun examples discussed in §2 illustrate this temporal asymmetry between form and meaning in the development of future tenses. As we saw in examples §2-§4, the future tenses in the N1, E2, and E3 dialects are ambiguous between the lexical meaning of a movement verb and the grammatical meaning of future tense. What this suggests is that there must have been a semantic shift from verbal to grammatical meaning and now both coexist in the dialects concerned. But this semantic shift does not appear to have been accompanied so far by corresponding structural (morphosyntactic and/or morphophonological) shift. Accordingly, the only reasonable conclusion is that there was semantic but so far no morphosyntactic change – in other words, structural change lags behind semantic change (Heine forthcoming).

5 Conclusions

Our starting point was the situation in the "Khoisan" language !Xun of southwestern Africa, where speakers of a number of different dialects appear to have moved in the same direction in designing a future tense category. In doing so, they appear to have drawn on a crosslinguistically common conceptual pathway whereby a verb for directed spatial movement belonging to the basic vocabulary

---

[10] We are grateful to Andrej Malchukov (p.c.) for having drawn our attention to this point.

[11] To be sure, structural change can also be instrumental to inducing semantic change, as demonstrated, for example, in the work on degrammaticalization (see especially Norde 2009), but this does not normally appear to apply to the evolution from lexical to grammatical categories.

in the sense of Swadesh (1952) in combination with another verb over time gives rise to a grammatical category expressing prediction, that is, a future tense. Thus, the paper was restricted to one specific pathway of grammaticalization, ignoring other pathways that have movement verbs as their source or future tense as their target. Whether, or to what extent, the findings made can be generalized beyond this pathway is a question that is beyond the scope of the present paper.

According to the findings presented, it is neither the context nor inferential mechanisms or analogy that seem to provide a sufficient basis for explaining the "macro-shift" from lexical source to the grammatical target of a future tense category. What appears to be involved most of all is some fixed asymmetric semantic relation between source concepts for 'go (to)' and 'come (to)' and the grammatical target concept of future tense. The causal nature of this relation is in need of much further research, it is presumably shaped or influenced by discourse functions, e.g. by the fact that the source meaning is functionally useful "in a discursively secondary role" (Harder & Boye 2011: 65).

Thus, the hypothesis in (9), proposed on the basis of observations made in the "Khoisan" language !Xun does not appear to be falsified by the data examined in §3 on movement-based future tenses in Germanic languages.

Much of what was discussed in the paper could have been phrased within the framework of Construction Grammar, that is, as an instance of constructional change (or constructionalization). A considerable part of work within this framework has in fact been devoted more recently to issues of grammaticalization (see, e.g., Traugott 2003; Noël 2007; Trousdale 2008; Hilpert 2008; 2013; 2015; Bisang 2010; De Smet 2010; Gisborne & Patten 2011; Van Bogaert 2011; Trousdale 2013; Hüning & Booij 2014; Traugott & Trousdale 2014). This work has brought about a wealth of information on the history of the constructions concerned, including the history of constructions that were the topic of this paper.

The main reason for not drawing on this framework here is that the goals of Construction Grammar and grammaticalization theory are not the same and, hence, entail a different perspective of what grammatical change is about. The former is concerned with how constructions change, and most of all with what happens on the way from source to target construction. The latter, by contrast, is ultimately concerned with the following questions: What induces interlocutors in discourse across the world to draw on much the same lexical resources to create a new functional category for future tense, and why is this semantic process essentially regular, e.g., why is it fairly unlikely that there will be a process in the opposite direction? To our knowledge, the only explanatory account that exists so far is one with reference to the cognitive-communicative strategies

that speakers and hearers have when they design their discourse contributions (Heine, Claudi & Hünnemeyer 1991).

Thus, grammaticalization theory is concerned with the "macro-shift" from source to target meaning whereas the main concern of Construction Grammar is with the process leading from the former to the latter, that is, with the constructional history of the process. Accordingly, neither the perspective underlying these two frameworks nor the results obtained are the same. Nevertheless, both frameworks are needed for a comprehensive understanding of grammatical change.

Acknowledgements

Abbreviations

1, 2, 3	first, second, third person	PROG	progressive aspect
		SG	singular
FUT	future tense	T	transitive suffix (valency-
ICPL	incompletive aspect		increasing marker)
N1, N2, N3, N4	noun class 1, 2, 3, or 4	TOP	topic marker
PAST	past tense marker	TR	linker.

References

Bisang, Walter. 1998. Grammaticalization and language contact, constructions and positions. In Anna G. Ramat & Paul J. Hopper (eds.), *The limits of garammaticalization*, 13–58. Amsterdam, Philadelphia: Benjamins.

Bisang, Walter. 2010. Grammaticalization in Chinese. A construction-based account. In Elizabeth C. Traugott & Graeme Trousdale (eds.), *Gradience, gradualness and grammaticalization*, 245–277. Amsterdam: John Benjamins.

Bernd Heine, Tania Kuteva & Heiko Narrog

Börjars, Kerstin & Nigel Vincent. 2011. Grammaticalization and directionality. In H. Narrog & B. Heine (eds.), *The Oxford handbook of grammaticalization*, 162–176. Oxford: Oxford University Press.

Brems, Lieselotte. 2007. The grammaticalization of Small Size Nouns: Reconsidering frequency and analogy. *Journal of English Linguistics* 35. 293–324.

Bybee, Joan, R. D. Perkins & W. Pagliuca. 1994a. *The evolution of grammar: Tense, aspect and modality in the languages of the world.* Chicago-London: Chicago Univ. Press.

Bybee, Joan L. 2003. Mechanisms of change in grammaticization: The role of frequency. In Brian D. Joseph & Richard D. Janda (eds.), *The handbook of historical linguistics*, 602–623. Oxford: Blackwell.

Bybee, Joan L. 2006. From usage to grammar: The mind's response to repetition. *Language 82* 4. 711–733.

Bybee, Joan L. 2011. Usage-based theory and grammaticalization. In H. Narrog & B. Heine (eds.), *The Oxford handbook of grammaticalization*, 225–238. Oxford: Oxford University Press.

Bybee, Joan L., William Pagliuca & Revere D. Perkins. 1991. Back to the future. In Elizabeth C. Traugott & B. Heine (eds.), *Approaches to grammaticalization*, vol. 2, 17–58. Amsterdam, Philadelphia: Benjamins.

Bybee, Joan L., Revere D. Perkins & William Pagliuca. 1994b. *The evolution of grammar: Tense, aspect, and modality in the languages of the world.* Chicago: University of Chicago Press.

Campbell, Lyle. 1987. Syntactic change in Pipil. *International Journal of American Linguistics* 53(3). 253–280.

Dahl, Östen. 2000. Verbs of becoming as future copulas. In Östen Dahl (ed.), *Tense and aspect in the languages of Europe* (EALT/EUROTYP 20-6), 309–328. Berlin: Mouton de Gruyter.

De Smet, Hendrik. 2010. Grammatical interference. Subject marker *for* and the phrasal verb particles *out* and *forth*. In Elizabeth C. Traugott & Graeme Trousdale (eds.), *Gradience, gradualness and grammaticalization*, 75–104. Amsterdam: John Benjamins.

Dickens, Patrick. 2005. *A concise grammar of Ju/'hoan: With a Ju/'hoan-English glossary and a subject index* (Quellen zur Khoisan Forschung 17). Cologne: Köppe.

Disney, S. J. 2009. The grammaticalisation of 'be going to'. *Newcastle Workinng Papers in Linguistics* 15. 63–79.

Fischer, Olga. 2000. Grammaticalisation: Unidirectional, non-reversable. In O. Fischer, A. Rosenbach & D. Stein (eds.), *Pathways of change. Grammaticalization in English*, 149–169. Amsterdam, Philadelphia: John Benjamins.

Fischer, Olga. 2013. An inquiry into unidirectionality as a foundational element of grammaticalization: On the role played by analogy and the synchronic grammar system in processes of language change. *Studies in Language* 3. 515–33.

Gisborne, Nikolas & Amanda Patten. 2011. Construction: Grammar and grammaticalization. In H. Narrog & B. Heine (eds.), *The Oxford handbook of grammaticalization*, 92–104. Oxford: Oxford University Press.

Greenberg, Joseph H. 1963. *The languages of Africa.* The Hague: Mouton.

Güldemann, Tom. 2014. 'khoisan' linguistic classification today. In Tom Güldemann & Anne-Maria Fehn (eds.), *Beyond 'Khoisan'. Historical Relations in the Kalahari Basin*, 1–40. Amsterdam & Philadelphia: Benjamins.

Harder, Peter & Kasper Boye. 2011. Grammaticalization and functional linguistics. In H. Narrog & B. Heine (eds.), *The Oxford handbook of grammaticalization.* 56–68. Oxford: Oxford University Press.

Heine, Bernd. 1995. Conceptual grammaticalization and prediction. In John R. Taylor & Robert E. MacLaury (eds.), *Language and the cognitive construal of the world*, 119–135. Berlin & New York: Mouton de Gruyter.

Heine, Bernd. 1997. *Cognitive foundations of grammar.* New York: Oxford University Press.

Heine, Bernd. Forthcoming. Grammaticalization in Africa: Two contrasting hypotheses. In *Grammaticalization and typology.* Oxford: Oxford University Press.

Heine, Bernd, Ulrike Claudi & Friederike Hünnemeyer. 1991. *Grammaticalization: A conceptual framework.* Chicago: University of Chicago Press.

Heine, Bernd & Henry Honken. 2010. The Kx'a family. *Journal of Asian and African Studies* 79. 5–36.

Heine, Bernd & Christa König. 2005. Grammatical hybrids: Between serialization, compounding and derivation in ! Xun (North Khoisan). In Wolfgang U. Dressler, D. Kastovsky, Oskar E. Pfeiffer & Franz Rainer (eds.), *Morphology and its demarcations: Selected papers from the 11th Morphology Meeting, Vienna, February 2004*, 81–96. Amsterdam, Philadelphia: John Benjamins.

Heine, Bernd & Christa König. 2015. *The !Xun Language: A Dialect Grammar of Northern Khoisan* (Quellen zur Khoisan-Forschung). Cologne: Rüdiger Köppe.

Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization.* Cambridge: Cambridge University Press.

Hilpert, Martin. 2008. *Germanic future constructions: A Usage-Based approach to language change* (Constructional Approaches to Language 7). Amsterdam, Philadelphia: Benjamins.

Hilpert, Martin. 2013. *Constructional change in English: Developments in Allomorphy, Word-Formation, and Syntax*. Cambridge: Cambridge University Press.

Hilpert, Martin. 2015. From hand-carved to computer-based: Noun-participle compounding and the upward strengthening hypothesis. *Cognitive Linguistics* 26(1). 113–147.

Hoffmann, Sebastian. 2004. Are low-frequency complex prepositions grammaticalized? On the limits of corpus data – and the importance of intuition. In H. Lindquist & C. Mair (eds.), *Corpus approaches to grammaticalization in English*, 171–210. Amsterdam: Benjamins.

Hoffmann, Sebastian. 2005. *Grammaticalization and English complex prepositions: A Corpus-Based study*. London, New York: Routledge.

Hopper, Paul J. & Elizabeth C. Traugott. 2003. *Grammaticalization*. Cambridge: Cambridge University Press.

Hüning, Matthias & Geert Booij. 2014. From compounding to derivation: The emergence of derivational affixes through "constructionalization". *Folia Linguistica* 48(2). 579–604.

Kiparsky, Paul. 2012. Grammaticalization as optimization. In D. Jonas, J. Whitman & A. Garrett (eds.), *Grammatical change: Origins, nature, outcomes*, 15–51. Oxford: Oxford University Press.

König, Christa. 2010. Verb serialization in ! Xun. In Matthias Brenzinger & Christa König (eds.), *Khoisan languages and linguistics. The riezlern Symposium 2003* (Quellen zur Khoisan-Forschung 17), 152–185. Cologne: Köppe.

Lehmann, Christian. 1982. *Thoughts on Grammaticalization. A Programmatic Sketch. Vol.1* (AKUP/ Arbeiten des Kölner Universalien-Projekts 48). Cologne: Universität zu Köln, Institut für Sprachwissenschaft.

Mair, Christian. 2004. Corpus linguistics and grammaticalization theory: Statistics, frequency, and beyond. In H. Lindquist & C. Mair (eds.), *Corpus Approaches to Grammaticalization in English*, 121–150. Amsterdam, Philadelphia: Benjamins.

Mair, Christian. 2011. Grammaticalization and corpus linguistics. In Heiko Narrog & Bernd Heine (eds.), *The Oxford handbook of grammaticalization*, 239–50. Oxford: Oxford University Press.

Newmeyer, Frederick J. 1998. *Language form and language function*. Cambridge, MA: MIT Press.

Noël, Dirk. 2007. Diachronic construction grammar and grammaticalization theory. *Functions of Language* 14(2). 177–202.

Norde, Muriel. 2009. *Degrammaticalization.* Oxford: Oxford University Press.

Norde, Muriel & Karin Beijering. 2014. Facing interfaces: A clustering approach to grammaticalization and related changes. *Folia Linguistica* 48(2). 385–424.

Ramat, Paolo. 2015. Grammaticalization. In Neil J. Smelser & Paul B. Baltes (eds.), *International encyclopedia of the social and behavioral sciences. Second edition*, 330–335. Oxford: Elsevier.

Sapir, Edward. 1921. *Language: An introduction to the study of speech.* New York: Harcourt, Brace & Company. http://www.gutenberg.org/etext/12629.

Swadesh, Morris. 1952. Lexicostatistic dating of prehistoric ethnic contacts. *Proceedings American Philosophical Society* 96. 452–463.

Torres Cacoullos, Rena & James A. Walker. 2011. Collocations in grammaticalization and variation. In H. Narrog & B. Heine (eds.), *The Oxford handbook of grammaticalization*, 225–238. Oxford: Oxford University Press.

Traugott, Elizabeth C. 2003. Constructions in grammaticalization. In Brian D. Joseph & Richard D. Janda (eds.), *The handbook of historical linguistics*, 624–647. Oxford: Blackwell.

Traugott, Elizabeth C. 2012. *Grammatical constructionalization: Rethinking grammaticalization in the light of constructionalization.* Paper presented at the University of Santiago de Compostela, 17 October 2012.

Traugott, Elizabeth C. & Richard B. Dasher. 2002. *Regularity in semantic change* (Cambridge Studies in Linguistics 96). Cambridge: Cambridge University Press.

Traugott, Elizabeth C. & Graeme Trousdale. 2014. Contentful constructionalization. *Journal of Historical Linguistics* 4(2). 254–282.

Trousdale, Graeme. 2008. Constructions in grammaticalization and lexicalization: Evidence from the history of a composite predicate in English. In Graeme Trousdale & Nikolas Gisborne (eds.), *Constructional approaches to English grammar*, 33–67. Berlin: Mouton de Gruyter.

Trousdale, Graeme. 2013. Multiple inheritance and constructional change. *Studies in Language* 37(3). 491–514.

Ultan, Russell. 1978. Some general characteristics of interrogative systems. In Joseph H. Greenberg, Charles A. Ferguson & Edith Moravcsik (eds.), *Universals of human language. Volume 3: Word structure*, 211–248. Stanford: Stanford University Press.

Van Bogaert, Julie. 2011. *I think* and other complement-taking mental predicates: A case of constructional grammaticalization. *Linguistics* 2. 295–332.

Bernd Heine, Tania Kuteva & Heiko Narrog

Vedder, Heinz. 1910-1911. Grundriß einer Grammatik der Buschmannsprache vom Stamm der ! Ku-Buschmänner. *Zeitschrift für Kolonialsprachen* 1(5-24). 106–117.

Chapter 3

The origin of comitative adverbs in Japhug

Guillaume Jacques

CNRS-CRLAO-INALCO-EHESS

The aim of this paper is threefold. First, it provides a description of the morphological and syntactic properties of comitative adverbs in Japhug and other Gyalrong languages, a class of adverbs derived from nouns by a combination of prefixation and reduplication. Second, it argues that they derive from a two-step derivation, first from noun into proprietive denominal verb, then from that verb into a participial form. The resulting form is later reanalyzed as a single morphological derivation from the noun. Third, this paper contributes to the study of language contact within the Gyalrongic group by showing how one of the two processes to build comitative adverbs in Japhug is borrowed from the neighbouring Tshobdun language.

Introduction

This paper discusses the origin of comitative adverbs in Japhug and other Gyalrong languages. These adverbs, only attested in the core Gyalrong languages, are a relatively recent innovation in these languages, and provide an interesting case study to investigate the origin of comitative constructions in the world's languages.

The paper contains four sections. First, I provide background information on Japhug and the other Gyalrongic languages. Second, I describe the morphological expression of possession in Japhug nouns, which must be taken into consideration in all types of denominal derivations, including that of comitative adverbs. Third, I discuss the syntactic morphological and syntactic properties of comitative adverbs. Fourth, I propose a grammaticalization hypothesis to account for their origin, involving comparison with the closely related Tshobdun language,

and show that the pathway in question has not been previously proposed for comitative markers.

1 Japhug and Gyalrongic languages

Japhug (in Chinese Chapu 茶堡) is a Gyalrong language spoken in Mbarkham county, Rngaba prefecture, Sichuan, China. The present study is based on the Kamnyu dialect, whose location is indicated in Figure 1. In addition to Japhug, there are three other Gyalrong languages, Tshobdun (in Chinese Caodeng 草等, Sun 2003), Zbu (aka Showu, in Chinese Ribu 日部, see Sun 2004; Gong 2014) and Situ, the language with the greatest number of speakers and dialectal variation (四土, Lin 1993; Huáng & Sūn 2002; Prins 2011). Gyalrong languages in turn belong to the Gyalrongic branch of Trans-Himalayan, which also includes Stau and Khroskyabs (see Sun 2000 and Lai 2015).
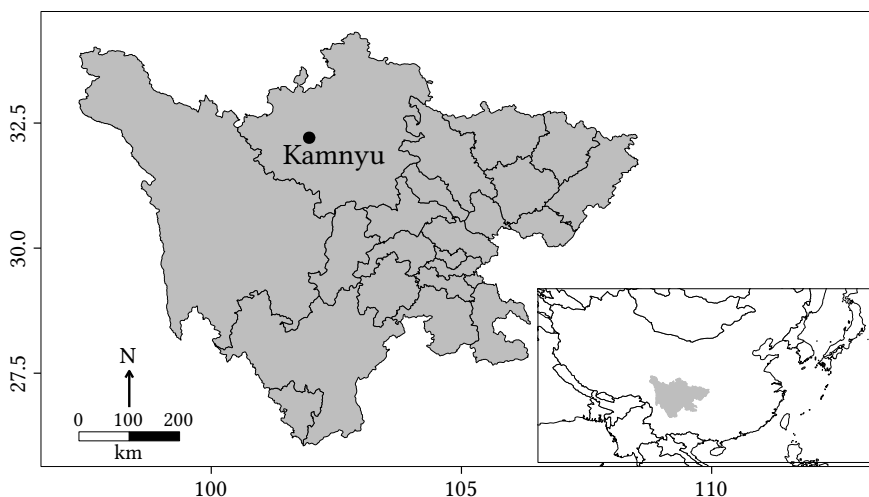
Figure 1: Location of Kamnyu village

Gyalrong languages, unlike most other languages of the Trans-Himalayan family, are polysynthetic languages with a rich derivational and inflectional verbal morphology (Jacques 2012b; Sun 2014a) and direct-inverse indexation (DeLancey 1981; Sun & Shidanluo 2002; Jacques 2010; Gong 2014) that are argued to be of proto-Trans-Himalayan origin (DeLancey 2010; Jacques 2012a). This morphology is typologically unusual in being mainly prefixing despite Gyalrong languages having strict verb-final word order (Jacques 2013).

2 Inalienably possessed nouns

Japhug nouns can be divided into inalienably possessed nouns (IPN) and non-inalienably possessed nouns (NIPN). IPNs differ from NIPNs in that they require the presence of one of the possessive prefixes (Table 1), while NIPNs can appear as their bare stem without any possessive prefix. The IPN / NIPN distinction is not completely predictable: although all body parts and kinship terms are IPNs, we also find nouns referring to (but not all) clothes (*tɯ-ŋga* 'clothes', *tɯ-xtsa* 'shoes' etc), some implements (*tɤ-mkɯm* 'pillow'), and abstract concepts (*tɯ-sɯm* 'thought', *tɯ-ʑɯβ* 'sleep', *tɯ-pʰɯ* 'price', *tɯ-nɲa* 'debt', *tɯ-kʰɯr* 'official position' etc). Note that IPNs can refer to entities or properties that are not necessarily permanently and definitively associated with the possessor, as is the case with clothes and concepts like 'debt' or 'official position', but that are not freely removable during at least a specified time (the time of wake in the case of 'clothes', the time of sleep in the case of 'pillow', the period between contracting the debt and repaying it in the case of 'debt' etc).

When no definite possessor is present, IPNs take one of the indefinite possessive prefixes *tɤ–* or *tɯ–*. The citation form of IPNs is built by combining one of the indefinite prefixes with the noun stem (*tɤ-lu* 'milk', *tɯ-ŋga* 'clothes', *tɤ-rpɯ* 'uncle', *tɯ-ci* 'water'). The distribution of the prefixes *tɤ–* vs *tɯ–* is lexically determined. When a specific possessor is present, the indefinite prefix is replaced by the appropriate possessive prefix (*ɯ-lu* 'her/its milk (from her nipple)', *a-ŋga* 'my clothes', *nɤ-rpɯ* 'your uncle', *ɯ-ci* 'its juice').

Although the generic possessive prefix *tɯ–* is homophonous with one of the indefinite possessive prefixes, the two are semantically distinct (compare *tɤ-se* INDEF.POSS-blood 'blood' with *tɯ-se* GENR.POSS-blood 'one's/people's blood').

It is possible to turn an IPN into a NIPN by prefixing a definite possessive prefix to the indefinite one, as in *ɯ-tɤ-lu* 3SG.POSS-INDEF.POSS-milk 'his milk (to drink)', *ɯ-tɯ-ci* 3SG.POSS-INDEF.POSS-water 'its water (of irrigated water, to a plant)'.[1] NIPNs cannot take indefinite possessive prefixes. However, they are compatible with the human generic possessor prefix *tɯ–*, as in (1), where the nouns *kha* 'house' and *laχtɕha* 'thing' are NIPNs.

(1) *wuma ʑo*   *tɯ-kha*   *cho tɯ-laχtɕha*   *ra*
    really EMPH GENR.POSS-house and GENR.POSS-thing PL

---

[1] A similar phenomenon is reported in Tshobdun (Sun 1998: 140)

Guillaume Jacques

Table 1: Possessive prefixes

Prefix	Person
a−	1SG
nɣ−	2SG
ɯ−	3SG
tɕi−	1DU
ndʑi−	2DU
ndʑi−	3DU
i−	1PL
nɯ−	2PL
nɯ−	3PL
tɯ−, tɣ−	indefinite
tɯ−	generic

sɯ-ɴqhi
CAUS-be.dirty:FACT

(Flies) make one's house and one's things dirty. (25 akWzgumba, 62)

3 Comitative derivation

In Japhug, adverbs meaning 'having X' or 'together with X' can be productively built from various types of nouns.[2] In this section, I first describe the morphological processes involved in the noun to adverb derivation, and then provide an overview of the use of these adverbs in context.

3.1 Morphology

Comitative adverbs are formed by reduplicating the last syllable of the noun stem and prefixing either *kɣ−* or *kɣɣɯ−*, as in examples such as *χɕɣlmuɣ* 'glasses' ⇒ *kɣ-χɕɣlmɯ~lmuɣ* / *kɣɣɯ-χɕɣlmɯ~lmuɣ* 'together with glasses'.[3] No semantic

---

[2] Comitative adverbs in Japhug have been briefly mentioned in a previous publication (Jacques 2008: 51), but this paper is the first detailed description of this derivation and its uses.

[3] Japhug *χɕɣlmuɣ* 'glasses' is a loanword from Tibetan *ɕel.mig*; note that reduplication disregards morpheme boundaries (*χɕɣl* 'glass' (Tibetan *ɕel*) is also attested in Japhug).

34

difference between the comitative adverbs in *kɣ–* and those in *kɣɣɯ–* has been detected; both are fully productive and can be built from the same nouns.

When the base noun is an IPN, it is possible to build a comitative adverb with the indefinite possessor prefix or with the bare stem. For instance, from *tɤ-rte* 'hat' one can derive both *kɣ-rtɯ~rte* / *kɣɣɯ-rtɯ~rte* 'with his/her hat' and *kɣ-tɤ-rtɯ~rte* / *kɣɣɯ-tɤ-rtɯ~rte* 'with a hat' with the indefinite possessor prefix *tɤ–*. The inalienable/non-alienable distinction is present in these forms: *kɣ-rtɯ~rte* / *kɣɣɯ-rtɯ~rte* means 'wearing one's hat' (2), while *kɣ-tɤ-rtɯ~rte* / *kɣɣɯ-tɤ-rtɯ~rte* implies that one is not wearing the hat (3).

(2) *kɣɣɯ-rtɯ~rte ʑo   kha   ɯ-ŋgɯ   lɤ-tɯ-ɣe*
    COMIT-hat   EMPH house 3SG-inside PFV-2-come[II]
    You came inside the house with your hat. (You were expected to take it off before coming in)

(3) *laχtɕha kɣɣɯ-tɤ-rtɯ~rte   ʑo   ta-ndo*
    thing   COMIT-INDEF.POSS-hat EMPH PFV:3→3'-take
    He took the things together with the hat. (Not wearing it)

Cognates of the Japhug comitative adverbs have been reported in other Gyalrong languages, in particular Tshobdun *ko–* (Sun 1998: 107), and the comitative adverb derivation can thus be reconstructed back at least to proto-Gyalrong. However, given the dearth of data on languages other than Japhug (in particular in terms of text examples), little external data will be discussed in this paper. A full comparative assessment of the hypotheses laid out here will have to wait the publication of fully-fledged grammatical descriptions of all Gyalrong languages.

Comitative adverbs, in any case, appear to be unattested outside of core Gyalrong languages (even in Khroskyabs, their closest relative, see Lai 2013), and is probably one of the many common Gyalrong morphological innovations.

3.2 Syntactic uses

The comitative adverb can either follow (4) or precede (5, 6) the noun over which it has scope. Alternatively, a comitative adverb can occur without a corresponding overt noun (7). However, if the noun is overt, the comitative adverb is contiguous to the NP to which it belongs.

The NP in question can either correspond to the P (4, 6), the S (5, 7) or even the A (8). This last option is not attested in the text corpus, but speakers have no trouble producing sentences of this type.

Guillaume Jacques

(4) tɤ-sno   kɤ́-juɯ~jaʁ nɯ lu-ta-nɯ
    INDEF.POSS-saddle COMIT-hand DEM IPFV-put-PL
    (Then), they put the saddle with its handles.

(5) pɣɤkhɯ nɯ ɯ-ku   nɯnɯ lɯlu tsa   ɲɯ-fse,
    owl   DEM 3SG.POSS-head DEM   cat a.little SENS-be.like
    ɯ-mtsioʁ   ɣɤʑu   ma   kɤ́-rnɯ~rna lɯlu
    3SG.POSS-beak exist:SENS a.part.from COMIT-ear   cat
    ɯ-tɯ-fse   ɲɯ-sɤre   ʑo.
    3SG.NMLZ:DEGREE-be.like SENS-be.extremely/be.funny EMPH
    The owl's head looks a little like that of a cat, apart from the fact that it
    has a beak, it looks very much like a cat with its ears.

(6) kɤ́-thɤlwɯ~lwa ɯ-zrɤm   ra kɯnɤ chɯ́-wɣ-ɣɯt pjɯ́-wɣ-ji   ri
    COMIT-earth   3SG.POSS-root PL also IPFV-INV-bring IPFV-INV-plant but
    maka tu-łoʁ   mɯ́j-cha
    at.all IPFV-come.out NEG:SENS-can
    Even if one takes its root with earth (around it) and plant it, it cannot
    grow.

(7) kɤ́-snɯ~sno ʑo kɤ-rŋgɯ
    COMIT-saddle EMPH PFV-lie.down
    (The horse) slept with its saddle. (elicited)

(8) lɯlu kɤ́-rɟɯ~rɟit   ra kɯ ʑo βʑɯ to-ndza-nɯ.
    cat COMIT-offspring PL ERG EMPH mouse IFR-eat-PL
    The cat and its young ate the mouse. (elicited)

Nouns incorporated into comitative adverbs lose their nominal status and can-
not be determined by relative clauses (including attributive adjectives), numerals
or demonstratives. In a sentence such as (9) for instance, the attributive particip-
ial relative [kɯ~kɯ-ŋɤn] 'all the ones who are evil' does not determine kɤɣɯ-
ŋkhɯ~ŋkhor 'with his subjects', a syntactic structure which would correspond
to the translation 'with all his evil subjects'. Rather, it determines the head noun
together with the comitative adverb rɟɤlpu kɤɣɯ-ŋkhɯ~ŋkhor 'the king with his
subjects', which implies the translation given below.

(9) rɟɤlpu kɤɣɯ-ŋkhɯ~ŋkhor [kɯ~kɯ-ŋɤn]   ʑo to-ndo tɕe,
    king COMIT-subjects   TOTAL~NMLZ:S/A-be.bad EMPH IFR-take LNK
    tɕendɤre kɯ-mɤku   nɯ sɤtɕha
    LNK   NMLZ:S/A-be.before DEM place

kɯ~kɯ-sɤ-scit   *ʑo*   *jo-tsɯm*   *ɲɯ-ŋu ri*
TOTAL~NMLZ:S/A-DEEXP-be.happy EMPH IFR-take.away SENS-be LNK
kɯ-maqhu   *tɕe, kɯ~kɯ-sɤɣ-mu*   *ʑo*   *jo-tsɯm*
NMLZ:S/A-be.after LNK TOTAL~NMLZ:S/A-DEEXP-fear EMPH IFR-take.away
tɕe
LNK

She took the king and his subjects, all the evil ones, in the beginning she took them to nice places, but later she took them to fearful places. (slobdpon)

4 Grammaticalization pathway

In this section, I first present the proprietive denominal derivation in *aɣɯ‒* and the infinitival and participial prefix *kɯ‒*. Then, I show that comitative adverbs are actually synchronically formally ambiguous with the infinitive and the S/A-participle of proprietive denominal verbs in some contexts. Finally, I propose that comitative adverbs diachronically derive from the participial forms of proprietive denominal verbs, and was then extended to other contexts after reanalysis.

4.1 Denominal derivation

Japhug has a rich array of denominal prefixes (Jacques 2014b). One of these prefixes, *aɣɯ‒*, derives stative intransitive verbs from either inalienably possessed and non-inalienably possessed nouns. As illustrated by the examples in Table 2, verbs derived with the prefix have meaning such as 'having X', 'producing a lot of X' or 'having the same X ' (with plural S). The noun stem is sometimes reduplicated, especially for the first type of meaning.

Table 2: The denominal prefix *aɣɯ‒*

Base noun	Meaning	Denominal verb	Meaning
tɯ-ɣli	excrement, manure	*aɣɯ-ɣli*	producing a lot of manure (of pigs)
tɤ-lu	milk	*aɣɯ-lu*	producing a lot of milk (of cows)
tɯ-mɲaʁ	eye	*aɣɯ-mɲaʁ*	having a lot of holes
tɯ-ɕnaβ	snot	*aɣɯ-ɕnɯ~ɕnaβ*	be slimy
ɯ-mdoʁ	colour	*aɣɯ-mdoʁ*	having the same colour
tɯ-sɯm	thought	*aɣɯ-sɯm*	get along well
smɤn	medicine	*aɣɯ-smɤn*	have a medical effect
tɯ-ɕna	nose	*aɣɯ-ɕnɯ~ɕna*	having a keen sense of smell

Guillaume Jacques

In some cases, the semantic relationship between the base noun and the derived verb is more metaphorical and not predictable. For instance, from the noun *tɯ-jaʁ* 'hand' one can derive either *aɣɯ-jɯ~jaʁ* 'having a lot of hands' (of a bug), while the non-reduplicated form *aɣɯ-jaʁ* means 'who steals anything (that comes near his hand)'.

4.2 S/A participle and infinitive

In Japhug, stative verbs (including the denominal verbs in *aɣɯ–* presented in the previous section) have two homophonous non-finite forms with a prefix *kɯ–*, the S/A-participle (Jacques 2014b: 5) and the infinitive.[4] The participle appears mainly in participial relatives (including all forms corresponding to attributive adjectives in European languages), as in example (10).

(10) *tɕheme ci kɯ-pɯ~pe   kɯ-mpɕɯ~mpɕɤr,*
      girl a NMLZ:S/A-EMPH~be.good NMLZ:S/A-EMPH~be.beautiful
      nɤ-ɕɣa   kɯ-xtɕɯ~xtɕi   ʑo a-nɯ-tɯ-ɣβzu
      2SG.POSS-tooth NMLZ:S/A-EMPH~be.small EMPH IRR-PFV-2-become
      smɯlɤm
      prayer
      May you become a nice and beautiful girl with short teeth. (Slobdpon, 261)

The infinitive is used (by some speakers) as the citation form of verbs, and appears in some types of complement clauses and manner subordinate clauses (Jacques 2014a: 271-2; 321-5), as in (11) where *kɯ-pɯ~pe*, meaning here 'nicely', is a manner subordinate clause comprising a single verb.

(11) *ɕɤr   tɕe ʁzɤmi   ni kɯ-pɯ~pe   ʑo*
      evening LNK husband.and.wife DU INF:STAT-EMPH~good EMPH
      ɕ-ko-nɯ-rŋgɯ-ndʑi
      TRANSL-IFR-AUTO-lie.down-DU
      In the evening, the husband and the wife laid down in bed nicely.

---

[4] The morphological evidence for distinguishing between participle and infinitive is clearer with dynamic verbs, whose infinitive in *kɤ–* differs from the S/A-participle. Cognates of the participle and the infinitive *kɯ–* are found in all Gyalrong languages, with only mild differences (see in particular Sun 2014b).

4.3 Potential ambiguity

Due to the sandhi rule according to which *kɯ–* combined with *a–* initial verbs yields /*kɤ–*/ in Japhug (Jacques 2004), S/A-participles or infinitive forms of denominal verbs in *aɣɯ–* are formally homophonous with comitative adverbs in *kɤɣɯ–*. For example, the form *kɤɣɯrtɯrtaʁ* 'together with its branches' from *tɤ-rtaʁ* 'branch' is identical to the participle *kɤɣɯrtɯrtaʁ* 'the one which has many branches' found in example (12).

(12) *si   kɯ-ɣɯrtɯrtaʁ   ki kɯ-fse*
      tree NMLZ:S/A-have.many.branches this NMLZ:S/A-be.this.way
      ɲɯ-ɕar-nɯ
      IPFV-search-PL

      They search for a tree having a lot of branches like this. (NOT: 'a tree with its branches' in this particular context)

Examples (13) and (14) present a minimal pair contrasting the comitative adverb 'with his/her children' on the one hand and the participle 'having many children' on the other hand (both derived from the possessed noun *tɤ-rɟit* 'child').

(13) *iɕqha   tɕʰeme nɯ kɯ-ɣɯrɟɯrɟit   ci*
      the.aforementioned woman DEM NMLZ:S/A-have.many.children INDEF
      pɯ-ŋu
      PST.IPFV-be

      This woman had a lot of children.

(14) *kɤɣɯ-rɟɯ~rɟit ʑo jo-nɯ-ɕe-nɯ*
      COMIT-children EMPH IFR-VERT-go-PL

      She/They went back with their children.

4.4 Reanalysis

The formal ambiguity between the comitative adverbs on the one hand, and the participles and infinitives of *aɣɯ–* denominal verbs on the other hand, together with the semantic proximity of the two forms, raise the question of their potential historical relatedness.

An obvious possibility is that comitative adverbs originate from the reanalysis of the S/A-participles of reduplicated *aɣɯ–* denominal verbs. Ambiguous sentences like (12) actually constitute the pivot constructions which allow reanalysis in contexts where both proprietive ('having X') and a comitative ('with X') interpretations were possible.

Guillaume Jacques

(15) si   kɤɣɯrtɯrtaʁ   ɲɯ-ɕar-nɯ
      tree NMLZ:S/A-have.many.branches//COMIT-branch IPFV-search-PL
      'They search for a tree having a lot of branches' ⇒ 'They search for a tree
      and/with its branches'

Starting from such ambiguous sentences, the comitative adverb was extended
to nouns without a corresponding proprietive denominal verb. In addition, comitative adverbs incorporating the indefinite possessive prefix were created (such
as kɤɣɯ-tɤ-rtɯ~rte 'with his hat'). Forms of this type are clearly distinct from
infinitives or participles of denominal verbs, as indefinite possessive prefixes are
always deleted during denominal derivation.

I therefore propose the pathway (16) to account for comitative adverbs in
kɤɣɯ– in Japhug:

(16) NOUN + PROPERTY DENOMINAL DERIVATION + infinitive/participle → COMI-
      TATIVE

Among the possible origins of comitative markers, Heine & Kuteva (2002: 91,
139, 287) includes nouns meaning 'comrade' or verbs such as 'follow' and 'take'
and make no mention of proprietive markers. However, Sutton (1976) noted etymological connections between proprietive and comitative markers in several
languages of Australia, and although none of the standard references on comitative constructions (Stassen 2000; Stolz, Stroh & Urdze 2006; Arkhipov 2009)
explicitly mention a pathway PROPRIETIVE → COMITATIVE, they all noticed the
close functional relationship between these two categories, notably in languages
of Australia where both comitative and proprietive cases may exist in the same
language (for instance, Djabugay, see Patz 1991).

The pathway presented above accounts well for the kɤɣɯ– type comitative
adverbs, but does not explain the kɤ́– variant, which is actually more common
in the corpus.

Comitative adverbs in kɤ́– are anomalous in Japhug in being among the very
few prefixes attracting stress, a feature that could indicate fusion of two syllables
(for instance the negative sensory marker mɯ́j- probably historically results from
the fusion of the negative mɯ– and the sensory prefix ɲɯ–).

If the sound laws of Japhug (Jacques 2004) are applied backwards, the prefix
kɤɣɯ– would go back to pre-Japhug *kɛwə–. We know that in Tshobdun, *wə
regularly corresponds to o. It is in particular the case of the inverse prefix o–
(Sun & Shidanluo 2002) which originates from proto-Gyalrong *wə). Through
vowel fusion (which also occurs with the inverse prefix), ko–, the actual form of

the comitative prefix (Sun 1998: 107), is the expected outcome of * kɐwə–. We can therefore safely conclude that (1) the comitative prefixes kɤɣɯ– in Japhug and ko– in Tshobdun are cognate and (2) that the grammaticalization in (16) took place before the split of Japhug and Tshobdun, and show be reconstructed back at least to their common ancestor.

The comitative prefix kɤ́– in Japhug, on the other hand, makes no sense from a Japhug-internal perspective. A possible way to explain it however, is to suppose *borrowing* from Tshobdun ko–. Japhug, and especially the Kamnyu variety described in the present paper, has borrowed a few nouns from Tshobdun, as shown by forms such as *qro* 'ant', *qaliaʁ* 'eagle' and *tɯɟo* 'demon' instead of expected * qroʁ, * qarɟaʁ (attested in some dialects of Japhug) and * tɯʑu, following the sound laws set out in Jacques (2004).

Borrowing of Tshobdun ko– as Japhug kɤ́– is not surprising phonologically. The stress on the prefix in Japhug is probably a trace of the stress on that prefix in pre-Tshobdun, lost due to the strong tendency of Gyalrong languages to stress the final or penultimate syllable (Sun 2005). The vowel ɤ rather than o is a consequence of the fact that derivational prefixes in Japhug have strong phonotactic constraints: the only possible vowels are either ɤ or ɯ (and a, but only in the case of stem-initial a–).

The borrowing hypothesis also accounts for the absence of any discernible difference in function between the two comitative prefixes in Japhug.

Conclusion

The contribution of this paper is threefold. First, it provides the first detailed description of comitative adverbs in any Gyalrong language. Second, it shows that language contact between Gyalrong languages is not restricted to the lexicon, but actually also involves clear cases of borrowing of grammatical morphemes. Third, it provides an example of evolution with clear directionality from PROPRIETIVE to COMITATIVE.

Acknowledgements

Guillaume Jacques

References

Arkhipov, Alexandre. 2009. Comitative as a cross-linguistically valid category. In Patience Epps & Alexandre Arkhipov (eds.), *New Challenges in Typology: Transcending the Borders and Refining the Distinctions*, 223–246. Berlin, New York: Mouton de Gruyter.

DeLancey, Scott. 1981. The category of direction in Tibeto-Burman. *Linguistics of the Tibeto-Burman Area* 6(1). 83–101.

DeLancey, Scott. 2010. Towards a history of verb agreement in Tibeto-Burman. *Himalayan Linguistics* 9(1). 1–39.

Gong, Xun. 2014. Personal agreement system of Zbu rGyalrong (Ngyaltsu variety). *Transactions of the Philological Society* 112(1). 44–60.

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

Huáng, Liángróng & Hóngkāi Sūn. 2002. 漢嘉戎詞典 Han jiarong cidian (Chinese- Rgyalrong dictionary). (黃良榮、孫宏開). 北京：民族出版社.

Jacques, Guillaume. 2004. *Phonologie et morphologie du japhug (Rgyalrong)*. Université Paris VII - Denis Diderot PhD thesis.

Jacques, Guillaume. 2008. 嘉絨語研究 Jiāróngyǔ yánjiū (Study on the Rgyalrong language). Beijing: Minzu chubanshe.

Jacques, Guillaume. 2010. The inverse in Japhug Rgyalrong. *Language and Linguistics* 11(1). 127–157.

Jacques, Guillaume. 2012a. Agreement morphology: the case of Rgyalrongic and Kiranti. *Language and Linguistics* 13(1). 83–116.

Jacques, Guillaume. 2012b. From denominal derivation to incorporation. *Lingua* 122(11). 1207–1231.

Jacques, Guillaume. 2013. Harmonization and disharmonization of affix ordering and basic word order. *Linguistic Typology* 17(2). 187–217.

Jacques, Guillaume. 2014a. Clause linking in Japhug Rgyalrong. *Linguistics of the Tibeto-Burman Area* 37(2). 263–327.

Jacques, Guillaume. 2014b. Denominal affixes as sources of antipassive markers in Japhug Rgyalrong. *Lingua* 138. 1–22.

Lai, Yunfan. 2013. *La morphologie affixale du lavrung wobzi*. Université Paris III MA thesis.

Lai, Yunfan. 2015. The person agreement system of Wobzi Lavrung (Rgyalrongic, Tibeto-Burman). *Transactions of the Philological Society* 113(3). 271–285.

Lin, Xiangrong. 1993. *Jiarongyu yanjiu* 嘉戎語研究 (A study on the Rgyalrong language. Chengdu: Sichuan minzu chubanshe.

Michailovsky, Boyd, Martine Mazaudon, Alexis Michaud, Séverine Guillaume, Alexandre François & Evangelia Adamou. 2014. Documenting and researching endangered languages: the Pangloss Collection. *Language Documentation and Conservation* 8. 119–135.

Patz, Elizabeth. 1991. Djabugay. In Barry J. Blake (ed.), *The Handbook of Australian Languages,* vol. 4, 244–347. Oxford: Oxford University Press.

Prins, Marielle. 2011. *A web of relations: A grammar of rGyalrong Jiaomuzu (Kyom-kyo) dialects*. Leiden University PhD thesis.

Stassen, Leon. 2000. AND-languages and WITH-languages. *Linguistic Typology* 4(1). 1–54.

Stolz, Thomas, Cornelia Stroh & Aina Urdze. 2006. *On Comitatives and Related Categories: A Typological Study With Special Focus on the Languages of Europe*. Berlin, New York: Mouton de Gruyter.

Sun, Jackson T.-S. 1998. Nominal Morphology in Caodeng rGyalrong. *Bulletin of the Institute of History and Philology* 69(1). 103–149.

Sun, Jackson T.-S. 2000. Stem Alternations in Puxi Verb Inflection: Toward Validating the rGyalrongic Subgroup in Qiangic. *Language and Linguistics* 1(2). 211–232.

Sun, Jackson T.-S. 2003. Caodeng rGyalrong. In Graham Thurgood & Randy LaPolla (eds.), *The Sino-Tibetan languages*, 490–502. London: Routledge.

Sun, Jackson T.-S. 2004. Verb-Stem Variations in Showu rGyalrong. In Ying-chin Lin, Fang-min Hsu Hsu, Chun-chih Lee, Jackson T.-S. Sun, Hsiu-fang Yang Yang & Dah-an Ho (eds.), *Studies on Sino-Tibetan Languages: Papers in Honor of Professor Hwang-cherng Gong on His Seventieth Birthday*, 269–296. Language & Linguistics Monograph Series W4.

Sun, Jackson T.-S. 2005. Jiāróngyǔzǔ yǔyán de yīn'gāo: liǎnggè gè'àn yánjiū 嘉戎語組語言的音高：兩個個案研究 (On Pitch in the rGyalrongic Languages: Two Case Studies). *Yuyan yanjiu* 語言研究 25(1). 50–59.

Sun, Jackson T.-S. 2014a. Sino-Tibetan: Rgyalrong. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford Handbook of Derivational Morphology*, 630–650. Oxford: Oxford University Press.

Sun, Jackson T.-S. 2014b. Typology of Generic-Person Marking in Tshobdun Rgyalrong. In Richard VanNess Simmons & Newell Ann Van Auken (eds.),

Studies in Chinese and Sino-Tibetan Linguistics: Dialect, Phonology, Transcription and Text, 225–248. Taipei, Institute of Linguistics, Academia Sinica.

Sun, Jackson T.-S. & Shidanluo. 2002. Caodeng Jiarongyu yu rentong dengdi xiangguan de yufa xianxiang 草登嘉戎語與「認同等第」相關的語法現象 (Empathy Hierarchy in Caodeng rGyalrong grammar). *Language and Linguistics* 3(1). 79–99.

Sutton, Peter. 1976. The having affix and other morphemes in fifty Australian languages. In R.M.W. Dixon (ed.), *Grammatical categories in Australian languages*, 297–305. New Jersey: Humanities Press.

Chapter 4

Copulas originating from the imperative of 'see/look' verbs in Mande languages

Denis Creissels

University of Lyon

This paper analyzes Mande data that suggest a grammaticalization path leading from the imperative of verbs 'see/look' to ostensive predicators (i.e. words functionally similar to French *voici*, Italian *ecco*, or Russian *vot*), and further to copulas. Clear cases of copulas cognate with 'see/look' verbs are found in several branches of the Mande family, and there is convincing evidence that they did not develop from the semantic bleaching of forms originally meaning 'is seen/found' (another plausible grammaticalization path leading from 'see' verbs to copulas), but from the routinization of the ostensive use of the imperative of 'see/look'. Comparison of the Mande data with the Arabic data provided by Taine-Cheikh (2013) shows however that this is not the only possibility for imperatives of verbs 'see/look' to grammaticalize into copulas, since in the Arabic varieties in which the imperative form of 'see' has become a plain copula, the most plausible explanation is that a modal/discursive particle resulting from the grammaticalization of the imperative of 'see' has undergone a process of semantic bleaching in the context of an equative or locational predicative construction that initially included no overt predicator.

1 Introduction

The grammaticalization path leading from the imperative of verbs 'see/look' to ostensive predicators or to copulas is not mentioned in the inventory of grammaticalization processes provided by Heine & Kuteva (2002), and 'see/look' verbs are not mentioned as a possible source of copulas in general accounts of non-verbal predication such as Hengeveld (1992) or Pustet (2003) either. However, French *voici/voilà* constitute a well-known example of the grammaticalization of the imperative of a verb 'see' as an ostensive predicator, and additional examples

Denis Creissels

can be found for example among Chadic languages (see Hellwig (2011: 380-382) on Goemai; Jaggar (2001: 468-469) and Newman (2000: 181-182) on Hausa). As regards the possibility that the imperative form of a verb 'see/look' grammaticalizes as a copula, this possibility has been recognized so far in two language families, and in one of these two cases, the first stage in this evolution is the reanalysis of the imperative of a 'see/look' verb as an ostensive predicator: (a) As discussed by Taine-Cheikh (2013), in Arabic languages, the grammaticalization of the imperative form of verbs cognate with Classical Arabic *raʔā* 'see' has developed in different directions, with the creation of a copula as one of its possible outcomes.[1] (b) As observed by Westermann & Melzian (1930), Monteil (1939), Heydorn (1940-1941) Heydorn (1949-1950), Welmers (1974), Creissels (1981), and Tröbs (2003), Mande languages provide evidence that copulas may result from the evolution of ostensive predicators whose origin is the imperative of a verb 'see'. This is however not the only possible type of evolution resulting in the creation of a copula or an existential verb from a 'see' verb. Cross-linguistically, the translation equivalents of 'see' may be polysemous verbs expressing the meanings commonly expressed in English as 'find' or 'get', and it is easy to imagine a process of semantic bleaching converting a form meaning 'is found' into a locational copula. As rightly observed by a reviewer, in Sanskrit the root VID 'see/know' (from Indo-European *weid) in passive form (*vid-ya-te*) was used in the classical language with the meaning 'there is', and more generally, the pathway (IS_SEEN~)IS_FOUND > LOCATIONAL COPULA (or variants thereof)[2] may be more common cross-linguistically than the creation of copulas from the imperative form of 'see/look'. In this article, after clarifying the notion of ostensive predicator (§2) and providing some background information on Mande languages, and in particular on Mande predicative constructions (§3), I present comparative data on copulas originating from 'see/look' verbs in Mande languages (§4). §5 compares the Mande data with the Arabic data provided by Taine-Cheikh (2013). In §6, I discuss the details of two possible grammaticalization paths by which the imperative of a 'see/look' verb may be converted into a copula. In §7, I discuss evidence against the alternative hypothesis according to which the Arabic and Mande copulas analyzed in this article might have resulted from the pathway (IS_SEEN~)IS_FOUND > LOCATIONAL COPULA. §8 summarizes the main conclusions.

---

[1] The grammaticalization of the imperative of 'see' into an ostensive particle, and further into a copula, in some Arabic varieties, was already briefly signaled by (Rubin 2005: 43).

[2] The same reviewer signals that the pathway APPEAR > COPULA/EXISTENTIAL VERB is unambiguously attested in some varieties of Tibetan, where the reflexes of Classical Tibetan *snang* 'appear' are used as copula and evidential marker (Suzuki 2012).

2 Ostensive predicators

I define *ostensive predicators* as grammatical words or expressions whose combination with a noun phrase constitutes the core of clauses aiming to draw the attention of the addressee to the presence of some entity in the situation within which the speaker-addressee interaction takes place (speech situation), such as French *voici*, English *here is*, Italian *ecco*, Russian *vot*, etc. Ostensive predicators are more commonly called 'presentative particles' (Petit 2010), but this term is ambiguous in two respects: on the one hand, 'presentative' is sometimes used as an equivalent of 'existential', and on the other hand, the label 'presentative particle' is sometimes used for words that have a different distribution (in particular, for interjections). Ostensive predicators entail meanings typically expressed by copulas: identification of a referent, and presence of a referent at some place. They differ from copulas in two crucial respects: the deictic component of their meaning, and syntactic constraints following from the particular illocutionary force they carry. The argument of an ostensive predicator must be located in the speech situation, and ostensive clauses can be neither negated nor questioned, since their function is to draw the addressee's attention to an obvious fact. In this respect, some similarity can be recognized between ostensive clauses and exclamatory clauses. In addition to their use in clauses that consist of just the ostensive predicator and a noun phrase, ostensive predicators often occur with the same deictic meaning in constructions in which they combine with a complement clause – Ex. (1b), or in constructions that can be described as including a secondary predication (or 'small clause') – Ex. (1c).

(1) French

    a. *Voici nos amis.*
       OST   our friends

       'Here are our friends.'

    b. *Voici que   nos amis   arrivent.*
       OST   COMP our friends arrive

       'Behold, our friends are coming.' (lit. 'Here is that our friends are coming!')

    c. *Voici nos amis   qui arrivent.*
       OST   our friends REL arrive

       'Behold, our friends are coming.' (lit. 'Here are our friends that are coming!')

Denis Creissels

3 Verbal predication and copulas in Mande languages

3.1 Some background information about the Mande language family

The Mande language family includes about 50-60 languages (depending on whether relatively close varieties are counted as distinct languages or dialects of a single language) whose common ancestor is evaluated as dating back 5-6000 years. The unity of the Mande language family was recognized very early in the history of African linguistics, because of its remarkable typological homogeneity. Its validity as a genetic grouping is uncontroversial, but the nature of its relationship to other language families of Subsaharan Africa remains an open question. The Mande language family was included by Greenberg in the Niger-Congo phylum, but the evidence supporting this decision is rather slim, and the Niger-Congo affiliation of Mande is considered questionable by many specialists – on this question, see Dimmendaal (2011). A simplified version of the current classification of Mande languages is given in Table 1 with the names of the languages mentioned in this article in italics.

For more details on the internal classification of Mande languages, see Vydrin (2009).

3.2 Verbal predication in Mande languages

In Mande languages, verbal predication can be schematized as S (O) V (X).[3] No variation is possible in the linear order of constituents. Predicative constructions with two or more terms encoded in the same way as the patient of typical mono-transitive verbs (so-called 'multiple object constructions') are not possible. In Mande languages, an important characteristic of verbal predication is the existence of paradigms of grammatical words (or clitics), called *predicative markers* in the mandeist tradition, occupying a fixed position immediately after the subject. They express TAM, transitivity and polarity distinctions, either by themselves or in interaction with morphological variations of the verb. The division of labor between predicative markers and suffixal or tonal verb inflection varies greatly from one Mande language to another. For example, in Soninke, the paradigm of predicative markers includes (among others) *má* 'completive, negative', *dà* 'transitivity marker', and the locational copula *wá* (negative *ntá*), which in combination with verbs in the gerundive fulfills the function of incomplete auxiliary – Ex. (2). Verb inflexion is limited to the gerundive suffix *-nV́* (where *V* represents a copy of the preceding vowel), and a tonal alternation by which an entirely L

---

[3] S = subject, O = object, V = verb, X = oblique.

48

South-East Mande	South Mande	Dan
		Guro
		Mano
		etc.
	East Mande	Bisa
		San
		Busa
		etc.
West Mande	Soninke-Bozo	*Soninke*
		Bozo languages
	Bobo-Samogo	Bobo
		Dzuun
		etc.
	Central	*Manding languages*
		Jogo-Jeri
		Kono-Vai
		etc.
	Soso-South-West-Mande	Soso-Jalonka
		South-West Mande languages
		(Mende, Kpelle, Loma, etc.)

Figure 1: The Mande language family (adapted from Vydrin 2009)

contour substitutes for the inherent tonal contour of the verbal lexeme. The slot for predicative markers (immediately after the subject NP) is left empty if one of the following two combinations of values is intended: 'intransitive, completive, positive' or 'intransitive, imperative singular, positive'; in all other cases an overt predicative marker must be present.

(2) Soninke (pers. doc.)

    a. *Ké yúgó xàrá.*
       DEM man study
       'This man studied.'

    b. *Ké yúgó má   xàrà.*
       DEM man CPL.NEG study[L]
       'This man did not study.'

Denis Creissels

   c. *Lémínèn dà í   hàabá ŋàrí.*
      child-D   TR REFL father see
      'The child saw his father.'

   d. *Lémínèn má   í   hàabá ŋàrì.*
      child-D   CPL.NEG REFL father see[L]
      'The child did not see his father.'

   e. *Ó wá   táaxú-nú dàagó-n kànmá.*
      1PL LOCCOP sit-GER   mat-D   on
      'We will sit on the mat.'

   f. *Ń   dà dòròkê-n qóbó sáxà-n   ŋá.*
      1SG TR dress-D buy market-D POSTP
      'I bought a dress at the market.'

   g. *Ń ntá   dòròké qòbò-nò án dà.*
      1SG LOCCOP.NEG dress   buy-GER[L] 2SG for
      'I will not buy a dress for you.'

The rigid constituent order is crucial for the recognition of grammatical relations. In Mande languages, the flagging of core syntactic terms is either totally inexistent, or very marginal. As regards argument indexation, some Mande languages have subject indexes attached or incorporated to the predicative marker (never to the verb itself), the others have no subject indexation at all. A mechanism that can be described as object indexation is found only in some languages in which the third person singular object pronoun has fused with the verb, and in which 'third person singular object' is encoded by a modification of the initial of the verb – see (15) below.

3.3 Copulas in Mande languages

In most Mande languages, non-copular equative or locational clauses (i.e. equative or locational clauses without any explicit predicator) are marginal. Equative predication and locational predicative constructions in Mande languages can be schematized as S COP X. S is an unflagged NP sharing with the subject of verbal predication its obligatoriness and its clause-initial position. X shares with the obliques in verbal predication the following two properties: it follows the predicative element, and its most common form is that of an adpositional phrase, even in equative predication.[4] The position occupied by the copula is comparable

---

[4] In Mande languages, the second term of equative predication is commonly flagged by means of the posposition that marks 'functive' phrases (i.e. the equivalent of *as*-phrases in English)

to that occupied by the verb in intransitive verbal predication, and in terms of possible syntactic operations, copular clauses are not different from intransitive verbal clauses. The only difference is that the copulas have no inflexion, and do not combine with predicative markers, which makes it impossible for copular clauses to express the TAM variations expressed by verb inflection and predicative markers in canonical verbal predication. The use of verbs 'become' (in the case of equative predication) and 'be found' (in the case of locational predication) constitutes the usual strategy to bypass this impossibility. Typically, Mande languages have (at least) two distinct positive copulas: an equative copula and a locational copula. In the negative, they may correspond to two distinct negative copulas, as in Soninke – Ex. (3), but it may also happen that the same negative copula is used in equative and locational predication. As a rule, negative copulas bear no resemblance to their positive counterparts.

(3)   Soninke (pers. doc.)

    a. *Ké yúgó nì tàgé-n ñà yí.*
       DEM man EQCOP blacksmith-D FOC POSTP

       'This man is a blacksmith.'

    b. *Ké yúgó hètí tàgé yì.*
       DEM man EQCOP.NEG blacksmith POSTP

       'This man is not a blacksmith.'

    c. *Múusá wá kónpè-n dí.*
       Moussa LOCCOP room-D in

       'Moussa is in the room.'

    d. *Múusá ntá kónpè-n dí.*
       Moussa LOCCOP.NEG room-D in

       'Moussa is not in the room.'

3.4 Copulas in auxiliary function

As already illustrated in Ex.(2)e and (2)g above, it is common in Mande languages that locational copulas in incompletive auxiliary function combine with verbs, in constructions that lend themselves to a straightforward analysis according to which the copula fulfilling this function occupies the slot for predicative markers. In some Mande languages (for example, in Soninke), the distinction between the use of copulas as predicative markers in verbal predication, and periphrases in

---

in verbal predication – see Creissels (2014).

which the complement of the copula is a nominalized verb, is quite clear-cut, but in some others, this distinction may be more or less problematic. This is not unexpected, since diachronically, periphrases in which the complement of the copula is a nominalized verb are a source from which constructions with copulas in predicative marker function can develop.

4 'See/look' verbs, ostensive predicators, and copulas in Mande languages

Ostensive clauses formally analyzable as imperative clauses headed by a 'see/look' verb are common in Mande languages. Clear cases of copulas originating from the imperative of 'see/look' verbs can be found in Southwestern Mande languages and in the Manding dialect cluster. Moreover, there is some evidence that the locational copula of Soninke might have the same origin.

4.1 Copulas originating from 'see/look' verbs in Southwestern Mande

Southwestern Mande is a group of closely related languages including Mende, Loko, Kpelle, Loma, Zialo, and Gbandi. A common root *káa* 'see' can be reconstructed for Proto-Southwestern-Mande (Valentin Vydrin, pers.com.). In Kpelle, *ka* 'see' is also an ostensive predicator, a locational copula and a progressive auxiliary – Ex. (4). A similar situation is found in Looma (Sadler 2006) and Gbandi (Heydorn 1940-1941).

(4) Kpelle (Westermann & Melzian 1930: 3, 10, 11, 12)

    a. *Ku ŋaloŋ ka bɛlɛi mu.*
        1PL man see house in

        'We saw a man in the house.'

    b. *I seŋkau ka!*
        2SG money OST

        'Here is your money!'

    c. *Ŋaloŋ ka bɛlɛi mu.*
        man COP house in

        'The man is in the house.'

    d. *Nɛni ka pai.*
        woman PROG come

        'The woman is coming.'

4 Copulas originating from the imperative of 'see/look' verbs in Mande

In addition to the coincidence between Kpelle *ka* 'see', *ka* ostensive predicator, *ka* locational copula, and *ka* progressive auxiliary, Westermann observed that the behavior of the NP preceding the locational copula *ka* or the progressive marker *ka* is different from the behavior of subjects in other predicative constructions, and the explanation he put forward is that the subject of the locational copula and the subject of verbs in the progressive construction were originally the object of *ka* 'see' in the imperative: "the *ka* in form No 3 [i.e. in the progressive construction] is perhaps the verb *ka* to see, so that the form really means 'see me coming', 'see him coming', etc." (Westermann & Melzian 1930: 11). In other words, the grammaticalization path analyzed is this paper was explicitly put forward for the first time in Westermann's description of Kpelle.

4.2 Copulas originating from 'see/look' verbs in Manding

Manding is a dialect cluster included in the Central sub-branch of the Western branch of the Mande family. The analysis of Manding as a single macro-language including some relatively divergent dialects, or as a set of distinct although closely related languages, is debatable. Manding varieties share a root for 'see' found as *yé* or *jé*, depending on the individual varieties, and a root for 'look' found as *félé*, *félé*, or very similar forms. As illustrated in Table 1, the use of the imperative of 'look' as an ostensive predicator is pervasive across Manding varieties, and most of them have a similar use of the imperative of 'see'. As will be discussed in §6.1.2, in several Manding varieties, *félé ~ félé* seems to be involved in an incipient grammaticalization process that could lead to the emergence of a new copula, but in all the Manding varieties for which I have the relevant data, copula-like uses of *félé ~ félé* are only sporadic. As regards *yé ~ jé* 'see', there are Manding varieties (for example, Sédhiou Mandinka) in which no grammaticalized use of this verb can be found, but most Manding varieties use *yé ~ jé* either as a locational copula (and incompletive auxiliary), as an equative copula, or both.

Kita Maninka illustrates the case of a Manding variety with the maximum range of grammaticalized uses of *yé ~ jé* 'see'. Note that, in (5), the notation of tone and nasality is phonetic, and only tones contrasting with the tone of the preceding syllable are explicitly noted, which means that *yé* may be transcribed as *yè*, *ɲé*, *ye*, etc. depending on the context.

(5)   Kita Maninka (Creissels 2009: 19, 78, 79, 87, 88)

Denis Creissels

Table 1: Grammaticalized uses of 'see' and 'look' in four Manding varieties

	Sédhiou Mandinka	Dantila Maninka	Bamako Bambara	Kita Maninka
'see'	*jé*	*jé*	*yé*	*yé*
'look'	*félé*	*félé*	*flɛ́*	*félé*
ostensive predicator	*félé*	*félé/jé*	*flɛ́ / yé*	*félé / yé*
equative copula	*mú*	*mú*	*dòn / yé*	*lè / yé*
locational copula	*bé*	*bé/jé*	*bɛ́*	*yé*
incompletive auxiliary	*bé*	*bé/jé*	*bɛ́*	*yé*

a. *Sékù dí tùbabu náni ye kunùn.*
   Sékou CPL European four see yesterday

   'Sékou saw four Europeans yesterday.'

b. *Móngon ɲè!*
   mango.D OST

   'Here is a mango!'

c. *Nénè yé Kìta.*
   cold.D COP Kita

   'It is cold in Kita.'

d. *Kóngò ye n na.*
   hunger.D COP 1SG POSTP

   'I am hungry.' (lit. 'Hunger is in me')

e. *Nònilí yɛ́ ku-jogɛ́ lè di.*
   insult.D COP thing-bad.D FOC POSTP

   'Insult is a bad thing.'

f. *Músa ye ɲó sène-la.*
   Moussa COP millet.D cultivate-INF

   'Moussa cultivates millet.'

g. *Sán nà-dó yè.*
   rain.D come-GER COP

   'Rain is coming.'

Heydorn (1949-1950) describes a similar situation in Manya (a Manding vari-

ety spoken in Liberia), and explicitly states that "Wie im Bandi und verwandten Sprachen ein deutlicher Zusammenhang zwischen 'sein' und 'sehen' besteht, so scheint dies auch im Manya, wo 'sehen' *yẹ* heisst, der Fall zu sein." (Heydorn 1949-1950: 57)

4.3 The locative copula and the verb 'see' in Soninke

The resemblance between the Soninke verb *wàrí* (or *ŋàrí, ŋèrí*) and the locational copula *wá* (also used in incompletive auxiliary function) is not very great, and might be due to mere chance. However, evidence of a possible etymological link is provided by the data of Azer, a now-extinct Soninke variety. Monteil (1939: 42-44) mentions the existence of variants of the locational copula/incompletive auxiliary such as *wari, war, wri,* and explicitly states that he considers this copula/auxiliary as a grammaticalized form of 'see'.

5 Comparison with the grammaticalization of 'see' in Arabic

In this section, I summarize the data on the grammaticalization of 'see' in Arabic languages that have been presented and analyzed in detail by Taine-Cheikh (2013), emphasizing the commonalities and differences with Mande languages that are directly relevant to the topic of this article.[5] An important particularity that distinguishes the predicative system of most Arabic varieties from that of most Mande languages is the systematic use of equative or locational predicative constructions including no overt predicators – (6).

(6) Classical Arabic

    a. *al-waladu ṣaġīru-n.*
       DEF-boy   small-INDEF
       'The boy is small.'

    b. *al-waladu fī l-madrasat-i.*
       DEF-boy   in DEF-school-GEN
       'The boy is at school.'

---

[5] In addition to Catherine Taine-Cheikh's 2013 article, this paper has benefited from the discussions I had with her about the Arabic data analyzed in her article, and I want to express my gratitude to her.

Denis Creissels

The grammaticalization of *raʔā* 'see', in particular in its imperative form, is a very common phenomenon across Arabic varieties. This verb is preserved in literary Arabic, but in most modern Arabic varieties, only grammaticalized forms of *raʔā* have subsisted, and the verb most commonly used in the sense of 'see' is *šāf*. A detailed analysis can be found in Taine-Cheikh (2013). In the present article, I concentrate on the aspects that are directly relevant to the current discussion. Plain ostensive predicators cognate with *raʔā* are not very common across Arabic varieties. However, Ḥassāniyya (the variety spoken in Mauritania) and a few other varieties illustrate this possibility – (7) & (8).[6]

(7) Ḥassāniyya (Taine-Cheikh 2013)
ṛâʕi xṛûv!
OST lamb
'Here is a lamb!'

(8) Yâfiʕ, Yemen (Vanhove 2010: 336-337)
raʕ ar-rābʕeh
OST DEF-jug
'Here is the jug!'

Particles expressing not only simultaneity ('right now'), but also various modal or discursive values derivable from an original ostensive meaning, constitute the commonest outcome of the grammaticalization of *raʔā* across Arabic varieties. Their contribution to the meaning of the clause can be variously rendered in English as 'indeed', 'really', 'certainly' 'don't you see that...?', 'and then', 'this is a fact', 'you must know that...', 'I remember you that...' etc. To the best of my knowledge, this grammaticalization path has no equivalent in Mande languages. Equative or locational clauses including an element whose etymon is the imperative of *raʔā* 'see' are common across Arabic varieties, but in most cases, as illustrated by (9) for a variety from the South of the Arabic Peninsula, it would not be correct to identify this element as a copula, since it is syntactically optional, its presence implies a marked modal or discursive value, and crucially, it can be added with the same value to verbal clauses. Although this is not easy to reflect in the translation of isolated examples, it is clear from the comments in the original source that, in this Arabic variety, *raˀ* cannot be analyzed as an

---

[6] Interestingly, an ostensive predicator *ša* originating from *šāf* (the verb most commonly used in the sense of 'see' in modern Arabic varieties) is attested in Syrian Arabic (Stowasser & Ani 1964: 115).

integral part of a particular type of predicative construction, and is rather an op-
tional particle used to emphasize a precise fact or a sudden appearance and to
express the reason behind something, or an event's consequences.

(9) Datînah Arabic (Landberg 1909: 485, 486, 488)

    a. *raˁ=nī ˁawaḍ.*
       raˁ=1SG ˁawaḍ

       ('Je suis ˁAwaḍ, moi.') 'Me, I am ˁAwaḍ.'

    b. *raˁ=ak fi arḍ   ˁöleh.*
       raˁ=2SG in country ˁOlah

       ('[…] c'est que tu [es] dans le pays des ˁOlah.') 'That's because you
       [are] in the country of the Olah.'

    c. *raˁ em=maṭar y-ehđil.*
       raˁ DEF=rain 3M.INCPL-drizzle.SG

       ('Voilà que la pluie tombe fine.') 'There goes the rain drizzling.'

A plain copula originating from the imperative of *raʔā* can only be found in
Algerian Arabic, and more precisely in the variety spoken in Algiers. This was
already observed by Cohen (1912: 252), and Boucherit (2002) confirms that, in the
equative and locational clauses of Algiers Arabic, *ra* does not express the values
carried by its cognates in most other Arabic varieties, and can be analyzed as the
suppletive present form of a copula whose past form is *kān*.

(10) Algiers Arabic (Boucherit 2002: 62)
    ra=ni   fi=l=kuzina.
    COP=1SG in=DEF=kitchen

    ('Je suis dans la cuisine.') 'I am in the kitchen.'

6 From the imperative of verbs 'see/look' to copulas

6.1 The grammaticalization path SEE/LOOK$_{imper}$ > OSTENSIVE PREDICATOR > COPULA

This grammaticalization path is strongly suggested by the Mande data, since the
creation of ostensive predicators from the imperative of 'see' or 'look' is very
common in Mande languages, and in the Mande languages that have copulas
cognate with a verb 'see', the same form is used as an ostensive predicator, and
has no other use that could constitute an intermediate stage in this grammatical-
ization path.

Denis Creissels

6.1.1 SEE/LOOK$_{imper}$ > OSTENSIVE PREDICATOR

As already mentioned in §1, the grammaticalization of the imperative or verbs 'see' or 'look' as ostensive predicators is common cross-linguistically. The creation of ostensive predicators from the imperative of 'see' or 'look' boils down to the routinization of an ostensive use of the imperative of 'see' or 'look'. In this use, *See/look at X!* is not interpreted in its literal meaning of an incitement to see/look at the referent of X, but as expressing awareness of the presence of the referent of X in the speech situation. Since uttering *See X!* or *Look at X!* in their literal meaning entails the presence of the referent of X, the routinization of the ostensive use of the imperative of a 'see/look' verb can be viewed as the semanticization of a pragmatic entailment. At an early stage of the evolution, there is no formal manifestation of the development of an ostensive reading of the imperative of 'see' or 'look', but subsequent changes may introduce formal distinctions. For example, in French, it is obvious that *Me voici!* comes from a construction whose equivalent in Modern French would be *Vois-moi ici!*, but synchronically, the position of the object index (which in Modern French cannot precede the verb in the imperative positive), and the coalescence of *vois + ici* into *voici*, distinguish the ostensive predicator from the imperative of 'see'. However, the persistence of the ambiguity is possible too. For example, in Mandinka (and other Manding varieties), *Ń félé* |1SG look| is ambiguous between its literal meaning 'Look at me!' and the ostensive reading 'Here I am!'.

6.1.2 OSTENSIVE PREDICATOR > COPULA

Ostensive predicators entail meanings typically expressed by copulas: identification of a referent, and presence of a referent at some place. They differ from copulas in two crucial respects:

(a) the argument of an ostensive predicator must be located in the speech situation;

(b) ostensive clauses express a particular type of speech act (drawing the addressee's attention to an obvious fact) distinct from plain assertion, and consequently do not lend themselves to operations such as negation, questioning, or relativization.

Consequently, the relaxation of these constraints, manifesting the loss of the deictic component of ostensive predication and the reinterpretation of ostensive

clauses as plain assertive clauses, is crucial in the evolution from the status of ostensive predicator to that of copula.

Interestingly, in Mande languages, in addition to copulas analyzable as resulting from the grammaticalization of an ostensive marker, it is possible to find ostensive markers that still cannot be analyzed as having grammaticalized into copulas, which however already occur more or less sporadically in contexts implying the weakening of their deictic component or the bleaching of their particular illocutionary force.

For example, contrary to *yé ~ jé* 'see', which has become a copula in many Manding varieties, I am aware of no Manding variety with a true copula cognate with *félé ~ félé* 'look'. However, in Kita Maninka, ostensive clauses lend themselves to relativization, which implies the cancellation of the particular illocutionary force normally carried by *hélé* in its use as an ostensive predicator – Ex. (11).

(11) Kita Maninka (Creissels 2009: 82)

    a. *Wórì   hele!*
       money.D look

       literal meaning 'Look at the money!' (imperative), can also be interpreted as 'Here is the money!' (ostensive reading)

    b. *Wórì   mín hèle, ŏ   tà!*
       money.D REL look DEM take

       'Take the money that is here!' lit. 'Look$_{imper}$ at which money, take that!'

Note that my consultant for Kita Maninka accepted this use of *hélé* in ostensive predicator function, but rejected other manipulations on ostensive clauses (for example, questioning) which would have been expected to be accepted if ostensive *hélé*-clauses had been fully reinterpreted as plain assertive clauses. Similarly, in Mandinka, (12) illustrates the sporadic occurrence of *félé* in contexts incompatible with the deictic value normally implied by *félé*: this sentence was extracted from a story about a village very far from the place where the story was recorded, which means that a plain locative copula could substitute for such an occurrence of the ostensive predicator without any difference in meaning.

(12) Mandinka (Creissels & Sambou 2013: 158)
    Jálájúw-òo félé   lòo-ríŋ   jĕe   háhí bĭi.
    jala_tree-D look stand-RES there even today

    'Up to the present day, a *jala*-tree stands there.' (lit. 'Look at a *jala*-tree

Denis Creissels

standing there even today!')

Similar observations can be made about Soninke *háyí* 'look' and Bozo *xai* 'see'. In Soninke, the imperative of *háyí* is routinely used as an ostensive predicator, but it is also sporadically found in contexts in which its deictic component or its special illocutionary force cannot be maintained, which points to an incipient process whose outcome could be the creation of a new locational copula. For example, my Soninke consultant accepts the use of *háyí* in interrogative clauses such as those in (13), which force a reading of *háyí* as a mere locational copula.

(13)   Soninke (pers. doc.)

   a.   *À   háyí màní ñàa-nà?*
        3sg look what do-GER$^{\text{L}}$
        'What is he doing?' (lit. 'Look$_{\text{imper}}$ at him doing what?')

   b.   *À   háyí sòxò-nó   bà?*
        3sg look cultivate-GER Q
        'Is he cultivating?' (lit. 'Look$_{\text{imper}}$ at him cultivating?')

   c.   *Kó   háyí sòxò-nò?*
        who look cultivate-GER$^{\text{L}}$
        'Is he cultivating?' (lit. 'Look$_{\text{imper}}$ at whom cultivating?')

As regards Bozo, Blecke's description of Tigemaxo suggests that, in this Bozo variety, there is a similar relationship between the locational copula *ga* (which incidently might well be cognate with the root **káa* reconstructed for Southwestern Mande) and *xai* 'see'. Blecke (1996: 206 et seq.) not only mentions an ostensive use of the imperative of *xai* 'see', he also repeatedly insists on the possibility of substituting *xai* for the locational copula *ga*.

6.2 The grammaticalization path SEE/LOOK$_{\text{imper}}$ > MODAL/DISCURSIVE PARTICLE > COPULA

The grammaticalization path discussed in §6.1 is consistent with the Mande data, but it does not provide a satisfactory explanation of the Arabic data, since plain ostensive predicators cognate with *raʔā* 'see' are not very common in Arabic. Given the pervasiveness of modal or discursive particles cognate with *raʔā*, it seems more plausible that the copula *ra* found in Algerian Arabic results from the reanalysis of such a particle in equative and locational constructions that initially included no overt predicator.

Across Arabic varieties, irrespective of the presence of an ostensive predicator cognate with *raʔā*, modal or discursive particles cognate with *raʔā* can be added to equative or locational clauses including no overt predicator exactly in the same way as they are added to verbal clauses, with the same semantic implications, as already illustrated by Ex. (9) above. Ex. (14) provides an additional illustration.

(14)   Ḥassāniyya (Taine-Cheikh 2013)
        (a)ṛâ=ni   merîḍ.
        (a)ṛā=1SG sick.M.SG
        'I remind you that…, remember that I am sick.'

The use of *ra* as a plain copula in Algerian Arabic (illustrated by Ex. (10) above) is most probably due to a process of semantic bleaching that affected *ra* in a construction originally similar to (14), leading to its reanalysis as a plain copula. This hypothesis is supported by the fact that, in Algerian Arabic, *ra* occurs in equative and locational clauses without any particular semantic or discursive implication, but is still found in verbal clauses with values similar to those found in other Arabic varieties.

7 An alternative grammaticalization path from 'see' verbs to copulas

This discussion of the grammaticalization of 'see/look' verbs into copulas would not be complete if another possible grammaticalization path from 'see' verbs to copulas were not evoked and confronted with the Mande data. The point is that, cross-linguistically, as already mentioned in the introduction, the translation equivalents of 'see' may be polysemous verbs expressing the meanings commonly expressed in English as 'find' or 'get'. This means that some forms of such verbs may be found with meanings such as 'is found' or 'is available', i.e. meanings very close to those typically expressed by locational copulas. Consequently, 'see' verbs and copulas can be diachronically related in at least three different ways:

(a) SEE$_{imper}$ /LOOK$_{imper}$ > OSTENSIVE PREDICATOR > COPULA

(b) SEE$_{imper}$ /LOOK$_{imper}$ > MODAL/DISCURSIVE PARTICLE > COPULA

(c) (IS_SEEN~)IS_FOUND > LOCATIONAL COPULA

Denis Creissels

Given the rich verbal morphology of Arabic, and in particular the clear-cut distinction between subject and object indexation, there can be no doubt that the grammaticalized uses or 'see' described by Taine-Cheikh (2013) developed from the imperative form of this verb. For example, in Algiers Arabic, the subject of the suppletive form of the copula resulting from the grammaticalization of the imperative of 'see' is indexed by the suffixes used in canonical verbal predication to index objects, which supports the hypothesis that the subject of the present form of the copula is a former object that has been reanalyzed. By contrast, for Mande languages, it is necessary to discuss the evidence supporting the hypothesis that, as assumed in the previous sections, copulas cognate with 'see' verbs in Mande languages were created according to Path (a) rather than Path (c).

In the case of Southwestern Mande languages (see §4.1), conclusive evidence can be found in the systems of consonant alternations affecting the initial of nouns and verbs. The point is that, in Southwestern Mande languages, the boundary between object NP's and verbs in transitive predication is characterized by sandhi phenomena that do not occur at the boundary between subject NP's and verbs in intransitive predication. Consequently, if a copula had been created according to Path (c), its subject would have been already a subject in the source construction involving a verb 'be found', and it would be therefore expected to behave as a normal subject with respect to its interaction with the initial consonant of the verb. By contrast, if a copula has been created according to Path (a), its subject is historically a reanalyzed object. Consequently, the subject of a copula created according to Path (a) can be expected to retain the type of interaction with the initial of the verb which normally characterizes objects, and this is precisely what can be observed.

For example, in Kpelle, in intransitive constructions in which the verb is immediately preceded by its subject, the third person singular pronoun is realized as a distinct segment, and the initial consonant of the verb does not change – ex. (15)a-b. In transitive constructions, with the object NP in immediate preverbal position, the third person singular object manifests itself by a change in the initial consonant (and the tone) of the verb – ex. (15)c-f, and the same phenomenon is observed with the third person singular subject of the copula cognate with 'see' – ex. (15)g-h.

(15)   Kpelle (Westermann & Melzian 1930: 4, 11, 21)

    a.   *Kú pâ.*
       1PL come
       'We came.'

b. *È   pà.*
   3SG come
   'He, she, it came.'

c. *Loa   tíe!*
   hole dig
   'Dig a hole!'

d. *Díe!*
   3SG.dig
   'Dig it!'

e. *Dì   kú   kâ.*
   3PL 1PL see
   'They saw us.'

f. *Dí   gà.*
   3PL 3SG.see
   'They saw him.'

g. *Kú ká   bɛ.*
   1PL COP here
   'We are here.'

h. *Gà   bɛ.*
   3SG.COP here
   'He is here.'

This is certainly why Westermann, who was the first to mention the imperative of 'see' as a plausible origin of copulas and incompletive auxiliaries in Mande languages, did not hesitate in putting forward this analysis of the Kpelle data. Things are less straightforward in Manding, since in Manding languages, the distinction between subjects and objects has no morphological correlate. However, evidence supporting the choice of path (a) can be found in Manding too. A first observation is that, in Manding, the imperative positive is the only tense in which the verb does not combine with an overt TAM marker (either suffixed to the verb or immediately following the subject). Consequently, the fact that the Manding copulas cognate with 'see' show no trace of TAM marking supports the hypothesis that they originate from an imperative form. A second observation is that there is no reason why a copula resulting from the semantic bleaching of a verb 'be found, be available' should not have a negative form created in a parallel way from the negative form of the same verb. By contrast, the meaning carried by ostensive predicators makes them incompatible with negation. Consequently, the

Denis Creissels

fact that no negative copula cognate with 'see' is found in Manding supports the hypothesis that the Manding copulas cognate with 'see' were created according to path (a).

8 Conclusion

In this article, I have tried to show that, in the Mande language family, clear cases of copulas cognate with 'see/look' verbs are found at least in Southwestern Mande languages and in the Manding dialect cluster, and I have discussed evidence that they did not develop from the semantic bleaching of forms originally meaning 'is seen/found', but from the routinization of the ostensive use of the imperative of 'see/look'. By comparing the Mande data with the Arabic data provided by Taine-Cheikh (2013), I have tried to show that this is however not the only possibility for imperatives of verbs 'see/look' to grammaticalize into copulas. In the Arabic varieties in which the imperative form of 'see' has become a plain copula, the most plausible explanation is that a modal/discursive particle resulting from the grammaticalization of the imperative of 'see' has undergone a process of semantic bleaching in the context of an equative or locational predicative construction that initially included no overt predicator.

Abbreviations

COMP	complementizer		LOCCOP	locational copula
COP	copula		M	masculine
CPL	completive		OST	ostensive predicator
D	default determiner		PL	plural
DEM	demonstrative		POSTP	multipurpose postposition
EQCOP	equative copula		PROG	progressive
FOC	focus marker		Q	interrogative particle
GEN	genitive		REFL	reflexive
GER	gerundive		REL	relativizer
INDEF	indefinite		RES	resultative
INF	infinitive		SG	singular
NEG	negative		TR	transitivity marker.
L	replacive morphotoneme 'low'			

References

Blecke, Thomas. 1996. *Lexikalische Kategorien und grammatische Strukturen im Tigemaxo (Bozo, Mande).* Cologne: R. Köppe.

Boucherit, Aziza. 2002. *L'arabe parlé à Alger. Aspects sociolinguistiques et énonciatifs.* Paris/Louvain: Peeters.

Cohen, Marcel. 1912. *Le parler arabe des juifs d'Alger.* Paris: Campion.

Creissels, Denis. 1981. L'étymologie des prédicatifs d'identification des parlers bambara et jula: *yé* et *dòn~dò~lò. Mandenkan* 1. 3–10.

Creissels, Denis. 2009. *Le malinké de Kita.* Köln: Rüdiger Köppe.

Creissels, Denis. 2014. Functive phrases in typological and diachronic perspective. *Studies in Language* 38(3). 605–647.

Creissels, Denis & Pierre Sambou. 2013. *Le mandinka: Phonologie, grammaire, textes.* Paris: Karthala.

Dimmendaal, Gerrit. 2011. *Historical linguistics and the comparative study of African languages.* Amsterdam/Philadelphia: Benjamins.

Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization.* Cambridge: Cambridge University Press.

Hellwig, Birgit. 2011. *A grammar of Goemai.* Berlin/New York: Mouton De Gruyter.

Hengeveld, Kees. 1992. *Non-verbal predication.* Berlin/New York: Mouton De Gruyter.

Heydorn, Richard. 1940-1941. Die Sprache der Bandi im nordwestlichen Liberia. *Zeitschrift für Eingeborenensprachen XXXI* 81-114. 188–217.

Heydorn, Richard. 1949-1950. Das Manya: Ein in Liberia gesprochener Mandingodialekt. *Zeitschrift für Eingeborenensprachen* XXXV. 25–44.

Jaggar, Philip J. 2001. *Hausa.* Amsterdam/Philadelphia: John Benjamins.

Landberg, comte Carlo de. 1909. *Sur les dialectes de l'Arabie Méridionale II. Datînah.* Vol. 2. Leiden: Brill.

Monteil, Charles. 1939. *La langue azer, d'après les documents recueillis par Th. Monod et D. Brosset.* Paris: Larose.

Newman, Paul. 2000. *The Hausa language: An encyclopedic reference grammar.* New Haven/ London: Yale University Press.

Petit, Daniel. 2010. On presentative particles in the Baltic languages. In Nicole Nau & N. Ostrowski (eds.), *Particles and connectives in Baltic*, 151–170. Vilnius: Vilniaus Universitetas, Acta Salensia 2.

Pustet, Regina. 2003. *Copulas.* Oxford: Oxford University Press.

Denis Creissels

Rubin, Aaron D. 2005. *Studies in Semitic grammaticalization.* Winona Lake, Indiana: Eisenbrauns.

Sadler, Wesley. 2006. A complete analysis of the Lɔɔma language, Interior Liberia, West Africa. *Mandenkan* 42. Edition of a manuscript written in 1949, 1–109.

Stowasser, Karl & Mokhtar Ani. 1964. *A dictionary of Syrian Arabic: English-Arabic.* Washington DC: Georgetown University Press.

Suzuki, Hiroyuki. 2012. Multiple usages of the verb snang in Gagatang Tibetan (Weixi, Yunnan). *Himalayan Linguistics* 11(1). 1–16.

Taine-Cheikh, Catherine. 2013. Grammaticalized uses of the verb raʔā in Arabic: A Maghrebian specificity? In M. Lafkioui (ed.), *African Arabic: Approaches to dialectology*, 121–159. Berlin/ New York: De Gruyter.

Tröbs, Holger. 2003. On the origin of some predicative markers in imperfective constructions in Manding. *Mandenkan* 38. 1–14.

Vanhove, Martine. 2010. Deixis, information structure and clause linkage in Yafiˁ Arabic (Yemen). In I. Bril (ed.), *Clause linking and clause hierarchy. Syntax and pragmatics*, 333–354. Amsterdam/Philadelphia: Benjamins.

Vydrin, Valentin. 2009. On the problem of the Proto-Mande homeland. *Вопросы языкового родства – Journal of Language Relationship* 1. 107–142.

Welmers, Williams E. 1974. *Manya.* Typescript. Los Angeles.

Westermann, D. & H. J. Melzian. 1930. *The Kpelle language in Liberia.* Berlin: Dietrich Reimer /Ernst Vohsen.

Chapter 5

Multiple argument marking in Bantoid: from syntheticity to analyticity

Larry M. Hyman
University of California, Berkeley

> "All evidence points to the... hypothesis that such [analytic] languages are the logically extreme analytic developments of more synthetic languages which because of processes of phonetic disintegration have had to reexpress by analytical means combinations of ideas originally expressed within the framework of the single word."
> (Sapir 1933[1949]: 18-19)

This paper addresses the mechanisms of change that lead from syntheticity to analyticity in the Bantoid languages of the Nigeria-Cameroon borderland area. I address the different strategies that are adopted as these languages lose applicative "verb extensions" found elsewhere in Bantu and Niger-Congo. I show that although historical recipient, benefactive, and instrumental applicative marking on verbs allowed multiple object noun phrases (send-APPL chief letter, cook-APPL child rice, cut-APPL knife meat), they have been replaced by adpositional phrases and/or serial verb constructions in all branches of Bantoid. I map out the different analytic strategies that have been adopted and reconstruct the original verbal, nominal and pronominal sources of the different grammaticalization processes. Of particular interest is the development of a recipient/benefactive preposition 'to, for' from the word for 'hand' and a comitative/instrumental preposition 'with' from a third person plural pronoun.

Larry M. Hyman

1 Establishing a Proto-Bantoid synthetic system

The general issue I address in this paper is how to account for the alternative grammaticalization strategies adopted as a highly synthetic (agglutinative) language develops towards analyticity. My focus will be on the multiple pathways that can be observed between the inherited head-marking verb structures of Proto-Bantoid and the more analytical structures found in most of the daughter languages spoken today.[1] As noted by Dimmendaal (2000: 187-8), among others, extensive head-marking occurs in at least some languages in all four of Greenberg's (1963) macro-stocks: Nilo-Saharan, Afro-Asiatic, "Khoisan", and, as exemplified in (1), multiple branches of Niger-Congo.

(1)   a. Seereer ["Atlantic" branch; Senegal]
          a   up-t-ik-t-ir-oox-k-a   *apeel*
          3p.SM bury-REV-GOAL-INST.APPL-REC-REFL-FUT-INFL shovels
          'they'll go unbury each other with shovels' (John Merrill, pers.comm.)

      b. Cicipu [Plateau/Central Nigerian branch; Nigeria)]
          zzá   nnà ù-tób-ìl-ìs-ìs-u-wò-wò-nò=mu
          person REL 3SG-cool-PL-CAUS-CAUS-V-ANTICAUS-APPL-PERF=1SG
          sháyì
          tea
          'the person who has caused tea to become cooled down in a forceful and iterative fashion for me' (McGill 2009: 209)

      c. Moro [Kordofanian; Sudan]
          oːwːa   g-ubəð-i-tʃ-ən-ə-ŋó
          woman SM.CL-run-CAUS-APPL-PASS-PERF-3SG.OM
          'the woman was made to run away from him' (Rose 2013: 49)

      d. Kinande (Bantu) [Bantoid subbranch; Democratic Republic of Congo]
          tu-né-mu-ndi-syá-tá-sya-ya-ba-king-ul-ir-an-is-i-á=ky-ô
          we-TNS/ASP COMPLEX-them-close-REV-APPL-REC-CAUS-CAUS-INFL=it
          'we will make it possible one more time for them to open it for each other'
          (Philip Mutaka, pers.comm to Nurse & Philippson 2003: 9)

---

[1] The term "Bantoid" is used in two senses in the literature. First, it refers to a node in the Niger-Congo family tree that includes both Bantu and non-Bantu languages; second, it refers to these latter non-Bantu languages themselves. In most of my discussion I will be citing such Bantoid languages which have evolved significantly further than their agglutinative Bantu cousins.

5 Multiple argument marking in Bantoid: from syntheticity to analyticity

The example in (1)d is of most relevance to the present study, as it illustrates several of the most common Bantu derivational suffixes known as verb extensions: causative, applicative etc. As I noted in Hyman (2003), the following valence-marking verb extensions tend to occur in the order Causative-Applicative-Reciprocal-Passive (C-A-R-P) in what I shall refer to as Canonical Bantu (CB):

	Causative	*Applicative*	*Reciprocal*	*Passive*	*(C-A-R-P)*
Proto-Bantu	-ɪc-	-ɪl-	-an-	-ʊ-	
Shona	-is-	-il-	-an-	-w-	
Makua	-ih-	-il-	-an-	-iw-	
Chichewa	-its-	-il-	-an-	-idw-	

Of the above extensions, the causative and applicative add valence, while the reciprocal and passive decrease valence. In considering what has occurred within the related Bantoid languages, I will be most concerned with how these languages compensate for the loss of valence-adding extensions, e.g. the applicative, which has multiple functions in CB, illustrated from Chichewa in (2).

(2) tum-ir- (send+applicative)
    'send for (s.o.), send to (s.o.), send with (sth.), send to (some place),
    BENEFACTIVE RECIPIENT   INSTRUMENT   LOCATIVE
    send for (some reason)'
    CIRCUMSTANCE

While CB languages are highly agglutinative, Northwest (NW) Bantu languages often have simpler structures, even extreme analyticity, as in Nzadi, a "Narrow Bantu" language spoken in the Democratic Republic of Congo which has lost valence-related suffixes, replacing them with the following analytic structures (Crane, Hyman & Tukumu 2011):

(3) a. causative:
        yà ó   líŋ   mwàán kè   lîì
        2SG PST want child SBJV cry
        'you made the child cry' (lit. you wanted that the child cry)

Larry M. Hyman

b. benefactive:
   bì ó súm mwàán òŋkàáŋ
   1pl PST buy child book
   'we bought the child a book' (double object)

c. recipient:
   bì ó pé mwàán fùfú
   1pl PST give child fufu
   'we gave the child fufu'

d. = *bì ó pé fùfú kó mwáàn*
   1pl PST give fufu to/for child
   'we gave fufu to the child'

e. instrument:
   ndé ó píŋ ntsúr tí mbyě
   3SG PST cut meat with knife
   'he cut meat with a knife'

f. circumstance:
   ndé á sâl sám ꜛé ndzíì
   3SG PRES work reason of money
   'he is working for money'

As can be seen, the above structures represent four different strategies for dealing with the loss of verb extensions: periphrasis (3a), unmarked double objects (3b,c), adpositions (3d,e), nominal constructions (3f). Missing in Nzadi is a fifth strategy, serial verb constructions, which will be become central in the discussion of the Bantoid developments discussed below.

> While the historical changes that have taken place in Nzadi definitely give it a 'non-Bantu' feel, it is clear that Nzadi derives from a quite canonical Bantu type. Nzadi 'feels' like a simplified Bantu language rather than a Bantu language which has developed West African Benue-Congo characteristics (e.g. Nzadi does not have the 'serial verb constructions' attested in Cameroon). (Crane, Hyman & Tukumu 2011: 3-4)

In this study I will assume that (pre-) Proto-Bantoid was like Proto-Bantu (PB) in having verb extensions (causative, applicative, etc.), multiple objects, and very few—perhaps even only one—adposition.[2] This naturally raises the question of

---

[2] Only *na* 'with, and' can be confidently reconstructed for PB and early Niger-Congo.

why synthetic head-marking languages like Kinande and Chichewa become analytic languages like Nzadi? That is, why do such languages undergo such a dramatic change of typology? As far as I know, there have been three proposals in the literature: The first is that the affixal morphology is lost through "processes of phonetic disintegration" (cf. the Sapir 1933[1949] quote at the beginning of the paper). Known as "erosion" (Heine & Reh 1984: 21-28) or "phonological attrition" (Lehmann 1985: 4) in the grammaticalization literature, the change in typology is an innocent by-product of natural sound changes, particularly phonetic weakening and loss at word edges: "The opposite historical directionality towards analyticity proceeds mostly by way of erosion and loss of phonological and morphological substance." (Güldemann 2011: 129) The second explanation attributes the development of analyticity to contact and imperfect learning by L2 speakers, ultimately leading to creolization.

> ... we [should] at least consider that these [analytic] languages' grammars were incompletely acquired at some point in their history. This is a known cause of analyticity, whereas the idea of generations of first-language speakers 'dropping' *all* of the affixes used by previous ones is peculiar at best and implausible at worst. (McWhorter 2011: 226)

Table 1: Syllable length of verb stems in Chichewa vs. Nzadi

	1σ	2σ	3σ	4σ	$5\sigma+$	Totals
Chichewa	30 (1.4%)	650 (31%)	906 (43.2%)	477 (22.8%)	22 (1.1%)	2095
Nzadi	291 (83.9%)	51 (14.7%)	2 (0.6%)	1 (0.3%)	0 0%	347

In McWhorter's account, phonetic erosion would have played little, if any role, in the development of the type of "radical analyticity" seen in Nzadi. The third account proposed in Hyman (2004) and subsequent papers is that morphology was lost as a result of imposing templatic constraints on stems (in this case, verb stems, which consist of a root + suffixes). Whereas PB did not have such limitations, the changes which took place included imposing a strong-weak structure highlighting the stem-initial CV and maximal size constraints on stems, which limited the ability of verb roots to occur with derivational suffixes. As will have been noted in (3), words are very short in Nzadi. Compare in Table 1, the number of verb stems having one to five syllables in Chichewa vs. Nzadi.[3] As seen, the

---

[3] The numbers from Chichewa are based on a lexical database of 5,862 entries in Filemaker Pro™

Larry M. Hyman

vast majority of Nzadi verbs are monosyllabic, with most of the bisyllabic verb consisting of relic derived forms, e.g. *dɛf* 'borrow' → *dɛfsa* 'lend' (< 'cause to borrow'). That monosyllabicity is the endpoint of a gradual process of limiting stem size can be seen from the following continuum in NW Bantu:

(4)   a.   four (~five) syllable maximum in Yaka (Hyman 1998), Bobangi (Whitehead 1899) Punu (Fontaney 1980, Blanchon 1995)

     b.   three (~four) syllable maximum in Koyo (Hyman 2004), Eton (Van de Velde 2008)

     c.   three-syllable maximum in Tiene (Ellington 1977), Basaa (Lemb & de Gastines 1973, Hyman 2003), Kukuya (Paulian 1975)

     d.   two (~three) syllable maximum in Mankon [Grassfields Bantu (GB)] (Leroy 1982)

     e.   one (~two) syllable maximum in Nzadi (Crane, Hyman & Tukumu 2011)

However, it is not just maximal stem size that is innovated, but also templatic prosodic constraints. This is most clearly seen in Tiene, which allows a maximally trisyllabic stem having the following properties (Ellington 1977, Hyman 2010):

(5)   a.   five stem shapes: CV, CVV, CVCV, CVVCV, CVCVCV

     b.   in the case of $C_1V_1C_2V_2C_3V_3$:
        i. C_2 must be coronal
        ii. C_3 must be non-coronal
        iii. C_2 and C_3 must agree in nasality
        iv. V_2 is predictable (with few exceptions)

The effects of prosodic constraints on morphology can be quite dramatic. Thus, the coronal + non-coronal constraint on C_2 and C_3 can result in infixation, as in (6b,c).

(6)   a.   CB -ik- 'stative': ból-a 'break' → ból-ek-ɛ 'be broken'

     b.   CB -is- 'causative: láb-a 'walk' → lásab-a 'cause to walk'

     c.   CB -il- 'applicative: bák-a 'reach' → bálak-a 'reach for'

While McWhorter's and my explanations have in common that more needs to be involved than phonetic erosion, it is unlikely that the innovated infixation

---

based on Scott & Hetherwick (1970) and tone-marked by Al Mtenje. The much smaller Nzadi lexicon of 1,035 entries can be found in (Crane, Hyman & Tukumu 2011: 281–298).

process in (6b,c) would have resulted from "incomplete acquisition". Instead, as I argued in Hyman (2004), Niger-Congo languages become analytic by the stages outlined in (7).

(7) a. start with a full set of (stacked) verb extensions (causative, applicative etc.) and multiple objects

b. size (and other prosodic) constraints come to be imposed: $4\sigma > 3\sigma > 2\sigma$ maximum

c. such maximality constraints result in longer verbs not being able to take extensions

d. to accommodate these verbs, analytic alternatives are favored (and created, if not preexistent)

e. these alternatives come to be used even with shorter verbs, with extensions becoming less favored

f. former valence-related extensions take on new, especially aspectual functions (e.g. various pluractional meanings), or drop out

Turning to Bantoid, as an example of (7f), causative *-sə* has become an iterative extension in Bangwa [GB, Bamileke; Cameroon] (Nguendjio 1989: 243) in (8), while several of the inherited verb extensions have taken on pluractional meanings in Kejom [GB, Ring subgroup; Cameroon] (Jisa 1977, Akumbu 2008) in (9).

(8) sò   'laver'   → sò-sə   'laver plusieurs fois'
     fák   'tourner'   → fák-sə 'tourner plusieurs fois'
     cí-   'casser'   → cí-sə   'casser plusieurs fois'
     yàʔ   'couper'   → yàʔ-sə 'couper plusieurs fois'
     ghɛ   'partager'   → ghɛ-sə 'partager plusieurs fois'

(9) a. tsɔʔɔ 'jump'
       tsɔʔ-mə 'jump one after the other'
       tsɔʔ-kə 'jump time and again'
       tsɔʔ-lə 'jump across things'
       tsɔʔ-tə 'jump gently' (= attenuative)

   b. dì 'cry, cackle'
      dì-mə 'lots of children crying'
      dì-kə 'cry time and again'
      dì-lə 'lots of chickens cackling'

    c. zhwí 'kill'
        zhwí-tə 'kill one by one, bit by bit'
        zhwí-lə 'kill lots of people, one after the other'
    d. sù 'stab'
        sù-tə 'stab lots of things one by one, or one thing many times'
        sù-lə 'stab with lots of things at one time'

To summarize, major changes transformed an originally agglutinative proto language into much more analytic daughter languages in some of NW Bantu and Bantoid. As a result, non-Bantu Bantoid languages differ considerable from CB, as summarized in (2).

Table 2: Comparison of Canonical Bantu with Non-Bantu Bantoid

	Canonical Bantu	Non-Bantu Bantoid
phonology	minimum word = 2 syllables	maximum stem =mostly 2~3 syllables
morphology	highly synthetic, agglutinative	less so, gradual move towards analyticity
verb extensions	many, mostly marking valence	few, mostly marking aspect
unmarked objects	multiple	at most two, ultimate limitation to one per verb
object marking	head marking on verb	various prepositions and/or serial verbs [diversity!]
ditransitive verbs	a few (*pá 'give')	few or none

Having established that Proto-Bantoid had a range of verb extensions, I now consider the structures which come to replace them in the daughter languages.

2 Analytic replacements of the lost Proto-Bantoid synthetic structure

In this section I examine what has replaced the verb extension system inherited by languages in the Bantoid area of Cameroon. In order to control the study, I focused exclusively on the marking of valence by head marking on the verb, specifically benefactives ('for someone'), recipients ('to someone') and instruments ('with something'). As will be seen, Bantoid languages either innovate adpositional phrases, serial verb constructions, or both. This therefore raises two questions. First, where did Bantoid languages get their prepositions (or, in a few cases, postpositions)? Recall that the proto language may have only had one preposition *na, which occurs widely in Niger-Congo.

> A feature common to languages that have obligatory applicatives and to languages that have the type of complex predicates presented in section 4.3.6 [serial verb constructions] is that, in comparison with other languages, they make only a very limited use of adpositions, since adpositions typically encode the semantic role of obliques, and both mechanisms result in giving the status of direct objects to various semantic types of complements that in other languages tend to be treated as obliques. (Creissels et al. 2008: 124)

The second question concerns how Bantoid languages developed their serial verb constructions (SVCs)? In order to investigate these questions, I decided to survey what has replaced the benefactive and recipient functions of the CB applicative extension -il- and the common -an- suffix which marks reciprocal in CB, but also instruments in Cameroonian NW Bantu:

(10)   a.   Mokpe [A22] -sos- 'wash' -sos-an- 'wash with' (Henson 2001)
        b.   Akoose [A15C] -kób- 'catch' -kób-ɛn- 'catch with' (Hedinger 2008: 90)

From the available literature, aid of colleagues over email, and my own work, the goal was to fill out the following questionnaire for as many as possible of the ca. 100 Bantoid languages in this small area of Cameroon.

1.   How are benefactives expressed? Which of the following are possible for the meaning 'he cooked rice for the child'?
     a. *double object:* "cook child rice"
     b. *benefactive preposition:* "cook rice for child" [if yes, what is the preposition?]
     c. *serial verb construction:* "cook rice give child"

Larry M. Hyman

2. How are recipients expressed? Which of the following are possible for the meanings 'he gave the child a book' or 'he sent/wrote the woman a letter'? [They are not necessarily the same]
   a. *double object:* "write woman letter", "give child book"
   b. recipient preposition: "write letter to woman", "give book to child" [if yes, what is the preposition?]
   c. *comitative preposition:* "write woman with letter", "give child with book" [if yes, what is the preposition?]
   d. *serial verb construction:* "write letter give woman", "take book give child"

3. How are instruments expressed? Which of the following are possible for the meaning 'he cut the meat with a knife'?
   a. *instrumental preposition:* "cut meat with knife" [if yes, what is the preposition?]
   b. *serial verb construction:* "take knife cut meat"

The table in the Appendix presents findings from 27 languages. Concerning the marking of ditransitives (benefactives, recipients, instruments), the following generalizations were noted:

(i) In all subareas there is at least some resistance to multiple objects, which are often restricted to only a few verbs.

(ii) There is no applicative or instrumental valence-marking by verb extensions, whereas there are identifiable, though not necesarily productive causative extensions in many Bantoid languages.

(iii) Virtually all of the flagging and word order strategies summarized by Malchukov, Haspelmath & Comrie (2010) are found in this small area, e.g. both adpositions and serial verb constructions (SVCs), which represent different responses to the change from synthetcity to analyticity.

As mentioned, Bantoid languages do retain verbs with recognizable causative suffixes. However, causative *-sə*, which corresponds directly to CB *-is-*, is usually restricted to intransitive roots due to the widepread resistance to double object constructions. In the few transitives that have been found with a causative extension, the verb does not become ditransitive:

(11) a. Babungo
       ŋwə́ fèe zɔ̀ 'he was afraid of (i.e. feared) a snake'
       mə̀ fè-sə̀ ŋwə́ (nə̀ zɔ̀) 'I frightened him (with a snake)'
       (Schaub 1985: 211)

b. Bafut
má shwìʔì ŋki 'I am pouring water'
má shwìʔì-sɜ̀ ŋkì 'I am making water to pour'
(Bila 1986: 102)

While causative extensions are attested, reflexes of the CB applicative suffix *-il-* are virtually absent in the Bantoid area. One possible exception concerns six out of Ngum's (2004) lexicon of 262 verbs in Meta [GB; Momo subgroup]:

(12) ghàb   'share'   ghàb-rɨ   'share to'
      cob   'donate'   cob-rɨ   'donate for'
      sòm   'cut'   sòm-bɨ   'cut into'
      wí   'refund'   wíí-rɨ   'reply, refund to'
      wub   'crave'   wub-rɨ   'crave for'
      dìì   'pity'   dìì-rɨ   'pity for'

However, since *-rɨ* has other functions, it is not clear if this suffix is cognate with PB applicative *-ɪl-*. The only other applicative I have found in the area comes from Vute (Mambiloid), which has innovated a new extension *-nà* from the main verb 'to give'. "*-nà* is added to a verb to indicate that there is an indirect object or benefactive NP present in the clause. Its function is similar to a Bantu applicative extension in this way. *-nà* is derived from the verb *nà-nɨ* 'to give'." (Thwing 2006: 8) Table 3 summarizes the different constructions that replace former applicative and instrumental verb extensions.

Table 3: Benefactive, Recipient, and Instrumental Constructions in Bantoid

alignment	schema	Benefactive	Recipient	Instrument
neutral	verb + X + Y	*cook child rice*	*write woman letter*	
indirective	verb + X + [prep Y]	*cook rice for child*	*write letter to woman*	
secundative	verb + Y + [prep X]	*cook child with rice*	*write woman with letter*	*cut meat with knife*
co-verb (Y)	verb + X + [give Y]	*cook rice give child**	*write letter give woman*	
co-verb (X)	[take X] + verb + Y			*take knife cut meat*

Although some languages do maintain unmarked double objects, assumed to be inherited, the more pervasive strategies are to replace head marking with adpositions and/or SVCs, with subareal distributions (see below). Let us first consider prepositions, then serial verbs. As mentioned, the proto language had perhaps only one preposition, **na* 'with' whose various reflexes *nə, nɨ, ni, nɛ* may expand to take on all three functions 'for', 'to', and 'with', as in Limbum [Eastern GB] (Fransen 1995: 259):

Larry M. Hyman

(13) a. *wìr bí   fàʔ   nì   Tàrī*
         we FUT₀ work for lord
         'we will serve [work for] the Lord'

     b. *mě fā   ŋwàʔ nì mūū wǎ*
         I   give-PERF book to child my
         'I have given a book to my child'

     c. *mě gwàr   cī   nì   ndyàà*
         I   cut-PERF tree with axe
         'I have cut the tree with an axe'

Table 4: please provide a caption

a.	cl. 3	wáy	w-ɛ́m	'my market'	('at market of me')
		fɔ̀-wǎy	fɔ̄ mē	'at my market'	
b.	cl.9	jɔ̀ɔ̀	y-ɔ̀	'your sg. stream'	('in stream of you sg.')
		ɛ̀-jɔ́ɔ̀	jɔ̄ wɔ̀	'in your sg. stream'	
c.	cl.9	còn	y-è	'his/her hut'	('in hut of him/her')
		cōǹ	dvū wvù	'in his/her hut'	

   In other cases the source of the preposition is from a locative. Bantu languages have locative noun classes that condition agreement. These are also present in certain Bantoid languages, although not always easy to identify with PB. Thus, Noni [Beboid] *fɔ* in Table 4a is cognate with PB **pa*, while the other two locative noun classes in Table 4b,c have no known PB correspondence (Hyman 1981). A comparison of the possessor marking in these examples reveals that independent pronouns are used instead of possessive pronouns with the above locative classes, indicating that they are prepositions. I suggest that the same locative source is involved in the development of the widespread preposition *á ~ án* which comes to be used as a benefactive and/or recipient preposition, e.g. in Noni, where the synchronic reflex of **a* is [ɛ] (Hyman 1981: 80):

(14) a. *mē nɔ́ɔ̀   ndɛ̀ɛ̀ wǎn bèŋkfǔ*
         I   PERF.FOC cook child yams
         'I have cooked the child yams'

     b. *mē nɔ́ɔ̀   ndɛ̀ɛ̀ bèŋkfǔ ɛ̄   wān*
         I   PERF.FOC cook yams   for child
         'I have cooked yams for the child'

Assuming an earlier NP prep NP structure explains the unusual verb + X + Y word order in Medumba [GB; Bamileke], which has lost the preposition *á, but still uses the independent pronominal forms as "indirect object pronouns" (Voorhoeve 1976: 22):

(15)   a. *á ᵗfá éᵗé bó*
       he gave it   them
       'he gave it to them' (cf. direct object pronoun *yób* 'them')

    b. *a⁵ fɔ³ bum² bu³*
       he give egg   dog
       'he gives an egg to the dog' (Caroompas 2014: 2)

Two other areal developments can be noted from the data in the Appendix (and compared with the accompanying map (Figure 1). First, in a contiguous area involving two subgroups of Grassfields Bantu (Eastern Grassfields and Momo), the benefactive/recipient preposition is reinforced by the noun 'hand' (cf. PB *-bókò*); hence, 'to the hand(s) of s.o.' becomes a new, fuller preposition.[4] Elizabeth Magba (pers.comm.) thus points out the following two possibilities in Mundani [GB; Momo]:
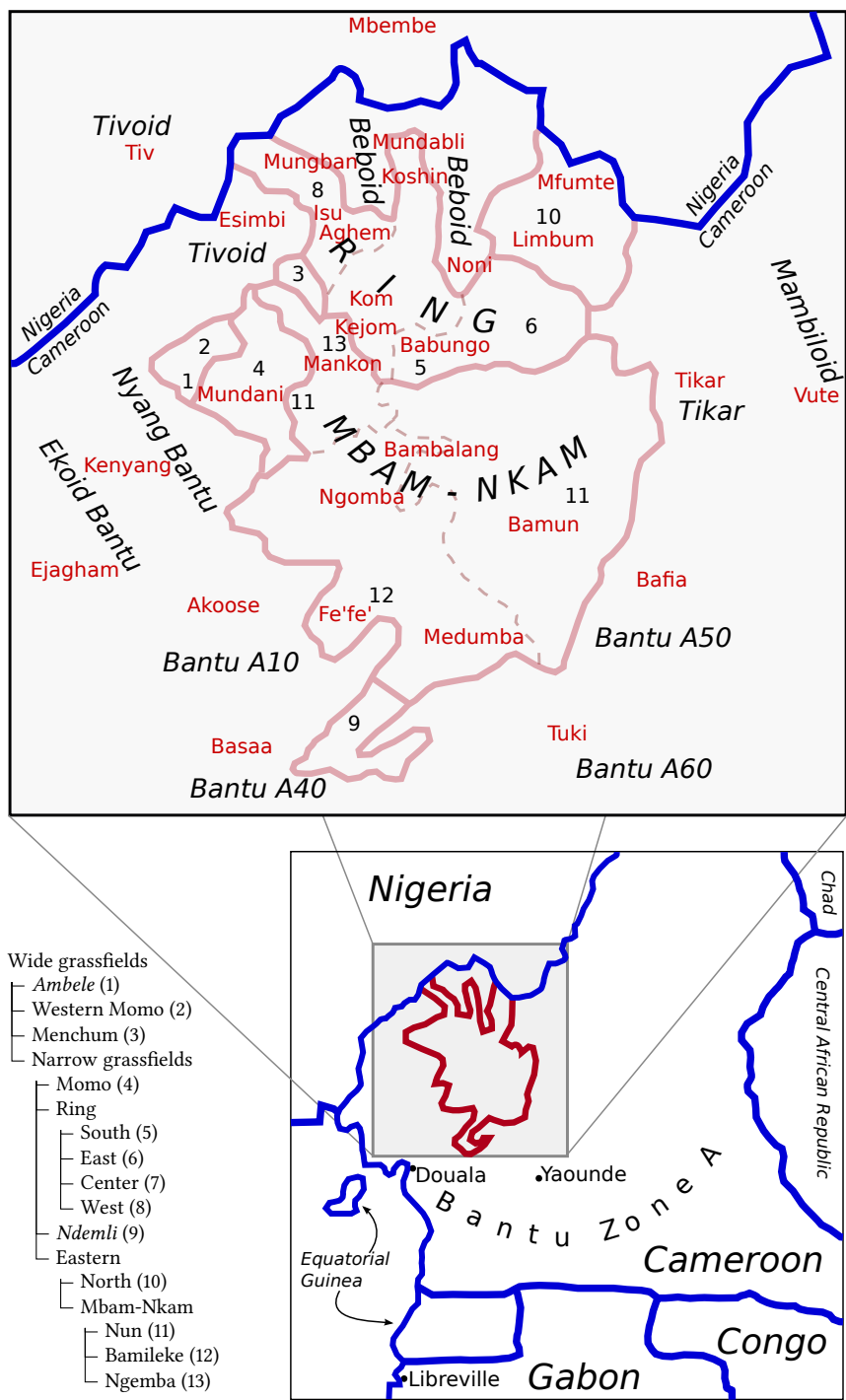
(16)   a. *tà tsaa   àkate yu abua tò*
       he has-sent letter the to   her
       's/he has sent the letter to her/him'

    b. *tà tsaa   a tò àkate yu*
       he has-sent to her letter the
       (idem)

> The difference between examples [16a and [16b] in terms of marking the recipient role has to do with a difference in focus: in [16a] *abua tò* (the recipient) is in focus, appearing in clause-final position; in [16b], *àkate yu* (the item sent...) is brought into focus by being shifted to the clause-final position.... The origin of *abua* variously translated as 'to, for, from, with' is likely to be the noun *àbu* 'hand, arm', possibly suffixed by the Class 7 genitive marker *-a*. (Elizabeth Magba, pers.comm.)

It is likely that Isu [GB; Ring] *áwɔ̀* 'for' (benefactive) derives from *á + kɔ̀-wɔ́* 'hand' (with common prefix-deletion and tonal change) and that *áwɔ̀* subsequently developed into *â* 'to' (recipient) (Roland Kießling, pers.comm.):

---

[4] Note that Heine & Kuteva (2002: 166) have 'hand' > LOCATIVE, but not RECIPIENT.

Larry M. Hyman

Figure 1: Map of Languages Surveyed (base map from Watters 2003: 226)

(17) a. *ɣú fàʔà   áwɔ̀ dɔ̀ŋ k-ìy*
          3pl work.IPF for   king 7-OF
          'they worked for the king'

      b. *ú   kɔ̀ʔ yɔ̀ wè dzài yɔ̀   â wè*
          3s.PST3 see. CFG 3SG tell   CFG to 3sg
          'he saw her and told her'

Locative *á* is implicated in the similar development of the benefactive and re-
cipient preposition â in closely related Aghem [GB; Ring group] (Watters 1979:
152-8), but also marks instruments by itself (Hyman 1979: 45):

(18) a. á fɪ́ghàm 'on the mat'
          á kɪ́⁺tú 'on the head'

      b. á fɪ̀⁺ñɨ 'with a knife'
          á kɪ́kɔŋ 'with a stirring stick'

The second areal development concerns a new instrumental preposition *ꞌbɔ́*
which replaces *ꞌna* 'with' in the North (Jukunoid, Yemne-Kimbi, Beboid, North-
ern subbranch of Eastern Grassfields). As seen in the Noni examples in (19) *bɔ́* is
used with persons, instruments and secundative 'give Y with X':

(19) a. *me ntɔ́ɔ́ bɔ́   wān*
          I   come with child
          'I am bringing the child'

      b. *me nɔ́ɔ̀ nsɛ̀ɛ̄ ñàm bɔ́   fèñɔ̄*
          I   PERF cut   meat with knife
          'I have cut meat with a knife'

      c. *me cí   ñá bɔɔm   bɔ́   kèŋgɔ̀m*
          I   PST₂ give children with plantains
          'I gave the children plantains'

Consistent with earlier speculations, the likelihood is that this preposition
comes from the third person plural pronoun of the same shape: incorporative
'they-with s.o.' > associative 'they-with sth.' ('they left they-with load of yams')
> instrumental 'with'. "... perhaps *bɔ́* 'with' comes from *bɔ́* 'they'." (Hyman 1981:
81, re Noni) "This conjunction [*bɔ́* 'and'] is identical in form to the third per-
son plural pronoun from which it is probably derived." (Hedinger 2008: 72, re

Larry M. Hyman

Akoose)[5] The likely starting point is incorporative pronouns, widespread in this area, e.g. Akoose (Hedinger 2008: 73):

(20)   a.   *bɔ́   awi mwaád*
             they his wife
             'he and his wife' (i.e. they including his wife)

       b.   *bɔ́   María*
             they Mary
             's/he and Mary' (i.e. they including Mary)

       c.   súmə̄
             's/he and I' (lit. we-(s)he')

A diachronic development of comitative > instrumental is a very common one cross-linguistically (Creissels & Voisin-Nougier 2008: 292). As seen in (21), both the new preposition < 'they' and inherited **na* form secundative *verb Y with X* in the North and Ring groups:

(21)   a.   *mə̄ fà   wɔ̀ bɔ́   ndì*
             I   give you with water
             'I give you some water'
             (Koshin [Yemne-Kimbi]; Ousmanou 2014: 309)

       b.   *mɔ̀ kɔ́   Làmbí nɔ̀   fá*
             I   give Lambi with thing
             'I give something to Lambi'
             (Babungo [GB; Ring]; Schaub 1985: 60)

It can be noted that no Bantoid language has secundative 'Y with X' without also having an alternative 'X to Y'. Leaving prepositions, another areal development is serial verb constructions (SVCs) which have also been innovated to express multiple arguments in Bantoid:

(22)   a.   Benefactive 'give' (Bamun [GB; Nun])
             nasha   na   malori mfa ne pon
             my.mother cook.PST rice   give to children
             'my mother cooked rice for the children' (Abdoulaye Nchare,
             pers.comm.)

---

[5] Note that 'and' and 'with' are often expressed with the same morpheme in Bantu languages.

5 *Multiple argument marking in Bantoid: from syntheticity to analyticity*

    b. Benefactive 'give' (Mundani)

       tà lè lạạ ėghɨdzɨ ŋa abua tò

       she PST₃ cook food give to   him

       'she cooked food for him' (Elizabeth Magba, pers.comm.)

    c. Instrumental 'take' (Ngomba [GB; Bamileke])

       n dɔ̌k   níi   ŋ́-kxuɣʏ̌ʔ tʉ́

       I take.PST machete CNS-cut tree

       'I cut the tree with a machete' (Satre 2010: 60)

From the table in the Appendix, we can make the following observations concerning the distribution of SVCs: (i) 'give' and 'take' SVCs are definitely in the minority (see the numbers in the bottom row of the table); (ii) except for Mbembe [Mambiloid] in the North and Ejagham [Ekoid Bantu] in the South, SVCs are found throughout the Grassfields area except the Ring group; (iii) although 'give' and 'take' SVCs are absent, Ring Grassfields Bantu exploits SVCs in other functions. This is extensively documented by Kießling (2011) for Isu and can also be seen in the following example from closely related Aghem (Hyman 1979: 204):

(23) *sǒogɔ̀ʔ vʉ́ ndùu nùŋò èkɔ́ʔ zɨ̀ghà mà?à tsùghò áwɛ́,   nùŋò èndú*

     soldier that go leave ascend leave throw descend children leave go

    ndùu kɔ̀ʔ ndùu nùŋɔ̀   vʉ̀

     go see go woman that

    'the soldier went and abandoned his children and went to see the woman'

The absence of valence-related serial verbs in the Ring subgroup is consistent with Foley & Olson's (1985) observation that SVCs are expected to be acquired in the specific order: motion/directional verbs > postural verbs > stative/process verbs > valence. "On the grammatical side, phonological attrition causes gradual loss of the bound morphemes…. At this verbal morphological is lost, a new device for valence adjustment must be found. Verb serialization begins to be used in this function, *provided serial constructions already exist in the language*." (Foley & Olson 1985: 51) [my emphasis] Concerning the order in which different valence SVCs are acquired, the present survey of Bantoid languages suggest two generalizations. First, 'give' SVCs are acquired before 'take' SVCs. Thus, Mfumte [EG; North] uses a 'give' SVC for benefactives, but a preposition *wɔ́* 'with, to' instead of an instrumental 'take' SVC (Greg McLean, pers.comm.):

Larry M. Hyman

(24) a. *yɔ́ tó fá mɔ̀ nku*
       3SG call give 1SG chief
       's/he called the chief for me'

    b. *yɔ́ sɨ ngyaʔ wɔ́ mbyì*
       3SG cut meat with knife
       's/he cut meat with a knife'

Second, benefactive 'give' SVCs are acquired before recipient 'give' SVCs. Evidence for this has already been seen from Mundani (16a) 'send to' vs. (22b) 'cook give', repeated below (Elizabeth Magba, pers.comm.):

(25) a. *tà tsaa àkate yu abua tò*
       he has-sent letter the to her
       's/he has sent the letter to her/him'

    b. *tà lè lạạ èghɨdzɨ ŋa abua tò*
       she PST₃ cook food give to him
       'she cooked food for him'

Fe'fe' [GB; Bamileke] also supports the idea that 'give' is initially oriented towards the benefactive rather than the recipient (Hyman 1971; pers.notes):[6]

(26) a. *à kà láh càk náh nsàʔ mbú à*
       3SG PST₂ take pot &take &come to me
       's/he brought the pot to me'

    b. *à kà láh càk náh nsàʔ hā ā*
       3SG PST₂ take pot &take &come give me
       's/he brought the pot for me'

    c. *à kà láh càk náh nsàʔ hā mbú à*
       3SG PST₂ take pot &take &come give to me
       's/he brought the pot to me'

    d. *à kà láh càk náh nsàʔ mbú à hā ā*
       3SG PST₂ take pot &take &come to me give me
       's/he brought the pot to me for me' (helped get pot to me)

---

[6] In these examples *náh* is a common simplification of *ndáh*, the consecutivized form of *làh* 'take'. The RECIP marker *mbú* is derived from the plural 'hands'.

The Fe'fe' data underscore that there are alternatives—and combinations, e.g. 'verb + give + to'. In addition, there is a preposition *ma* 'with' which has the same functions as *láh* 'take' (Hyman 1971: 33-37).

(27) a. *à   kà fáʔ mà žínù*
         3SG PST₂ work with intelligence

         'he worked intelligently' (he worked with intellligence)

     b. *à   kà láh žínù   mfáʔ*
         3SG PST₂ take intellligence &work

         'he worked intelligently' (he took intelligence &worked)

     c. *à   kà láh žínù   náh mfàʔ*
         3SG PST₂ take intelligence &take &work

         'he worked intelligently' (he took intelligence &took &worked)

This leaves us with the question: Why do Bantoid (and other) languages develop multiple strategies in the passage from syntheticity to analyticity? I take this up in the final section.

3 Conclusion

In response to why languages might develop alternative analytic structures, first consider the use of serialized 'take' as a "linker" in Fe'fe' in (28).

(28) a. *à´ mfáʔ   náh nghŭ nkāa*
         3SG work.PRES &take &make money

         's/he works and thereby earns money'

     b. *à´ ncēh   náh njíʔsī wū*
         3SG read.PRES &take &learn thing

         's/he reads and thereby learns'

As seen, I have translated '&take' as 'thereby', since it refers back to a proposition, not to a noun phrase. This is something that *mà* 'with' cannot do. Besides its ability to express a wider range of semantic roles than the preposition 'with', 'take' can also acquire an aspectual function, e.g. marking completive aspect in Gwari, a Nupoid language of Nigeria (Hyman & Magaji 1970):

(29) a. *wo si shnamá*   (present habitual)
         3SG buy yam

         's/he buys a yam'

    b. *wo si shnamá lo*   (present progressive)
       3SG buy yam   go
       's/he is buying a yam'

    c. *wó lá shnamá si*   (present perfect)
       3SG take yam buy
       's/he has bought a yam'

However, I don't think this is why SVCs develop. Rather, they originate as offering something different from the constructions with which they compete—and may ultimately replace. Much of the discussion concerned with defining SVCs has centered around VCs represent a single "event" (see Bohnemeyer et al. 2007, Bisang 2009 and references cited therein). However, speech communities differ in how much detail of an event they customarily express. Thus consider the function of 'take' as a "custody transfer" verb in Mungbam [Yemne-Kimbi] (Lovegren 2013):

(30)   a. *m̄*   *mū*
       take.IRR drink.IRR
       'take and drink!'
       [cup is within reach and at the level of the listener's hands, in front of him]

    b. *m̄*   *jɔ́*   *à*   *mū*
       take.IRR ascend.IRR 2S.TOP drink.IRR
       'take and drink!'
       [cup is on the floor and has to be "ascended"]

    c. *mū*
       drink.IRR
       'drink!'

As Lovegren puts it:

> "In an event description of this type, the absence of a custody transfer coverb usually indicates that no custody transfer took place (because the theme was already in the agent's custody at the outset of the event, because the action was performed without the agent taking custody, because the theme ceased to exist at the end of the event, etc.), and not that the custody transfer event is left unspecified. The only situation where a simple imperative *mū* 'drink!' is felicitous is a case where the addressee is already holding a drinking cup." (Lovegren 2013: 222)

5 *Multiple argument marking in Bantoid: from syntheticity to analyticity*

This raises the question of whether there could be comparable distinctions in expressing multiple arguments, e.g. benefactives and instruments in the following situations, all representing a single event:

(31) a. he cooked rice for child [the rice is still in the pot]
          he cooked rice give child [the rice is in the child's possession]
      b. he cut meat with knife [the knife was in his hand prior to the cutting]
          he took knife cut meat [the knife was not in his hand prior to the cutting]

A quite logical subsequent step would be for the SVCs in (31) to become the obligatory structure for expressing benefactives and instruments. Thus, in addition to Foley & Olson's (1985) demonstration that valence marking SVCs develop last, languages that have developed benefactive, recipient and instrumental SVCs may be at different stages: those like Fe'fe' which have alternate structures are "younger" serial verb languages than those like Mundani which lack prepositional alternatives.[7] It is however likely that Bantoid developed its SVCs fairly recently. As I pointed out in earlier work (Hyman 1975: 139-141), the type of SVCs surveyed above are an areal phenomenon in West Africa. However, the Bantoid distribution suggests there are micro-areas, since within the area surveyed, valence-marking SVCs are restricted to Eastern Grassfields Bantu and Momo languages. Such discontinuities probably hold in other parts of the continent as well. To conclude, I would like to draw the perhaps obvious moral that some languages care about certain things more than others. That some languages such as Mungbam care more about expressing the individual components of an action than English is not a new observation. Consider in this connection what Pawley (1993: 87) notes about Kalam, a language of New Guinea: Kalam speakers are markedly more analytic and explicit than speakers of European languages in their reporting of the action components of events. (Pawley 1993: 87) Kalam speakers thus say "food consume" for 'eat' and "water consume" for 'drink' (p.107) and have such elaborate SVC constructions as the following, which Pawley translates with one English verb (p.88):[8]

---

[7] This would of course suggest that more westerly Benue-Congo and Kwa languages which only have SVCs have had their serial verbs much longer.

[8] Thanks to Woodbury (2015) for bringing Pawley (1993) to my attention. An example closer to home might be the expression of motion events in "satellite-framed" Germanic languages which encode more about manner than "verb-framed" Romance languages (Talmy 1991, Slobin 2003).

Larry M. Hyman

(32) *pk   wyk d   ap   tan   d   ap   yap   g-*
      strike rub hold come ascend hold come descend do
      'to massage'

It is clear that different speech communities adopt different conventions for expressing similar events. While English has the compact verb "fetch", other languages require a tripartite SVC "go take come". Once a speech community starts to move in such an analytic direction the "drift" can on a life of its own. I would like to suggest a change in conversational conventions is not only responsible for the development of SVCs, but also for their areal diffusion: communities in contact borrow the speech styles of others, and thereby their grammar.

Abbreviations

ANTICAUS	anticausative	PB	Proto-Bantu
APPL	applicative	PERF	perfect(ive)
CAUS	causative	PL	plural
CB	Canonical Bantu	PRES	present
CFG	centrifugal	PST	past
CNS	consecutive	REC	reciprocal
FOC	focus	REFL	reflexive
FUT	future	REL	relative
GB	Grassfields Bantu	REV	reversive
INFL	inflection	SBJV	subjunctive
INST	instrumental	SG	singular
NP	noun phrase	SM	subject marker
OM	object marker	SVC	serial verb construction
PASS	passive		

5 Multiple argument marking in Bantoid: from syntheticity to analyticity

Table 5: Benefactive, dative & instrumental structures in Cameroonian Bantoid

(L. Hyman, 2015)

#	Language	Group	BENEFACTIVE			RECIPIENT				INSTRUMENT		Source
			V Y X	X for Y	V X give Y	V Y X	X to Y	Y with X	V X give Y	V X with Y	take Y V X	
1	Mbembe	Jukunoid	+	ké	+ (1)	+	ké	wō	+ (1)	wō	-	Richter (2015)
2	Mungbam	Yemne-Kimbi	- ?	á … nə̀	- ?	- ?	á … nə̀	bɛ̀, bɔ̀ (7)	- ?	bɛ̀, bɔ̀	-	Lovegren (2013)
3	Mundabli	Yemne-Kimbi	-	í	-	-	í … lā	ā	-	ā	-	Rebecca Voll
4	Koshin	Yemne-Kimbi		-lɔ̀	-	-	-lɔ̀	bɔ́		bɔ́	-	Ousmanou (2014)
5	Noni	Beboid	+	ɛ̀, -lé (8)	-	-	ɛ̀	bɔ́	-	bɔ́	-	Hyman (1981)
6	Mfumte	EG-North	+	-	+ (1)	+	wɔ́	-	+	wɔ́	- (2)	Greg McLean
7	Yamba	EG-North				+				bɔ́		Nassuna (2001)
8	Limbum	EG-North		nì			nì			nì		Fransen (1995)
9	Bamun	EG-Nun	- (5)	nə̀	+ nə̀ (3)		nə̀		+ (3)	nə̀	+ (4)	Abdoulaye Nchare
10	Medumba	EG-Bamileke		-	-	+ (5)	-	-	-	bûù	- (2)	Ariane Ngabeu
11	Feʼfeʼ	EG-Bamileke	+	mbú	+	+	mbú	-	+ mbú	mɔ̀	+	Hyman (1971)
12	Ngomba	EG-Bamileke		+ mbɔ̀ (3)	+ mbɔ̀ (3)		mbɔ̀			nɛ́	+	Satre (2004, 2010)
13	Bambalang	EG-Bamileke					nì			nì		Wright (2009)
14	Mankon	EG-Ngemba	- ?	nì	+	- ?	nì, á m̀bó	- ?	-	nì	- ?	Leroy (2007)
15	Mundani	Momo			+ abua (3)		abua, á		-	-	+	Elizabeth Magba
16	Babungo	Ring-South	-	tí	- ?	-	tí	nə̀	- ?	nə̀	- ?	Schaub (1985)
17	Kejom	Ring-Central	-	à	-	-	à	nə̀	-	nə̀	-	Pius Ajumbu
18	Kom	Ring-Central	-	sɔ̀	-	-	sɔ̀	nə̀	-	nə̀	-	Blasius Chiatoh
19	Aghem	Ring-West	-	â	-	-	â	-	-	á(n) (6)	-	Hyman (1979)
20	Isu	Ring-West	-	áwɔ̀	-	-	â	nə̀	-	nə̀	-	Roland Kießling
21	Esimbi	Tivoid?		ɔ̀hɔ̀ kV	-	+	ɔ̀hɔ̀ kV	-	-	ótɔ̀	-	Brad Koenig
22	Tiv	Tivoid	+	shá	+	+			-	shá, á		Abraham (1940)
23	Kenyang	Mamfe		ǹtá, tɔ̀kɔ́	- ?	+	ǹtá			nɛ̀		Tanyi Eyongetah
24	Ejagham	Ekoid	+	m̀bâ	+	+	m̀bâ	-	+	nə̀	+	John Watters
25	Akoose	Bantu A10	+ (9)	áyə̀lɛ̀	-	+	wê	-	-	ne (9)	-	Hedinger (2008)
26	Tikar	Tikar	+	lɛ̀		+	lɛ̀			lɛ̀		Stanley (1991)
27	Vute	Mambiloid				+ (10)						Thwing (2006)
	totals:		7	16	6	11	22	8	5	24	5	

Larry M. Hyman

Notes on Table 5

In Table 5, + means the language has the construction (which can be general or limited to certain verbs); - means it doesn't have it; blank = no info; EG = Eastern Grassfields (1) The Mfumte and Mbembe structure is *V give Y X* + resumptive 'with';
(2) The Mfumte and Medumba structure is *take X cut Y with(it)*, two events.
(3) The Bamun, Ngomba and Mundani structure is *V X give to Y*;
(4) The Bamun structure is *take X cut Y with(it)* = one event.
(5) The Medumba order is V X Y (the Y is from a PP, X, Y pronouns are distinct).
(6) Aghem *á(n)* is the general locative preposition, used also with instruments (but not comitatives, which use *à*);
(7) Mungbam Y with X also used for BEN.
(8) In Noni, ɛ̄ means 'to s.o.' or 'for s.o.'s benefit', while the locative suffix *-lé* means 'for s.o.' (in s.o.'s stead).
(9) Akoose has productive verb extensions: applicative *-e* producing V-e Y X and an *-ɛn* instrumental verb extension producing V-ɛn X Y (Y = the instrument NP).
(10) Vute has an applicative extension *-ná* from the verb 'to give'.

References

Abraham, R. C. 1940. *The principles of Tiv*. London: Crown Agents for the Colonies.

Akumbu, Pius Wuchu. 2008. *Kejom (Babanki) - English lexicon. Kay Williamson Educational Foundation*. Bamenda: Ga'a Kejom Development Committee.

Bila, Emmanuel Neba. 1986. *A semantic-syntactic study of the Bafut verb*. Doctoral dissertation. Yaounde: University of Yaounde.

Bisang, Walter. 2009. Serial verb constructions. *Language and Linguistics Compass* 3. 792–814.

Blanchon, Jean. 1995. *Punu lexical database in Filemaker Pro*.

Bohnemeyer, Jürgen, Nicholas J. Enfield, James Essegbey, Iraide Ibarretxe-Antuñano, Sotaro Kita, Friederike Lüpke & Felix K. Ameka. 2007. Principles of event segmentation in language: The case of motion events. *Language* 83. 495–532.

Caroompas, Jordan. 2014. *The dative alternation in Medumba*. Ms.

Crane, Thera Marie, Larry M. Hyman & Simon Nsielanga Tukumu. 2011. *A grammar of Nzadi [B865]* (University of California Publications in Linguistics 147).

Berkeley & Los Angeles: University of California Press. http://escholarship.org/uc/item/846308w2.

Creissels, Denis & Sylvie Voisin-Nougier. 2008. Valence-changing operations in Wolof and the notion of co-participation. In Ekkehard König & Volker Gast (eds.), *Reciprocals and reflexives: Theoretical and typological explorations*, 289–306. Berlin: Mouton de Gruyter.

Creissels, Denis, Gerrit J. Dimmendaal, Zymunt Frajzyngier & Christa König. 2008. Africa as a morphosyntactic area. In Bernd Heine & Derek Nurse (eds.), *A linguistic geography of Africa*, 86–150. Cambridge: Cambridge University Press.

Dimmendaal, Gerrit J. 2000. Morphology. In Bernd Heine & Derek Nurse (eds.), *African languages: An introduction*, 161–192. Cambridge: Cambridge University Press.

Ellington, John. 1977. *Aspects of the Tiene language*. Doctoral dissertation. Madison: University of Wisconsin.

Foley, William & Mike Olson. 1985. Clausehood and verb serialization. In Johanna Nichols & Anthony C. Woodbury (eds.), *Grammar inside and outside the clause*, 17–60. Cambridge: Cambridge University Press.

Fontaney, Louise. 1980. Le verbe. In F. Nsuka-Nkutsi (ed.), *Eléments de description du punu*, 51–114. Lyon: CRLS, Université Lyon2.

Fransen, Margo Astrid Eleonora. 1995. *A grammar of Limbum, a Grassfields Bantu language spoken in the North-West Province of Cameroon*. Doctoral dissertation. Amsterdam: University of Amsterdam.

Greenberg, Joseph H. 1963. *The languages of Africa*. The Hague: Mouton.

Güldemann, Tom. 2011. Proto-Bantu and Proto-Niger-Congo: Macro-areal typology and linguistic reconstruction. In O. Hieda, C. König & H. Nakagawa (eds.), *Geographical typology and linguistic areas*, 109–140. Amsterdam: Benjamins.

Hedinger, Robert. 2008. *A grammar of Akoose*. University of Texas, Arlington: SIL International.

Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

Heine, Bernd & Mechthild Reh. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Buske.

Henson, Bonnie. 2001. *The -an- extension and diachronic language change in Mòkpè, note =Paper presented at the Workshop on Comparative Benue-Congo, Berkeley, CA, March 26-27.*

Hyman, Larry M. 1971. Consecutivization in Fe'fe'. *Journal of African languages* 10. 29–43.

Larry M. Hyman

Hyman, Larry M. 1975. On the change from SOV to SVO: Evidence from Niger-Congo. In C. N. Li (ed.), *Word order and word order change*, 113–147. Austin: University of Texas Press.

Hyman, Larry M. 1979. *Aghem grammatical structure* (Southern California Occasional Papers in Linguistics 7). Los Angeles: University of Southern California. http://gsil.sc-ling.org/pubs/SCOPILS_6_7_8_9/Aghem_grammatical_structure.pdf.

Hyman, Larry M. 1981. *Noni grammatical structure* (Southern California Occasional Papers in Linguistics 9). Los Angeles: Department of Linguistics, University of Southern California.

Hyman, Larry M. 1998. Positional prominence and the 'prosodic trough' in Yaka. *Phonology* 15. 41–75.

Hyman, Larry M. 2003. Basaa A.43. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 257–282. London: Routledge.

Hyman, Larry M. 2004. How to become a Kwa verb. *Journal of West African Languages* 30. 69–88.

Hyman, Larry M. 2010. Affixation by place of articulation: The case of Tiene. In Michael Cysouw & Jan Wohlgemuth (eds.), *Rara & rarissima. Collecting and interpreting unusual characteristics of human languages*, 145–184. Berlin: Mouton de Gruyter.

Hyman, Larry M. & Daniel J. Magaji. 1970. *Essentials of Gwari grammar.* University of Ibadan: Institute of African Studies.

Jisa, Harriet. 1977. *Filecards of Babanki verbs (those with extensions digitized by LMH in Filemaker Pro™).*

Kießling, Roland. 2011. *Verbal serialisation in Isu (West-Ring) — a Grassfields Language of Cameroon.* Köln: Rüdiger Köppe Verlag.

Lehmann, Christian. 1985. Grammaticalization: Synchronic variation and diachronic change. *Lingua e Stile* 20. 303–318.

Lemb, Pierre & François de Gastines. 1973. *Dictionnaire basaa-français.* Douala: Collège Libermann.

Leroy, Jacqueline. 1982. Extensions en mankon, langue bantoue des Grassfields. In G. Guarisma, G. Nissim & J. Voorhoeve (eds.), *Le verbe bantou* (Oralité-documents 4), 111–141. Paris: Selaf.

Leroy, Jacqueline. 2007. *Le mankon.* Paris: Peeters.

Lovegren, Jesse Stuart James. 2013. *Mungbam grammar.* Doctoral Dissertation. University of Buffalo.

Malchukov, Andrej, Martin Haspelmath & Bernard Comrie. 2010. Ditransitive constructions: A typological overview. In Andrej Malchukov, Martin Haspel-

math & Bernard Comrie (eds.), *Studies in ditransitive constructions*, 1–64. Berlin: DeGruyter Mouton.

McGill, Stuart John. 2009. *Gender and person agreement in Cicipu discourse.* Doctoral dissertation. London: SOAS, University of London.

McWhorter, John H. 2011. Affixless in Austronesian: Why Flores is a puzzle and what to do about it. In John H. McWhorter (ed.), *Linguistic simplicity and complexity*, 223–60. Berlin: Mouton De Gruyter.

Nassuna, Nzenge Lucia. 2001. *Yamba: A morphosyntactic study of the basic sentence.* University of Yaounde I Dissertation for the Maîtrise.

Nguendjio, Emile-Gille. 1989. *Morphologie nominale et verbale de la langue baŋgwa.* Thèse de Doctorat de Troisième Cycle en Linguistique. Yaounde: Université de Yaounde.

Ngum, Comfort Che. 2004. *Verbal extensions in meta'.* University of Yaounde I Maîtrise in Linguistics.

Nurse, Derek & Gérard Philippson. 2003. *The Bantu languages.* London: Routledge.

Ousmanou. 2014. *How to disclose the environment through linguistic description: A basic linguistic analysis of Koshin ["Beboid" Bantu, Cameroon].* Doctoral dissertation. Yaounde: University of Yaounde I.

Paulian, Christiane. 1975. *Le kukuya: Langue teke du Congo.* Paris: SELAF.

Pawley, Andrew. 1993. A language which defies description by ordinary means. In William A. Foley (ed.), *The role of theory in language description*, 87–129. Berlin: Walter de Gruyter.

Richter, Doris. 2015. *A grammar of Mbembe.* Leiden/Boston: Brill.

Rose, Sharon. 2013. The morphological structure of the Moro verb. In Thilo C. Schadeberg & Roger Blench (eds.), *Nuba mountain language studies*, 83–104. Köln: Rüdiger Köppe Verlag.

Sapir, Edward. 1933[1949]. Language. Encyclopedia of the Social Sciences (New York). Reprinted 1949 in Mandebaum. In David G. (ed.), *Edward Sapir: Selected writings in language, culture, and personality*, 7–32. Berkeley & Los Angeles: University of California Press.

Satre, Scott A. 2010. The consecutive morpheme in Bamileke-Ngomba. *Studies in African Linguistics* 39(1). 35–76.

Schaub, Willi. 1985. *Babungo. Croom Helm descriptive grammars.* Beckenham, England: Croom Helm.

Scott, David & Alexander Hetherwick. 1970. *Dictionary of the Nyanja language.* London: National Educational Company of Zambia.

Larry M. Hyman

Slobin, Dan I. 2003. Language and thought online: Cognitive consequences of linguistic relativity. In Dedre Gentner & Susan Goldin-Meadow (eds.), *Language in mind: Advances in the study of language and thought*, 157–192. Cambridge, MA: MIT Press.

Stanley, Carol. 1991. *Description morpho-syntaxique de la langue tikar*. Yaounde: Société Internationale de la Linguistique.

Talmy, Leonard. 1991. Path to realization: A typology of event conflation. In *Papers of the seventeenth annual meeting of the Berkeley linguistics society*, 480–520. Berkeley: Berkeley Linguistics Society.

Thwing, Rhonda. 2006. *Verb extensions in Vute*. Ms.

Van de Velde, Mark. 2008. *A grammar of Eton*. Berlin: Mouton de Gruyter.

Voorhoeve, Jan. 1976. *Contes bamileke*. Tervuren: Musée Royal de l'Afrique Centrale.

Watters, John Robert. 1979. Focus in aghem: A study of its formal correlates and typology. *In Hyman* 1979. 137–197.

Watters, John Robert. 2003. Grassfields Bantu. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 225–256. London: Routledge.

Whitehead, John. 1899. *Grammar and dictionary of the Bobangi language*. London: Kegan P. & Trench.

Woodbury, Anthony. 2015. *The "genius" of the language: Discovering pervasive plan and unique design in linguistic description, note = Hale Lecture, July 21, 2015, Linguistic Summer Institute, University of Chicago*. Chicago.

Wright, Jennifer. 2009. *The noun and verb phrase in Chrambo (Bambalang)*. Ms. SIL Cameroon.

Chapter 6

Grammaticalization of participles and gerunds in Indo-Aryan: Preterite, future, infinitive

Annie Montaut

Inalco

The development of the Sanskrit passive past participle and gerund or passive verbal adjective of obligation in Indo-Aryan are up to a certain point parallel and resulted respectively in an ergative alignment in past sentences in Western languages and a nominative realignment in both future and past in Eastern languages. Only Eastern languages grammaticalized the old endings into the specific temporal markers -*l*- for past and -*b*- for future, while throughout the IA area the obligative passive verbal adjective also evolved into an infinitive. The aim of the paper is to account for the various grammaticalization paths of these forms in a unified manner, taking into account the whole range of other competing constructions in the various IA languages considered, as well as comparable instances of grammaticalization in Latin and Romance languages. Grammaticalization is understood here in the meaning of less grammatical to more grammatical as in Hopper & Traugott (2003) with a special attention to the shift in syntactic construction as in Benveniste (1966[1952]) and Kuryłowicz (1965).

1 Introduction

The ergative realignment of the Sanskrit passive past participle construction has been extensively studied in many Western Indo-Aryan languages, and is often taken to be a unique development in the area and the wider family of Indo-European. However, its nominative realignment in Eastern languages has raised far less interest. Moreover, the parallel development of the passive future verbal adjective of obligation into a future marker, and the infinitive, is still unstudied.

Annie Montaut

The aim of the paper, in continuation of Montaut (2016), is to account for the various grammaticalization paths of these forms in a unified manner. The words which developed in specific TAM markers with specific constructions are not lexical items but verbal adjectives, and grammaticalization is understood here in the meaning of less grammatical to more grammatical as in Hopper & Traugott (2003) with a special attention to the shift in syntactic construction as in Benveniste (1966[1952]) and Kuryłowicz (1965). To understand why different paths can lead to different outcomes in different languages, the whole range of other competing constructions is considered. The paper also includes a comparison with similar instances of grammaticalization in Latin and Romance languages. Indo-Aryan (IA) languages (apart from Dardic languages) are now conventionally classified into two main groups (Cardona & Suthar 2003), genetically associated respectively to the Western Sauraseni Prakrits in Middle Indo-Aryan (from which derive Gujarati, Hindi/Urdu, Punjabi, Rajasthani, Braj, among others), and to the Eastern Magadhean Prakrits (from which derive Bengali, Oriya, Assamese, Maithili, among others). I first analyse the different developments of the past verbal adjective and past passive participle, since the ergative development is limited to the first group, at the same time questioning theories of both passive and possessive origin of the perfect (§2). In §3, I present the parallel grammaticalization of the passive obligation participle (or gerund), and its different outcomes in Eastern and in Western Indo-Aryan, since the anomalous situation of Marathi invites to question the nature of this gerund. This question, along with the discussion of the gerund's grammaticalization as a verbal noun, is the subject of §4 This paper also attempts to clarify areal relations and the conditions for a given path of grammaticalization/reanalysis to actualize here and not there, while at the same time inquiring into the reasons for re-alignments.

2 The grammaticalization of the passive participle into finite past and "ergative alignments" in Indo-Aryan

As is well-known since Kellogg (1972[1875]) and Grierson (1903-1928), what is today called the ergative construction or alignment in Indo-Aryan, and what was in the 19th and early 20th centuries described as a "kind of passive construction", stems from a particular type of the Sanskrit nominal sentence, generalized in classical language[1]. The past passive participle (henceforth glossed PPP) or verbal adjective ending in *-(i)ta* was used as a predicate, replacing the finite Vedic synthetic past tense forms. In this predicative use, the PPP agreed in gender and number with the patient, the agent appearing in the instrumental. As the pred-

icative uses of PPP increased in frequency, it grammaticalized into the standard expression of past (cf. Bybee 2003 for the role of use frequency). The original instrumental case marking of the agent was later replaced by a postpositional marker, the so-called ergative case marker in Western IA languages, while the *-ita* (>*ia*) form acquired modern gender/number endings. The following example from classical Sanskrit (1) is a historical antecedent of the modern ergative construction in Hindi (2a), in contrast to the present or future nominative sentences (2b), a construction extensively studied since the 1980s for its syntactic as well as pragmatic properties (for Hindi cf. e.g. Kachru 1987; Montaut 2004; Davison 2002). In the ergative construction, the agent has most of the subject properties (particularly in control constructions) and the patient has only few discourse-related "subject properties".

(1) Sanskrit
 mayā / mama tat   kr̥tam
 1SG.INS / 1SG.GEN DEM NOM.N.SG do.PPP.NOM.N.SG
 'I did/have done that. (lit. 'by me/ of me this done')'[2]

(2) a. Hindi
  mãĩ.ne yah / apnā kām kiyā
  1SG.ERG DEM.M.SG REFL work.M.SG do.M.SG
  'I did this /my work.'

 b. Hindi
  mãĩ. yah / apnā kām kar.ũ̃.g.ī
  1SG DEM.M.SG REFL work.M.SG do.1SG.FUT.F
  'I will do this /my work.'

2.1 Early New Indo-Aryan (NIA) data: alignment shift and acquisition of temporal meaning

Equally well-known is the fact that this TAM-based pattern of split ergativity (found in definite past and derived tenses such as present perfect, pluperfect and all compound forms involving the past participle) is now restricted to the Western part of IA. What has been commented on less is the fact that it was prevailing throughout Indo-Aryan up to the middle stages of NIA (14th-16th century,

---

[1] In Late Classical Sanskrit, for instance in the tales of the Vampire (*Vetāla*), participial forms represent about 95% of past sentences (Bloch 1906: 60).

[2] The genitive marking of the agent is restricted to pronouns, instrumental being by far the most usual marker elsewhere.

Annie Montaut

depending on the regional variety). Here are early examples of the extension of the construction, both from Western and Eastern IA, right from the last stage of Old Indo Aryan or Prakrit (PRK); in (3) from Ashoka's first edict which displays both Western (3a: Girnar) and Eastern (3b: Jaugada) dialectal varieties, and (4), by the playwright Kalidasa in Saurasenic Prakrit, which shows the contrast between instrumental agent and nominal predicate in the past, and nominative agent and finite predicate in the present:

(3)   a.   *iyam dhammalipī devānampriyena priyadassina ranna lekhapita*

   b. | *iyam* | *dhammalipi* | *devānampiyena* | *piyadassina* | *[lajina]* |
        |---|---|---|---|---|
        | this | law-scripture | of-gods-friend | friendly-looking | king |
        | NOM.F.SG | NOM.F.SG | INST.M.SG | INST.M.SG | INST.M.SG |

   lekhita
   inscribed
   NOM.F.SG

   'The friendly looking king beloved of gods has (made) engraved this law-edict.' (PRK)

(4)   a. | *hau* | *pai* | *pucchimi* | ... | *diṭṭhī* | *pia* | *pai* | *sāmuha* |
        |---|---|---|---|---|---|---|---|
        | 1SG.NOM | 2.OBL | ask.PRS.1SG | | seen.F.SG | loved.F.SG | 2.OBL | in.front |

   jāntī
   passing.NOM.F.SG

   'I ask you... Did you see (my) beloved passing in front (of you)?' [3]
   (PRK)

All Western and Central IA languages, now ergative, displayed at an older stage contrasts similar to (4), with personal endings on the finite verb in the present whereas in the past the verb retains nominal morphology. In the latter pattern, the predicate shows gender and number agreement with the patient, while the agent is marked by the oblique case (a polyfunctional case, as a result of the usual syncretism ABL/DAT/LOC in the area). The pattern is attested throughout the Western and Central languages (5a-d); example (5e) from Kabir (13th c.) illustrates the so-called "saint language", a transregional literary koine which belongs to the vast category of old Hindi (Kellogg 1972[1875]):

(5)   a.   Old Punjabi

---

[3] The form *pai* for 2nd person is already used as a syncretic marker for several oblique cases.

    guri   *dānu*   *ditta*
    guru.OBL/LOC gift.M.SG given.M.SG

    'The guru gave the gift.' (*Guru Granth Sahib*)

b. Old Rajasthani
    sundari.nai   *Bharath.ai*   *rākhī*
    beautiful.lady.F.SG.ACC Bharath.OBL.M.SG kept.F.SG

    'Bharath kept the beautiful lady.' (Tessitori 1914–1916: 167)

c. *kṛpā*   *kel.i*   *tumhĩ (Old Marathi)*
    pity.F.SG do.l.F.SG 2.OBL

    'You have had pity.' (Jñaneśvari 11.255, in Bloch 1970[1920]: 261)

d. Old Braj
    maĩ   *nahĩ*   *mākhan khāyau*
    1.SG.INS NEG butter.M.SG eat.M.SG

    'I did not eat the butter.' (Surdas 25.1 mid 16th c.)

e. Sant Bhasha
    gur.i   *diyā*   *palītā*
    guru.LOC/OBL give.M.SG stick.M.SG

    'The guru gave the stick.' (Kabir 8.3)

f. Old Pahari
    Virrsigh Joysī.yã bhāṣ   *pāi*
    Virsingh Joshi.OBL proclamation/bond.F.SG get.F.SG

    'Virsingh Joshi (the king) received the bond.' (Stroński 2014: 283)

Some variation appears in this pattern already; note agreement with a marked object in Old Rajasthani (5b), a marking itself rather recent, as well as a *-l-* suffixation on the predicative participle in Old Marathi (5c), with variations in the oblique case marking (*ai, -hi /-i, -yã* ending). Still, the basic pattern is the same. This pattern was also maintained unchanged in the Eastern languages, with a similar contrast between nominative agent and finite form agreeing with agent in the present, as opposed to non-finite participial forms in the past with oblique agent. In the 16th century, the Bhojpuri first person pronoun still had a nominative form inherited from the Sanskrit nominative *aham* (*hau manus* 'I [am] a man'), whereas it displayed an oblique form stemming from the Sanskrit instrumental in the past (*maĩ pāi* 'I obtained') (cf. Tiwari 1966: 158). The examples in (6) illustrate a few of these Eastern varieties in early NIA, starting with a sample from the oldest Buddhist poems or *caryas*, in Old Bengali (Chatterji 1926), the

predecessor of Middle Maithili, Oriya, Bengali and Assamese, then other Eastern languages in their old or middle stages:

(6) a. Old Bengali
        mo.e   ghalil.i   hāḍerī-māli
        1SG.INS cast.off.*il*.F bone-garland.F
        'I have cast off the chaplets of bones.' (carya 10, in Chatterji 1926: 964)

     b. Old Bengali
        ebe maï   bujh.ila
        now 1SG.INS understand.*ila*
        'Now I have understood.' (carya 35, in Chatterji 1926: 964)

     c. Old Maithili
        bhala   na ka.la mañe   de.la   bisavāsa
        good.M.SG NEG do.*la* 1.SG.OBL give.*la* trust.M.SG
        'I did not [the] good, [that] I gave trust.'

     d. *Tirahuti   le.li   jānhi*
        Tirahuti.F.SG take.*li*.F.SG REL.OBL
        'By whom (the city) Tirahuti was taken = who took Tirahuti.' (from Jha 1958: 491)

     e. Old Awadhi
        eka rāta sapnā   mai dekhā
        one night dream.M.SG 1SG.INS see.M.SG
        'One night I saw a dream.' (Nur Mohammed 4)

Slight differences start emerging also here, such as the variety of oblique forms for the first-person pronoun (alternate forms of the instrumental in (6a-b), syncretic oblique in (6c-d), and the extension of a *−l/il* suffix to the predicative participle. This suffix, with no particular meaning, was used with nominal and adjectival basis and is now used for deriving adjectives in many IA languages.[4] All Eastern varieties, stemming from the Magadhean Prakrits (as opposed to those stemming from the Western Sauraseni Prakrits) display a progressive erosion of gender marking between the 14th and 16th century, so that agreement grows less distinctive. But again, the basic pattern is essentially maintained. As indicated by the translations of the examples in (5) and (6), by and large taken from the pub-

---

[4] For instance, *rangīlā* 'colourful' from *rang* 'colour' or *kānthīlā* 'thorny' from *kānthā* 'thorn' in Hindi.

lications providing the respective examples,[5] the temporal meaning of the form extends from present perfect and pluperfect to the mere representation of anterior events as a preterit. As soon as the old passive participle started developing as the only expression for past, it assumed both the original stative / resultative meaning (perfect) and a new anterior meaning. This change in meaning has been well documented in Peterson (1998: 190) for Pali and in Breunis (1990) for classical Sanskrit. I have accounted for it (Montaut 1996, 2007) along the lines of Bybee, Perkins & Pagliuca (1994) as a gradual process of grammaticalization of the new periphrastic form: as long as the nominal form, initially a marked innovation aiming at stylistic expressivity (Breunis), competed with the old tensed forms, it retained its original restricted meaning (resultative-stative). When the old forms disappeared, the new form, no longer stylistically expressive, occupied the whole space of past reference and acquired what Bybee, Perkins & Pagliuca (1994) calls an open meaning. This meaning conveyed the values of preterit or anterior, resultative, stative-resultative and stative, as already observed by Bloch (1906: 60).[6] Later on, through foregrounding, then conventionalizing, the implicature (action leading to the resulting state), the adjectival form acquired an anterior meaning, while a new periphrastic form with a copula emerged for the perfect (around the 17th-18th ct.). This foregrounding of dynamic aspect resulted in the agent becoming foregrounded. Apparently, the topicalizing fronting of the agent becomes the rule during Middle Indo-Aryan (MIA) (Breunis 1990: Chapter 6 on word order), although Bubenik & Paranjape (1996: 116-117) date the linguistic perception of the oblique noun as a semantic subject to late MIA. It should however be emphasized that the agent was endowed with control and binding properties from the very beginning of the use of the predicative past participle (cf. Hock 1992), which means that in Late Sanskrit too dynamicity and prominence of the agent where already present. Such a parallel acquisition of first dynamic perfect, then aorist and preterit meaning, and of syntactic subjecthood of the agent echoes the evolution of the North Russian and Circum- perfect (cf. Seržant 2012).

---

[5] Except when they were translated by passive sentences as (6a). As for the gloss, I adopted the gloss INS wherever the pronoun has a distinctively instrumental form, not a syncretic oblique form.

[6] "Du contexte et du sens de la racine dépendent la valeur active ou passive et la nuance temporelle et modale du participe. Il est donc le substitut de toutes les formes verbales du passé à tous les modes et à toutes les voix". As for the copula, it was originally used only for disambiguating an omitted agent in the first or second person.

Annie Montaut

2.2 Further grammaticalization in Modern Eastern New Indo-Aryan: a shift from passive to active or "possessive perfect"?

Whereas Western IA languages developed the non-nominative alignment with oblique agents further into a fully-fledged ergative pattern such as in (2a), Modern Eastern IA languages shifted back to a nominative alignment between the 14th and 16th century. Furthermore, most of them added personal endings to the participle, making it a finite tensed form, as in Bengali:

(7) a. Bengali
        āmi boi.ṭa   por.l.ām
        1SG book.DEF read.PST.1SG
        'I read the book.'

    b. Bengali
        tumi boi.ṭa   por.l.e
        2   book.DEF read.PST.2
        'You read the book.'

As noted by Chatterji (1926), the new past tense personal endings (1 -*ām*, 2-*i*, -*e*), derived from personal pronouns, are distinct from the inherited present tense personal endings (1 -*i*, 2 -*ish*, -*o*). As for the -*l*-, which is now analysed as a past tense marker, and which also occurs in Marathi, it originates from an adjectival suffix (cf. Hindi -*il*- in *rang.il.ā* 'coloured'). This suffixation is a further evidence for the adjectival nature of the predicate in the pre-ergative alignments (cf. Chatterji 1926: 928 and Tessitori 1914–1916). Its reanalysis as a past tense marker corresponds to the renewal of the "pre-ergative" alignment into a nominative alignment. Chatterji (1926), following the then-usual interpretation, considers this evolution a shift from passive to active,[7] but he clearly recognizes the resultative reading, as is also clear from his literal translation of this pattern, when the object is marked: 'there has been V by X as concerns Y'. Similarly, the evolution of this pattern in Eastern Hindi, which also displays the same suffix -*l*- in the definite past, is systematically interpreted as a shift from passive to active in Saxena (1937: 247 ff) for Bhojpuri, in Jha (1985[1958]: 492 ff) for Maithili, and in Tiwari (1966: 171) for Awadhi. The active reading, with concomitant acquisition of subject behavioural properties by the agent, was already present centuries before, so the modern shift in coding properties can be considered a mere final step in the grammaticalization of the construction. This evolution is in line with

---

[7] An "archaic" remnant of the old system survived in classical Bengali with the -*e* ending for agents in transitive past, and is still present in Assamese.

Haspelmath & Sims's (2010) principle of Behavioural before Coding Properties. This is also the case in other Indo-European languages such as Persian (cf. Cardona 1970) and Latin/Romance languages, which went through a similar cycle: finite past tense > nominal sentence with predicative participle agreeing with the patient and oblique agent > verbal sentence with nominative agent and verb agreeing with agent, as stated in Kuryłowicz (1931[1960]), 1965. Here is the Old (2.2a) and modern (2.2b) Persian data considered by the author:

(8) a. *mana   tyâ   karta.m*   > b. *man   in kar.d.am*
        1SG.GEN DEM.NOM.N.SG do.PPP.NOM.N.SG   1SG.NOM this do.PST.1SG
        'I have done that.'   'I did that.'

The corresponding Latin data involves a periphrastic past with participle and dative of the agent, renewed by means of the 'have' auxiliary and a nominative subject. In present-day French, a vestige of the old "pre-ergative" pattern is still available: if the direct object is preposed, the participle agrees with it, and not with the subject (*les lettres que j'ai écrit.es, je les ai écrit.es).* Examples (9a) and (9b) illustrate this shift to nominative alignment in Latin as analyzed in Kuryłowicz (1931[1960]), example (10) shows the output in modern Romance, with the French, Italian and Spanish translations of (9):

(9) a. Latin
        mihi   id   factum   (est)
        1SG.DAT DEM.NOM.N.SG do.PPP.NOM.N.SG (be.PRS.3S)
        'I did/have done that.' (*lit.* 'to.me this done (is)')

     b. Latin
        ego   id   factum   habeo
        1SG.NOM DEM.ACC.N.SG do.PPP.ACC.N.SG have.PRS.1SG
        'I have done that/it.' (*lit.* I have this done)

(10) French
      j' ai fait ceci
      Italian
      io ho fatto questo
      Spanish
      yo he hecho esto
      'I have done that'

As stated by Kuryłowicz (1931[1960]) in his paper on the formation of tenses in Romance languages, further developed in his study on the evolution of gram-

matical categories (Kuryłowicz 1965), "the decisive step is the replacement of the dative + *esse* [be] + nominative by nominative + *habere* [have] + accusative. The passive construction has been transformed into an active one" (Kuryłowicz 1931[1960]: 107). Against this classic analysis of a nominative shift paralleling the "active transformation", also proposed by the Indian scholarly literature mentioned above, Benveniste (1952), in a pioneering paper on the meaning of the perfect, proposed his own view of the perfect as basically possessive: the dative case marker in such periphrastic expressions with a participle was never an agent marker but a possessor marker, since classical Latin uses the dative and verb 'be' to represent possession.[8] The possessive construction *mihi filius est* ('I have a son') later shifted to the 'have' construction (*ego*) *filium habeo*, in the same way as the perfect. Consequently, the evolution of the Latin (and Persian) perfect has, according to Benveniste (1952), nothing to do with a shift from passive to active, but is a mere reversal ("renversement, retournement") of the old possessive pattern. Similar conclusions were presented later by Pirejko (1979) and by Trask (1979: 397) who associates the possessive origin with "the incorporation into the inflectional paradigm of a nominal form" with a genitive agentive complement.

2.3 Re-interpretation of the IA grammaticalization path: the new agent case markers in ergative Western languages as localizers

This famous analysis of the Indo-European perfect by Benveniste raises however a problem regarding the Indo-Aryan data, since it is exclusively based on case marking. As already mentioned, the agent in classical Sanskrit, particularly if it is a full NP, is standardly in the instrumental, the standard agent marker in passive sentences (see Montaut 2016 for details), and not in a possessive case. Yet, further developments of the "pre-ergative" pattern into a fully ergative alignment in Western IA languages provide arguments for considering the agent in a way similar to Benveniste's analysis of Latin. The new markers used to reinforce the old syncretic oblique case are indeed massively derived from location nouns. Moreover, they also served as a locative marker (such as the -*i* ending in Sant Bhasha in example (5e) above). Indeed, the most wide-spread form of the now ergative case markers is *ne* (*ni, nai, nẽ, ne*), found in Hindi/Urdu, Panjabi, Marathi, Gujarati, and it is derived from the locative (* *karṇasmin*, a renewal of the classical form *karne* by analogy with the major paradigm) of the noun *karṇa* "ear" (Tessitori 1914–1916: 65ff; for more details and examples see Montaut 2016). Then this form

---

[8] By contrast, the standard case marker for agents in passive clause is *a* + ablative in Latin, and similarly in Old Persian it is not a genitive but *hacama* + ablative.

underwent reduction along the following lines: *karnasmin > kannahĩ > kaṇṇhaĩ > kaṇṇaĩ > naĩ > nai > nẽ > ne*. Tessitori was the first scholar who identified the correct origin of *ne*, but already Trumpp (1872: 401) had traced *kane* ("near, at the edge", then "to") to the Sanskrit noun *karna* "ear". The origin of the ergative marker has later on been accepted by all traditional grammarians (Tiwari 1961, Tiwari 1966; Saxena 1937; Chatterji 1926; Chatak 1966). Tessitori (1914–1916: 68-70) gives examples such as (11a) with a clearly locative meaning and (11b) with an agentive meaning:

(11)   a.   *mithyādṛsthi loka   kanhai sravai   vasirau   nahĩ*
            false.look   people LOC   hermit.M.SG dwell.PRS.3M.SG NEG

            'A shravaka (hermit) should [does] not live near heretics.'

       b.   Old Rajasthani
            adiśvara   naï   diksā   lidhi
            Adishwara LOC/ERG consecration.F.SG take.F.SG

            'The Adishvara took the consecration.'

The second most widespread ergative marker is *le* (*lai, al*), found in Kumauni, Garhwali, early Nepali, and it is derived from the verbal root *lag* "to be in contact, touch" (*lagi/lāgi > laï, lai, le*) originally meaning "having come in touch with", "up to", "for the sake of". Reflexes of both markers are far more often used for dative (*nai/ne/nũ̃* in Rajasthani, Panjabi, Gujarati; *lā* in Marathi) than for instrument arguments (*ne* in Marathi, *-an/le* in Garhwali/Kumaoni). These connections should however not be overestimated, given the extremely weak semantic content of the initial etymon, and also given that the same ending −*i* was used in the Sant Bhasha of Kabir both for agents, as in (5e) above, and for locative complements, as in (12):

(12)   Sant Bhasha
       ābari   dīsai   ketā   tārā
       sky.LOC be.seen.3M.PL. how.many star.M.PL

       'How many stars we have seen in the sky!' (146.1)

Finally, there is the Gujarati ergative marker with its alternation *ne/e*, the first form being related to the above-mentioned *ne* whereas the *-e* ending is also used as a locative case marker (cf. Cardona & Suthar 2003: 678) and seems to have been the initial marking, since we find it in both functions in the *Jain Gurjar Kavyo*, dating to the 14-15th century (14):

Annie Montaut

(13) Modern Gujarati
Sītā.e   kāgal   vācyo
Sita.ᴇʀɢ letter.ᴍ.sɢ read.ᴍ.sɢ
'Sita read the letter.' (Mistry 1997: 426)

(14) a. Old Gujarati
jamunājī.n.e   tat.e
Yamuna.ʜᴏɴ.ɢᴇɴ.ʟᴏᴄ bank.ʟᴏᴄ
'On the bank of the Yamuna river.'[9] (Desai 1926: 630)

  b. *ame yamunā gayā~   (...) rokyã̄   nandanā   nānhaḍī.e*
1ᴘʟ Yamuna go.ᴍ.ᴘʟ   stop.ᴍ.ᴘʟ Nanda.ɢᴇɴ small.boy.ʟᴏᴄ/ᴇʀɢ
'We have been to the Yamuna (...), [we] were stopped by Nanda's boy /Nanda's boy stopped [us].' (Desai 1926: 630)

The specialization of a given case marker for a given function is clearly recent (and in some languages still very weak: Montaut (2015; 2016 and Stroṅksi (2010) quote various examples where the same marker is used for ergative, dative/accusative, and instrumental.[10] The original semantics of the words that grammaticalized to ergative case markers may explain such anomalous associations as agent and patient as distinct from subject (cf. Malchukov & Spencer 2009). Yet, a straightforward evolution of a construction with the passive participle into an ergative pattern with agents marked by originally localizing words is hardly plausible. One might be tempted to view the development from the passive participle in nominal sentences such as in (1) to the modern ergative construction given in (2a) and (13) as involving an intermediate stage which was neither a passive nor an active construction but a type of impersonal construction. This is a construction which Seržant's (2012: 374) calls a "free-dative-adverbial" complement as the agent (subsuming under "dative" the adhesive, dative, locative and genitive case). This construction becomes ambiguous with respect to subject properties (control and discourse-related) during the early NIA period, as shown in (14b). Later, marked objects blocked agreement, which then facil-

---

[9] Like in Hindi and all IA languages that maintained grammatical gender, the genitive postposition is an adjectiving suffix: the noun in the genitive agrees with the head noun in gender, number and case.

[10] Bangaru (Panjab-Haryana) is well-known for displaying the same case-marker, *nai* in certain dialects, *si* in others, for ᴇʀɢ, ᴅᴀᴛ/ᴀᴄᴄ and ɪɴs (for examples see Montaut 2007, Montaut 2015). Kului, a Western Pahari (Himachal, north of Panjab) also exhibits this peculiarity (example in Stroṅski 2014). Similarly, Maithili too displays "contradictory"uses of the postposition *so~* , ɪɴs/ᴀʙʟ, including for agents, and ᴅᴀᴛ/ᴀᴄᴄ (Jha 1958: 30).

itated reanalysis: This construction has been already identified as "impersonal" in constructions with a marked object in traditional grammars such as Kellogg (1972[1875]) and Chatterji (1926), as their literal translations such as "there has been V by X regarding Y" clearly demonstrate. It is in this abstract sense that the construction can be considered as a predication of localization: the process or the result is merely located in relation to the agent, not represented as the direct source as in the transitive model. Coming back to Benveniste and the Latin scenario involving the "have" auxiliary, Benveniste (1952) clearly states that "have" is but "an inverted be", that is, a stative verb, transitive only apparently. He later developed this notion in an article on French auxiliaries 1960 to explain the use of possessive predication in French and Romance languages for expressing transient states (*avoir faim, peur, mal* 'be hungry, cold, in pain'). On similar grounds (inverted stative predication rather than the direct transitive source-goal model), I suggested an analogy between ergative as well as dative subject sentences widespread in modern IA to experiential predicates (physiological or psychological processes) and predications of localisation (Montaut 2004). Currently, dative experiential subjects have case markers that are distinct from ergative agents in modern languages (except in Bangaru). Both fully-fledged ergative patterns and dative subjects arose, depending on the language concerned, between the 14th and 16th century They most probably derived from such localizing predications. This development is thus comparable to the reanalysis of impersonal intransitive constructions attested elsewhere; and it is particularly similar to the Slavic -*n/t* participle developing into perfect, as described by Seržant (2012). Significantly, constructions of the type exemplified in (1), (5), and (6) were also found with intransitive predicates, in Sanskrit as well as in early Western and Eastern NIA:

(15)   a. Sanskrit

ināhina [= *ina+*   *ahina*]   *sṛptam*
DEM.INS serpent.INS crawl.PPP.NOM.N.SG

'The serpent crawled.' (there has been crawling by the serpent) (Renou 1952: 198)

   b. *asmakam   abhipretam   bhavantam kimcid artham*
1PL.DUAL.GEN come.PPP.NOM.N.SG 2.ACC   INDEF for
abhiprastum
ask.INF

'We have come here to ask you something.' (*Mahabharata* IIIg9)

   c. Middle Bengali

Annie Montaut

> *duhẽ   thak.ila   eka paśe*
> both.ᴏʙʟ remain.*ila* one side.ʟᴏᴄ
> 'They both remained on one side.' (Chatterji 1926: 947)

Even the now highly anomalous clause *maĩ.ne gayā* 'I went', with the *ne* marker, reported to be frequently heard in the speech of migrants from Punjab to Delhi in the mid-20th century (Chatterji 1986: 71), can be explained by this scenario. This, as well as the diverging Eastern and Western evolutions, invites to question the now standard labelling of such patterns as "pre-ergative", also reflected in glossing of case markers, whether inflectional or postpositional (Assamese -*e* is never glossed ergative, and the language not considered ergative, Gujarati *ne/e* is glossed Agent in Cardona & Suthar (2003) - although the authors consider the language ergative - and ergative by Mistry (1997). Syncretic obliques are usually glossed ergative in modern studies on Old Hindi).

3 The grammaticalization of the modal verbal adjective: an areal complementary distribution?

Whereas the construction featuring the PPP developed into ergative alignments in Western languages only via the nominal sentence with an instrumental agent, this construction was realigned into nominative alignments in the East. In parallel, another passive participle also developed, also via nominal sentences, into patterns with non-canonical alignment, which were further realigned into nominative patterns in Eastern languages only. This passive participle is the form ending in -*tavya*, 'to be V.ed', also called gerund or verbal adjective of obligation. I will henceforth gloss this form as gerund, although it is morphologically a participle, inflecting like an adjective. This form underwent an evolution strikingly parallel to the evolution of past participle in Eastern languages, a fact which has long been noticed by the historians of those Eastern languages but not considered as a highly relevant fact before Montaut (1996; 2007; 2016). Let us first see this type of evolution before turning to its developments in Western languages.

3.1 The parallel grammaticalization paths of future and past in Bengali and other Eastern IA languages

In classical Sanskrit, the periphrastic construction involving a verbal adjective (sometimes called gerund or gerundive) was the standard way to express obligation. This pattern, illustrated in (16a), is parallel to the past tense pattern in

(1), but it is found both with transitive (16a) and intransitive verbs as shown in (16b) from Bloch (1906: 31). Example (17) illustrates both past and modal nominal sentences pervasive in the narrative register (*Tales of the Vampire*, in Bloch 1906: 59):

(16) a. Sanskrit
      mayā   tat   kartavyam
      1SG.INS DEM NOM.N.SG do.GER.NOM.N.SG
      'I have to/should do that. (lit. 'by me this to-be-done')'

   b. *yamayor   apramattayā   tvayā bhavitavyam*
      regulations.LOC.DUAL attentive.INS.F.SG 2.INS be.GER.N.SG
      'You should be attentive (non-distracted) regarding the regulations.'

(17) *mantriputreṇoktam (mantriputreṇa-uktam) "adya   tvayā*
      minister.son.INS.M.SG   say.PPP.N.SG now

      gantavyam";   tayoktam (=tayā.uktam) "gantavyam"
      2.INS   go.GER.N.SG 3.INS.F.SG   say.PPP.N.SG

      go.GER.N.SG
      'The son of the minister told her: 'Now you should go'; she said: 'should go'.'

This pattern remained unchanged in the various Prakrits of Middle Indo-Aryan such as the Magadhean variety in Ashoka (18), the direct ancestor of Bengali:

(18) *hida   no kimci   jive   alabhitu   pajohitavye no pi*
      here   no some   living   kill   sacrifice. no even
      NOM.N.SG CV GER.NOM.N.SG NOM.M.SG GER.NOM.M.SG
      ca samāje kattavye
      and assembly do

      'Here one should not sacrifice by killing a living creature nor hold a meeting here.'
      (lit. here living being killing should not be sacrificed nor meeting should be held)

The forms involving the gerund seem to have included a temporal meaning from as early as the 3rd century in the eastern region: according to Chatterji (1926: 966), the corresponding form in (19) from Ashoka's rock edict in Sarnath has "a vague mandatory sense, with an express future implication":

Annie Montaut

(19) a. Prakrit
      iyam   sāsane   vīnapayitavye
      DEM NOM.M.SG principle.M.SG make.known.GER.NOM.M.SG
      'This principle should/will be made known'. (made to be known)

Old Bengali (illustrated in (20a-c) from (Chatterji 1926: 967 ff) displays the same construction, both for intransitive and transitive verbs with the characteristic *-b-* suffix derived from the old *-tavya*, and agreement with the patient of transitive verbs, before the loss of gender blurred agreement and the ending *-ba* became invariable:

(20) a. Old Bengali
          toe   sāma kariba   maï   sānga
          2.OBL with do.*ba*(M.SG?) 1SG.OBL company.M.SG

          'I shall have union (do company) with you.'

      b. Old Bengali
          maï   dibi   piricha
          1SG.INS give.*b*.F question.F.PL

          'I shall ask questions.'

      c. Old Bengali
          tabe~ to.ka rakhiba   kona jāne
          then 2.ACC protect.*ba* which person.OBL

          'Then who will protect you?'

      d. *maĩ   jaivo (=jaiba?) govinda saha khelaṇa*
          1SG.INS go.*b*   Govinda with play.INF

          'I shall go to sport with Govinda.' (Chatterji 1986: 30)

Around the 15th century, this construction was transformed into a nominative one, with nominative (unmarked) agents and tensed verbs agreeing with the agent. The verbal form consists of the old *-ba–* form now suffixed with finite person endings. These suffixed person endings are the same as the ones used for past, both distinct from the present endings (–*ish* and *-o* for second familiar and second respectful person respectively):

(21) a. Modern Bengali
          tu boi.ta   por.bi
          2SG book.DEF read.F.2SG

          'You (familiar) will read the book.'

b. Modern Bengali
tumi   boi.ta   por.b.e
2HON book.DEF read.F.2HON
'You (respectful) will read the book.'

The verb morphology is now analysed as base + future marker *-b–* + person ending, in the same way as the past is now analysed as base + past marker *-l–* + person ending, both processes of reanalysis occurring at the same time. The above-mentioned grammarians also acknowledged this parallel in the evolution of Eastern IA languages as a common "active transformation". Chatterji (1926: 987) notes for Bengali that "the affixes are exactly on the lines of the past", in contrast with those for the present, and also that the shift in alignment occurred at the same time for future and past. Other Eastern IA languages also went through this "active transformation" but the respective outcomes are not as clear as in Bengali. In Awadhi, for instance, the sigmatic inherited future persisted and is still prevailing in certain persons (and similarly in Bhojjpuri), and Maithili, which now consistently displays a *-b–* future, has acquired a complex agreement system indexing several participants. But in older times the construction was maintained with the now-lost gender agreement still visible, for instance in Vidyapati (14th century):

(22)   a. Old Maithili
       sumarabi   mori name   … prema sumaraba
       remember.*b*.F my.F name.F   love.M remember.*b*.M
       'You will remember my name, you will remember my love.' (V 9, in Jha 1958: 494)

   b. *hamahũ nāgari sabe   sikhaūbi*
      1.OBL   lady.F all(=pl) teach.*b*.F
      'I shall instruct the ladies.' (V 52, Jha 1958: 495 495)

   c. *mañe   ki   bolaba sakhi apana geñāna*
      1SG.OBL how speak.*b* friend REFL experience.M.PL
      'O friend, how shall I speak out (my) feelings.'[11] (V 24, Jha 1958: 495)

Awadhi and Bhojpuri, which have now a complex paradigm in the future, also displayed the same morpho-syntactic pattern in their early stages before the shift

---

[11] In example (22b) *hamahũ* is the oblique form of the first person (singular and plural, today *hama*), whereas (22c) displays an oblique case of the base for first singular person, now gone out of use, with a case marking left unglossed by the author, but reminiscent of the palatal oblique forms for first person (Hindi *mujh, mujhe*).

Annie Montaut

to the nominative subject occurred. Similarly, Tiwari (1966: 171) observed for Bho-jpuri that "the future affixes for the first, second and third persons masculine and feminine singular and plural are in a line with those of simple past". Remarkably, in Awadhi, the -*ba* predicates combining with an oblique agent of the older pattern that prevailed till the 16th century could still convey necessity, which was no longer the case for the Bengali or Maithili equivalents. The resilience of the old sigmatic future in certain persons in Awadhi may explain why the -*b*– form took longer to grammaticalize into a future marker:[12]

(23) a. Awadhi
       kathā   bhāśā-baddha   karabi maĩ   so-saba hetu   kahaba
       story.F.SG language-ridden do.*b*.F 1SG.INS this-all because say.*b*.
       maĩ   gāī
       1SG.INS sing.CV

       'That story is to be composed by me in the vernacular, for this reason it is to narrated by me by singing.' (or: I shall compose, I shall tell by singing)[13]

    b. *ghara kaisai paiṭhaba maĩ   chũche, kaunu utara*
       house how   enter.*ba* 1.SG.INS deprived which answer.M.SG
       debau   ?
       give.*b*.M.SG

       'Empty how shall I enter the house, what answer shall I give?' (Saxena 1937: 261)

    c. *prāna-priya siya   jānibi,   nija kinkarī-kari mānibi*
       life-dear   Sita.F.SG know.*b*.F.SG REFL slave   consider.*b*.F.SG

       'Sita is to/will be regarded as beloved like life, she is to/will be accepted as thy slave.' (Chatterji (1986: 96) translates this form as obligative)

---

[12] Depending on the dialects: only the first person forms have the -*b*- ending in western dialects of Awadhi, and first and second person in eastern dialects. The -*h* form (inherited sigmatic future ṣ > s > h) prevails in the third person, but in early Awadhi -*b*- was found in all persons, although not systematically (Saxena 1937: 264-266).

[13] Same example is translated in Saxena (1937: 260) by a future form: *bhasa-badha karabi maĩ* "I shall render it in popular language".

3.2 Kuryłowicz's explanation of the parallel nature and development of perfect and future

This parallel evolution of past and future systems was also noticed by Kuryłowicz (1931[1960]; 1965) for Romance languages, based on different, but equally convincing morpho-syntactic evidence. The future in modern Romance languages is indeed a very peculiar innovation, involving the verb "have" constructed with the infinitive. The same auxiliary is used as in the perfect, the difference being only that in the future, the infinitive precedes the auxiliary, resulting in a fused form, whereas the auxiliary after the verb in the perfect is still free (cf. (10) above). Here are the forms in French and Spanish, with similar formation in Italian:

Table 1: First, second, and third person singular future forms of 'sing' in French and Spanish.

	1st person 'I will sing'	2nd person 'You will sing'	3rd person 'He will sing'
French	Je chanter.ai	Tu chanter.as	Il chanter.a
Spanish	Yo cantar.é	Tu cantar.ás	El cantar.á
	1SG sing.INF-1SG	2SG sing.INF-2SG	3SG sing.INF-3SG

The endings are either identical to (French) or derived from (Spanish) the present paradigm of 'have' (*j'ai, tu as, il a*). The French pattern is clearer, since past and future auxiliary have the same form (*j'ai chanté, tu as chanté, il a chanté*), whereas the past auxiliary displays an initial *h* in Spanish as a separate auxiliary (*he cantado, has cantado, ha cantado*). Historically, the "have" periphrastic future acquired its temporal meaning from the original meaning of necessity in Vulgar Latin, with the infinitive shifting from passive to active morphology (*cantari > cantare*). According to Kuryłowicz, 1931[1960]: 107, the striking parallelism between perfect and future (in contrast to the present) crucially relies on the origin of the *habere* periphrastic future: the 'have' periphrasis is for him the continuation of the Latin verbal adjective of obligation (sometimes called gerund or gerundive in -*nd*-):[14]

---

[14] The initially passive infinitive was replaced by the active infinitive as soon as late classical and vulgar Latin. The probable original pattern for (24b), as duly noticed by Benveniste, was then:*ego id cant.ari habeo* 1SG this sing.INF.PASS have.1SG 'I have this to be sung, to sing.'As for the dative pattern with obligation gerund, on the popular patterns in *delenda est Carthago* (Carthago is to be destroyed) or *mihi colenda est virtus* ('Virtue is to-be-culvated to me'), all the

Annie Montaut

(24) Latin
      a. mihi   cantandum est > b.   (ego)   cantare   habeo
         1SG.DAT sing.GER   be.PRS.3SG 1SG.NOM sing.INF.ACT have.1SG
      'I have this to be sung, to sing > I will sing' (Kuryłowicz 1931[1960]: 107)

The *-nd-* verbal adjective was used with a dative agent and agreed in gender
and number with the nominative patient, like the *-tavya* sentences in (16)-(17).
But unlike in the latter, the copula was required (agreeing with the nominative
patient), as shown in (25):

(25) a. Latin
          mihi   sit   referenda   omnis illa oratio
          1SG.DAT be.PRS.3SG relate.*nd*.F.SG all   this discourse.F.SG.NOM
          '[to which] I should relate this whole discourse.' (Cicero: *De Oratore*
          II.114)

      b. *nunc est   bibendum*
          now be.PRS.3SG drink.*nd*.N.SG
          'Now one should drink.' (Horatius)

Thus, we can summarize the Latin evolution of perfect and future in Table 2
below, which highlights similarities to the IA data. In a very illuminating paper
on the evolution of grammatical categories, Kuryłowicz (1965) further developed
the hypothesis of a deep similarity between these two tenses: on his account, per-
fect and future are both basically non-active because they do not aim at depicting
an action, but at representing viewpoints, from the present, on this action. This
contrasts with the present that aims at directly representing a process.

Benveniste (1966[1965]) reached similar conclusions regarding the symmetry
of past and future (both "orthogonal" relatively to the present, both representing
perspectives on an action rather than processes proper). However, Benveniste
(1966[1965]: 131) radically rejected Kuryłowicz' interpretation of a parallel evolu-
tion from Latin, arguing that the Latin construction involving *habere* never had
an obligative meaning and was used as a "future of predestination" in Christian
preachers, in the meaning "fated to happen". This latter argument has later on
been proven wrong, as well as the accusation against Kuryłowicz of mistaking

---

numerous examples quoted in Touratier (1994: 164 ff) as well as other examples in the most
classical writers such as Cicero, it always has a modal, not temporal, meaning : *senibus labores
corporis minuendi sint* ('[So that] old people should minimize physical works' Cic *De Officiis*
1.123) ; *cum haec nobis quaerenda sit in causis* ('As in our causes we have to research these
[goals]' Cic *De Oratore* II. 120).

Table 2: Alignment development from Early/Classical Latin to French.

alignment	perfect	future
Nominative: synthetic verb form, NOM subject, person agreement (Early/Classical Latin)	*(ego) feci/cantavi*	*(ego) faciam /cantabo*
Non-nominative: participial verb, DAT agent, no person agreement (Late Latin)	*mihi factum/cantatum est*	*mihi faciendum /cantandum est*
Nominative: V + have, NOM subject, Person agreement (Vulgar Latin)	*(ego) factum/cantatum habeo*	*(ego) fieri/cantari habeo*
Nominative: V+ have, NOM subject, person agreement (French)	*j'ai fait/ chanté*	*je fer.ai/chanter.ai*

a passive infinitive for an active infinitive. Yet, Benveniste's major reason for rejecting the symmetry of the two developments might have had something to do with his own thesis of the "possessive perfect", since it seems more difficult to derive the future meaning from the possessive (more details in Montaut 1997). The ambiguity between the original meaning of obligation (or potential) and the new meaning of future was attested between the 2nd and 3rd century by grammarians such as Tertullian and Pompeius. On the other hand, the future meaning was found only later (in conformity with Bybee, Perkins & Pagliuca (1994)'s generalizations), mainly in texts by Christian writers and grammarians, with no particular connotation of predestination. Moreover, it is almost exclusively used in learned texts, chiefly from Africa. Its earliest - and still ambiguous - instance in a colloquial register is found in Wâdi Fawâki's letters (cf. Adams 2011; 2013: 659).[15]

---

[15] Compare the ambiguous (future/alethic) example from Tertulian quoted in Adams (2011: 148): *si enim sustuleris istam tertiam, remanere habent duae* 'for if you take away the third (syllable) two will (have to) remain' (GL 129.6). Wâdi Fawâki' letter to Rustius Barbarus is also ambiguous

To my knowledge, no study points to a direct transformation of the Latin dative alignment with the *-nd-* gerund into the nominative alignment with *habere*, but this reanalysis is basically equivalent to the transformation regarding the perfect. The striking event in the history of the Romance languages is the total extinction of the Latin synthetic future and the subsequent formation of new futures, often on the basis of an obligative periphrasis (*habere*). Thus, grammaticalization of obligation into future is undoubted (Adams 2011; Bourova & Tasmowski 2007), not least because it conforms to a typologically common source for futures (Heine & Kuteva, this volume). In this respect, it parallels the Eastern IA data, which display an observable, continuous history from non-nominative alignment to nominative alignment.

3.3 Divergent evolution of gerund in Western IA languages

However, this evolution from modality to future is not pan-Indian, and this is a strong difference with the past in the general evolution of Indo-Aryan. Whereas the *-ta* form, being the normal expression of past, rapidly became the substitute for all verbal forms of the past, the *-tavya* form never became the normal expression of future because the old synthetic future was maintained in many regions and prevented the new periphrasis from extending to the field of future. In Western IA languages, the old verbal adjective (V.ADJ) of obligation was maintained in its original meaning up to the middle stage of NIA. For instance, the ergative-like pattern of Ashoka (early MIA), presented as the origin of the Eastern *-b-* future in (18) above, has a Western equivalent in Girnar (now Pakistan):

(26)   *hida*   *na kimci jivam arābhitpā prajuhitavyam na ca   samājo*
        here   no some living kill   sacrifice.   no and assembly
        NOM.N.SG   CV   V.ADJ.NOM.N.SG   NOM.M.SG
        kattavyo
        do
        V.ADJ.NOM.M.SG
        'Here one should not sacrifice by killing a living creature nor hold a meeting.'
        (lit. here should not be sacrificed killing a living being nor should meeting be held)

The same pattern, with a *-v–* form, continued till late MIA (Apabhramsha stage,

---

between future and deontic modality: *adferre habes* 'you have to bring' or 'you will bring' (Adams 2013: 659).

turn of the millennium) for instance in the well-known Western Jain text *Pau-macariu*, with or without oblique agent, clearly patterning like the past sentences exemplified in (27c):

(27) a. Apabhramsha
       annu   na nam.ev.au
       other.M.SG NEG respect.V.M.SG
       'No other is to be respected.' (Paumacariu 26.3.2)

    b. *navara ekku vau   maĩ pālevau*
       only   one vow.M.SG 1SG.INS keep.ev.M.SG
       'I shall/should observe only one vow.' (Bubenik 1998: 194)

    c. *tā   keumaiẽ   haũ gharaho nīya*
       then Ketumati.INS/OBL 1SG home.LOC conduct.PPP.M.SG
       'Then I was taken home by Ketumati'/'Then K took me home.'
       (Bubenik 1998: 148)

Those NIA languages that inherited this state of affairs and also retained the sigmatic future (e.g. Old Gujarati, Marathi, Rajasthani) maintained the *-tavya* pattern with an obligative meaning till their middle NIA stage, contrary to the Eastern languages which started shifting to a future meaning during their early NIA stage. Old Gujarati, for instance, shows agreement with patient and instrumental agent (*-im*), and the obligative meaning in the 16[th]century text *Upadeśamālā*:

(28) a. Old Gujarati
       isī   upamā   jāṇivi
       such.F.SG comparison.F.SG know.v.F.SG
       'Such a comparison should be known.' (Dave 1935: 64-65)

    b. *ṣiṣyiĩ te   kārya   tatkāla   ācarivaũ*
       pupil.INS DEM.M.SG work.M.SG immediately do-v-M.SG
       'The pupil should immediately do that work.' (Dave 1935: 94)

    c. *te pāpiu jāṇi.v.au*
       3M.SG.INS know.v.M.SG
       'He should be considered as a sinner.' (Dave 1935: 54)

This is a direct continuation of the 14th century language found in the *ṣadā-vaśyaka*, the oldest written testimony of Old Gujarati. According to Pandit (1976: 23), the "gerund" used as a predicate had an "imperative [= obligative] sense" in this variety: *rakhivaũ* in the masculine 'is to be saved', *karivaũ* 'is to be done',

vyavasthāpivi in the feminine 'is to be founded'. In Old Rajasthani, similar obligative constructions are still found in the 16th century, the time when the language is supposed to have diverged from Old Gujarati. The obligative construction is attested both with intransitive as shown in (29c) and transitive predicates, the latter both with and without an agent, given in (29b) and (29a) respectively:

(29) a. Old Rajasthani
          hisā   na   kar.av.ī
          violence.F.SG NEG do.*av*.F.SG
          'Injury is not to be done.' (Tessitori 1914–1916: 120)

      b. *anere   vidya   lete   vinay   kar.iv.um*
          other.PL.INS knowledge taking humility.M.SG do.*iv*.M.SG
          'Humility should be observed by others acquiring knowledge.'

      c. *taĩ   na   jāi.vu*
          2.INS NEG go.V.M.SG
          'It should not be gone by you.' (Khokhlova 2013: 101)

Old Marathi also displays similar constructions, which Bloch (1970[1920]: 264) presents as parallel to the past constructions, with example (30): the "syntax, with the logical subject in the instrumental, [is] very similar to that of the form for past".

(30) Old Marathi
      majhyāne~ cālavle~   / dhadā   sikhavlā
      1SG.INS   go.POT.PST.N.SG / lesson.M.SG learn.POT.PST.M.SG
      'I could/was able to go / to learn the lesson.' (Bloch 1970[1920]: 265)

Bloch also mentions a dative alternation for the agent (*majhyānẽ* or *mālā*).

3.4 The Marathi case: an exception to the areal complementary distribution

The -*tavya* obligative gerund was, as a rule, lost in modern Western IA languages, whereas it developed into a future with nominative realignment in the Eastern IA languages. On the other hand, the construction with the Sanskrit participle -*ita* in the past was maintained with the original alignment and developed into an ergative construction in the West, whereas it was realigned to a nominative pattern in the East. One could conceive of this as a complementary distribution between languages maintaining the non-nominative construction of the PPP on

the one hand and languages maintaining the *-tavya* form as a predicate on the other. Marathi is however an exception, since the modern language still displays the same pattern as in (30), albeit with the new case markers. Note that in modern grammars, the *-av-* suffix is usually glossed as "subjunctive" (sometimes as "potential"), and the marker *ne* is usually glossed as agent in such constructions. However, it is glossed ergative in past constructions (Dhongde & Wali 2009: 44), and now occurs only in the 3rd person, although in the unmarked 1st and 2nd person, the verb still agrees with the object.[16] Marathi is hence exceptional in two respects: one, unlike other Western languages, it maintained the *-tavya* predicative form; two, unlike Eastern languages, it maintained this form in its original obligative meaning and alignment.

(31) Marathi

 a. *tyāne~*   *ghari~*   *yā.v.e~*
    3M.SG.INS/AG home.LOC come.*av*.N.SG

    'He should come home.' Marathi

 b. *tyāne~*   *cĩc*   *khā.v.ī*
    3M.SG.INS/AG tamarind.F.SG eat.*av*.F.SG

    'He should eat tamarind.'

This construction has a nominative counterpart (32), with the verb agreeing with the subject. It is analysed as potential and not obligative in Wali (2004) and Dhongde & Wali (2009), but for Pandharipande (1997: 290) both meanings are conveyed by the (31) construction.

(32) *to*   *ghari~* *yā.vā*
    3M.SG.NOM home.LOC come.SBJV.3M.SG

    'He may come home.'

What is clear from these various evolutions is that wherever a different form for the future was available - be it the inherited sigmatic future in Gujarati, Rajasthani, and Braj, or a new form *-l-* as in Marathi - the *-tavya* verbal adjective retained its modal value (Marathi) or disappeared from the TAM paradigm altogether (other Western languages). All Western languages except Marathi indeed developed periphrastic constructions with verb "be" or "need" and dative 'subjects in non-canonical alignments (see ex. (36) below). The reason why only

---

[16] The verb also agrees with the agent in the 2nd person: *tyāne samayā ghas.l.yā* (3SG.ERG lamp.N.PL wash.PST.N.PL) 'he washed the lights', vs *tu samayā ghas.l.yā.s* (2SG.ERG lamp.N.PL wash.PST.N.PL2SG) 'you washed the lights'.

Annie Montaut

Marathi maintained the original construction and verbal form is most probably because it is also the only language which retained the old Sanskrit infinitive in -*tum*. In contrast we observe the presence of −*v*- /-*b*- infinitives inherited from the -*tavya* form in all languages which lost the obligative meaning of the predicative -*tavya*.

4 From the -*tavya* gerund to infinitive and verbal noun

The original, invariable Sanskrit infinitive (-*tum*) rapidly fell out of use in Indo-Aryan and the category itself almost disappeared, replaced in most modern languages by verbal nouns in -*an/ana* or, in the East, by former participles. Marathi is the only language to maintain the old form and category. Gujarati is another exception regarding the infinitive and verbal noun, since it never developed an -*an* verbal noun and instead marked both the infinitive and the verbal noun by a single form derived from the -*tavya* gerund. This development also occurred in other languages but was ultimately inhibited by the growth of the -*an* forms.

4.1 The Gujarati infinitive: infinitive and inflecting verbal noun

The single form of the infinitive in Gujarati is -*vũ*, which scholars relate to the old verbal adjective in -*tavya*: Dave (1935: 64) and Chatterji (1926: 966) for instance derive the Gujarati verbal noun *karvũ̃* from the verbal adjective or passive obligative participle *kartavyam*. This verbal noun agrees in case, similar to the -*an/ -ana* form in other IA languages: *jovũ* 'to see', *jovā lāyak* 'worth seeing'.

(33) a   *khā.v.ā   lāyak   vastu che*
      DEM eat.*v*.OBL worth thing be.3SG

      'This is a thing worth eating.'

In periphrastic permissive constructions with the verb *de* "give" and inceptive constructions with the verb *lag* 'touch/start', the verbal noun similarly displays the oblique form -*ā*: *karvā de* 'permit to do', *khāvā lag* 'start to eat' (cf. Cardona & Suthar 2003: 688; Dave 1935: 52 ff.).[17] Used in obligative constructions with 'be', it agrees in gender and number with the object, like the -*an* infinitive forms of other NIA languages in obligative constructions. But unlike other NIA languages, the Gujarati -*v*- infinitival form, in combination with auxiliary 'be' or 'need', can

---

[17] In these two constructions, languages with -*an/na* infinitives also display the oblique form of the verbal noun (Hindi *karne do* [do.INF.OBL give.IMPER] 'allow [x] to do/let X do', *khāne lagā* [eat.INF.OBL start.M.SG] '(he) started to eat'.

be used to convey not only obligation, but wish too, as shown in (34)b and (35)b, respectively. The meaning of obligation is conveyed by an extended form of the infinitive (-*v-an-*), with the relational suffix -*an* added (cf. Cardona & Suthar 2003: 677), whereas the desiderative meaning requires the short form -*v*-. Both forms are constructed with the copula:

(34) a. *mar.e   caupḍī   vanc.van.ī   che   / lekh   lakh.van.o*
          1SG.AG book.F.SG read.*van*.F.SG be.PRS.3 / article.M.SG write.*van*.M.SG
          che
          be.PRS.3S
          'I have to read a book / to write an article.'

      b. *tam.ne kyā   jā.vān.u   che*
          2.DAT where go.INF.N.SG be.PRS.3SG
          'Where do you have to go?' (Cardona & Suthar 2003)

(35) a. *mar.e   caupḍī   vanc.v.i   che   lekh   lakh.vo*
          1SG.AG book.F.SG read.v.F.SG be.PRS.3 article.M.SG write.v.M.SG
          che
          be.PRS.3SG
          'I want to read a book, to write an article.'

      b. *tam ne   kyā   ja.v.u   che*
          2   DAT where go.INF.N be.PRS.3
          'Where do you want to go?' (Cardona & Suthar 2003)

It should be noted that, although the meanings are very close to the Marathi obligative or potential sentences with finite forms in -*v*-, the morphology here is that of a verbal noun. This is exactly parallel to the -*an* formations of Hindi or other languages with gender/number agreement of the verbal noun. Interestingly in Gujarati, case marking is not limited to the dative as in other Indo-Aryan languages like Hindi, but exhibits an alternation between dative and agentive. Both are possible with the same pronoun as shown in example (36a), which expresses necessity with the 'need' auxiliary *joie* + verbal noun. The Hindi/Urdu equivalent, with verbal noun in -*n*- and *cāhie* (etymologically from a 'look' verb, like the Gujarati *joi*), has the same agreement pattern (default agreement with intransitives, object with transitives), but no alternation with the ergative/agentive. Punjab Hindi does display case alternation, depending on dialects (cf. Khokhlova 2013):

(36) a. Gujarati

Annie Montaut

mār.e / ma.ne   ghar jā.v.ũ   (gujarati bol.v.i) joie
1SG.AG / 1SG.DAT home go.INF.N.SG Gujarati.F speak.v.F need PRS.3SG

'I should go home. (speak Gujarati)' (Cardona & Suthar 2003)

b. Hindi/Urdu
mujhe ghar jā.nā   (gujarati bol.nī)   cāhie
1SG.DAT home go.INF.M.SG Gujarati.f speak.v.f need PRS.3SG

'I should go home (speak Gujarati).'

c. Punjab Hindi/Urdu
maĩ.ne / mujhe   ghar jānā   hai
1SG.ERG / 1SG.DAT home go.INF.M.SG be.PRS.3SG

'I should go home.'

The agent in (36c) is usually glossed ergative because of its identical form
(*ne*) with the specific ergative marker in transitive past clauses in Hindi/Urdu
and Punjabi. In sentences like (36a) and (33-34) above, the Gujarati *-e* is usu-
ally glossed agent whereas the same marker is glossed ergative in transitive past
clauses with similar agreement pattern (like in example 13 above). Despite the
variable glosses, it seems that Gujarati is shifting to dative agent, the standard
expression of obligation in all languages with the new *-an* verbal nouns, origi-
nating from a construction with an instrumental agent. The predicative *-tavya*
verbal adjective was used to express obligation in the old language (like through-
out middle IA), but a copula construction also emerged in the 16th century, where
the *–tavya* form is an infinitive and no longer a predicate by itself, similarly con-
structed with an instrumental agent:

(37) Old Gujarati
jīṇaĩ   jīviĩ   jīhã̄ jāi.v.aũ   chai
REL.INS person.INS there go.v.M.SG be.3M.SG

'The person who is designated to go there.' (= who should go) (Dave 1935:
67)

Outside this construction, verbal nouns in *-v-* were also as in the modern lan-
guage commonly used with the meaning 'worth of V', and as an action nominal
(*āsana naũ le.v.aũ* 'the taking of the seat', *viṇasi.v.ā nai-kāji* 'in order to be de-
stroyed', cf. Dave 1935: 54).

(38) a. *pāsachā   siũ aṇamila.v.aũ   bhalaũ*
depraved.OBL with not.mix.v.M.SG good.M.SG

'It is good (the good is) not to mix with the depraved.' (Dave 1935: 64)

b. *pāsachā   siũ boli.v.ũ,   ekaĩ   upāśrayi rahi.v.ũ*
   depraved.OBL with speak.*v*.N.SG one.LOC hostel.LOC stay.*v*.N.SG
   'To speak with the depraved, to stay at the same hostel.' (Dave 1935: 64)

4.2 Development and specialization of the -v/-b- infinitive

This -*v*- infinitive is present from late MIA onwards (Pischel 1971[1900]: 388). In the 11th century Jain Digambara texts (cf. Tagare 1948: 322) it is attested as a verbal noun, in parallel to its predicative use as in examples (27) and (28). Both uses are maintained throughout early NIA and continue to be used in many modern Western languages such as Braj (*māribaũ, māribau* 'to strike'), Rajasthani (*mārabo*), and Kanauji further north (*māribo*). This long tradition is documented for the older stages of Rajasthani Gujarati in Tessitori (1914–1916: 121):

(39) Old Rajasthani
      jīp.ava   vaṃchai
      win.*ava* want.PST.3SG
      'He wishes to conquer.' (Tessitori 1914–1916: 121; Chaya to Yogasāstra III-134)

In modern languages with two forms of the infinitive, the distribution is not quite clear. Braj for instance seems to select the -*n*- form for goal functions and as a complement of inceptive verbs (40c-d) and the -*b*-/-*v*- form for more nominal uses (40a-b), as illustrated by the 18-19th c. examples given by Kellogg (1972[1875]: 289):

(40) a. Braj
         sadā   kahu   saũ rah.iv.au   nāhi
         always somebody with stay.*iv*.M.SG NEG
         'One does not remain always in the same company.'
         (Lit. 'There is no staying always with anybody')

      b. *mere putrani kaũ pandit   kar.iv.e jog   hai*
         my   son.PL ACC wise.men do.*iv*.OBL worth be.SG
         'You are competent for making my sons wise men.'

      c. *tum saũ kah.an kaũ āyau haũ*
         2   SOC say.*an* DAT come PFT.1SG
         'I have come to tell you.'

Annie Montaut

    d. *rājā kah.ani lāgyau*
      king say.*an* start.M.SG
      'The king began to say.'

Snell (1991: 16) suggests that in classical Braj (16-18th ct.), the *-b-* form was chiefly used with inflected verbal nouns, while the *-na-* form occurred with semi auxiliaries such as *de-* (permissive), *pā* (ability) and *lag* (inception). Both forms were also employed to signal goal functions.[18] Bundeli, a South-western language sometimes considered a Hindi dialect, also has two types of infinitives. The *-na* form is preferred in permissive and inceptive constructions (*moṛā khõ khānā do* "let the boy eat" (cf. Jaiswal 1962: 132) while the *-b-* form typically occurs in nominal uses.[19]

(41) a. Modern Bundeli
        daur.b.o   ūke   cala.b.e   barābar hai
        run.*b*.M.SG 3.SG.GEN walk.*b*.OBL same   be.3.SG
        'Your running is equal to his walking.' (Jaiswal 1962: 133)

     b. *tumāe kara.b.e khõ ite   bohota hai*
        2.OBL do.*b*.OBL DAT here much be.3SG
        'There is much for you to do here.' (Jaiswal 1962: 133)

Whereas all Western languages at some point displayed, and many of them still display a *-v-/-b-* infinitive, the Marathi exception can be accounted for by the resilience of the inherited Sanskrit infinitive. It is also the only language which maintained the predicative gerund in its original obligative meaning. Since in Eastern languages the gerund developed into a future marker, one might also think that languages which lost the predicative construction turned the gerund into a verbal noun. On this account, three grammaticalization paths of the gerund are in complementary distribution: one, obligation and potential with new case markers as in Marathi; two, future with realignment in Eastern languages; and three, infinitive and verbal noun as in Gujarati and other Western languages. But it is obvious that the *−b-/-v-* verbal noun is also present in Eastern languages, although to a lesser extent than in the West.

---

[18] For instance, as a complement of movement verbs: *mokõ pakaran kõ āyo* 'he came to touch me', *mokõ spars karive kaū doryau* 'he ran to touch me'.

[19] The *-na-* form alternates with the -b form in presence of the agent suffix *-baro* (*karna baro* 'doer', *khabe baro* 'eater') whereas the *-b-* form is only allowed with the *-aiya* suffix (*dekhabaiya* 'seer').

4.3 The infinitive in the Eastern languages

Modern Maithili has three forms of infinitive (*-ana/āna, -al/ala, -aba/ab*), the two latter forms alternate with postpositions: *dekh.lā me sunnar* or *dekh.bā me sunnar* 'beautiful to look at' (Jha 1958: 519). Similar examples were already attested in Middle Maithili with the same *-ba/-va* form, and with a wider range of functions, such as *kopahū kara.vã joga* 'fit for showing anger' (Vidyapati 50, in Jha 1958: 519), or as a verbal noun in inceptive and volitive constructions:

(42) Middle Maithili
      kamal.āsana kichu   kahavã   lāgu
      lotus.seat   something say.*v*.INF;OBL start.M.SG
      'The lotus-seated [god] began to say something.' (Jha 1958: 611)

Assamese (North East) also has an infinitival form, ending in *-iba*, which is included into the suffix *-(i)ba.logya* 'worth to' (*sa.ba.logyia* 'worth to see'), and is required as a nominal formative and in concatenation with the modal 'be able':

(43) a. Assamese
          za.b.âr   xâm.ât
          leave.*b*.GEN time.LOC
          'Time of leaving.'

      b. *xi   saikel   sâla.bo par.e*
          3M.SG bicycle ride.*bo* be.able.3M.SG
          'He can ride a bicycle.' (Goswami & Tamuli 2003: 425)

Similarly, the standard Oriya infinitive ends in *-ibā* and is used as a noun before postpositions (*ās.ibā ku* '(in order) to come'). Although modern standard Bengali does not display *-b-/-v-* infinitives, the variety spoken in Assam (Tunga 1995) has similar forms used for complements of modal verbs (*karibar paro* 'I am able to do'). Grierson (1903-28) mentions Bengali constructions such as *ja.b.ār somoy.i* [go.*b*.GEN time.LOC] 'at the time of leaving', *korbā lāgil* 'he began to do', continuing from Middle Bengali *lage balibar* 'he began to say' (cf. Chatterji 1926: 1008). Besides, standard Bengali still uses it in nominal functions, particularly in the genitive (*-ar* ending): *ja.b.ar somoi* 'at the time of coming'; *as.(i)b.ar janye* 'for coming'. As for mid-Eastern languages such as Bhojpuri and Awadhi, they have *-be/-bu* verbal nouns alternating with the *-ana* forms in both their modern and early stages (cf. Saxena 1937: 282-5). So, it seems that, apart from Marathi, which maintained the old Sanskrit infinitive, only standard Hindi/Urdu and Panjabi,

the so-called central IA languages, do not have the -*b/v* infinitive. However, it should be remembered that Hindi and Urdu , sharing as common ancestors Braj, Avadhi, and, most importantly, the popular literary *koine* of the mystic preachers, the Sant Bhasha, and Panjabi also shares the Sant Basha ancestry.[20] They retained only the -*an/-ana* infinitive, which is used in obligative constructions as well as a verbal noun, and they developed a new periphrastic -*gā* future, so that the old verbal adjective has left no trace in the standard modern stage of these languages. But on the whole, more languages still display the infinitive more than any other tense or modal form derived from the old -*tvaya* verbal adjective. This can be seen as an indication that the basic meaning of the form was more nominal than verbal (or modal).

4.4 Original meaning of the so-called obligation gerund

A last piece of evidence for the remarkable persistence of the nominal meaning associated with the -*tavya* form in New Indo-Aryan comes from Romani, a language separated from West-central IA in the beginning of the 2nd millennium. In all Romani dialects, infinitive is notoriously absent as in most languages of the Balkans (cf. Boretzky 1996). While an infinitive form has been created out of contact with various local languages, the old -*tavya* form was however maintained as a noun of action in -*iben*, which is the only suffix for deverbal and deadjectival abstract nouns in the South Balkan dialects. For instance, the noun referring to a beverage, that is, something worth drinking, is *zspiben*, from the verbal IA root *pi* 'drink' (cf. Beniśek 2010). Similarly, in most Romani dialects, deverbal abstracts end in -*iben*,[21] a formation parallel to the other gerundival suffix -*niya/ñya* 'able to', 'worthy of' which also produces deverbal abstracts such as *pāj* (common Romani). These various evolutions cast doubt on the supposedly original meaning of the gerund, considered by Chatterji (1926: 966) to have developed "side by side" from obligation into infinitive and future.[22] Far more plausible from a typological perspective is the evolution from verbal noun to modal and future. Going back to

---

[20] Panjabi's founding texts (*Guru Granth Sahib*) displays a mixture of Sant Bhasha texts, many from Kabir and early Panjabi still not much differentiated from other dialects of what will become the Hindi language later. Although associated to Fariduddin Ganjshakar (13th ct.), the rise of Panjabi as distinct from various neighbouring "Old Hindi" dialects occurs only later. As for Hindi/Urdu, recent under these names, their earlier forms are regional languages from Braj to Awadhi.

[21] Deadjectival abstracts usually end in -*ipen*, a suffix inherited from the IA form -*pan*, and many dialects present a merger of the two forms.

[22] "The simple future notion evolved gradually; side by side with it, the old notion of an action to be done continued, and was modified into simply the notion of an act".

the origin of the *-tavya* form in Old Indo-Aryan, this form is itself derived from the verbal noun in *-tu*, the same which in the accusative was used throughout classical Sanskrit as an infinitive (*-tum*) and maintained in Marathi with the *-un* infinitive ending (*-idum* > Maharashtri *-iuṃ* > Mod. Marathi *-un*). In the dative case, *-tave*, the form was also used as an infinitive (*-tave* > MIA *-tae*, > Ardhamagadhi *-ttae*). This *-tave* verbal noun could also accept the gerundival suffix *-ya* (cf. Debrunner 1954: 612-615), which only later specialized in necessity constructions while the other gerundive occurring on the *-ana* nouns of action (*-anīya*) tended to be lexicalized (cf. Bubenik 1998: 190). It is also worth mentioning that the nominalising (deadjectival) suffix *-pan*, still used in NIA languages, is also supposed to derive from a gerundive ending *-tva* (Tagare 1948), suffixed with *-an* (*-tvana* > *-pan*). This suffix is very productive and is either directly suffixed to the base (such as in *bac.pan* "childhood" < *baccā* 'child') or to the oblique or direct form of the adjective (such as in *akele.pan* or *akelā.pan* 'solitude' < *akelā* 'alone, lonely'). Not surprisingly, it occurs in classical Sanskrit as a postpositional NP, patterning exactly like (43a):

(44) *na   ayam vaktavya.sya   kālaḥ*
       NEG this   speak-*tavya*.GEN time

      'It is not the time of speaking.' (from *Panchatantra*, in (Bloch 1970[1920]: 278)

   The Latin data leads to similar questions regarding the original meaning of the verbal adjective or gerund of necessity in *-nd-*. Whereas the well-known construction with copula and possible agent in the dative mentioned in §3.2 has always a necessity reading, conveying both passive voice and obligation, the form itself seems to be more general in meaning. This conclusion was first formulated by Ernout & Thomas (1951: 285): "il exprime simplement l'idée verbale", then developed by Touratier (1994: 164ff) who insists that the construction, not the form, conveys the meaning of necessity in predications with the verb 'be', with an optional dative agent (45a). Used as an attribute the *-nd-* verbal adjective obtains a vaguely abilitative meaning like the *-able* suffix (*orator legendus* 'an orator worth reading', 'a readable orator'), but it is mostly used, in Touratier's words, "merely as an infinitive" since the beginning of the common era, devoid of any modal meaning (45b):

(45) a. Latin

Annie Montaut

> *consola-nd-us*   *hic*   *mih-ist*   (= *mihi est*)
> console-*nd*-NOM.M.SG this.NOM.M.SG 1SG.DAT-be.3M.SG
>
> 'I have to console him.' (= he is to be consoled to me) (Plautus, in Touratier 1994: 165)

b. *de*   *consilio*   *relique-nd-i*   *Italiam*
   about decision.ABL leave-*nd*-GEN Italy.ACC

   'Regarding my project of leaving Italy.' (Cicero from Touratier 1994: 165)

In non-predicative constructions, the *-nd-* gerund is clearly devoid of obligation, a meaning which grammaticalized only in predicative constructions with the verb 'be'. Countless formations attest to this nominal behaviour, particularly its use as genitive complements, from the well-known *Ars ama.ndi* ('art of loving') to the three *libidines* ('desires') stigmatized by Augustinus (*libido domina.ndi, scie.ndi, frue.ndi* 'desire of domination, knowledge, enjoyment'). The difference with the Indo-Aryan evolution is that a further development into a standard infinitive never took place, because of the preservation of the inherited infinitive in *-re* throughout classical and vulgar Latin, still present in modern Romance languages. As a result, the gerund form disappeared altogether in Romance languages. But its initial meaning of expressing a mere verbal notion, common to both Old IA and Latin, nicely accounts for the various grammaticalization paths of this "gerund", richly illustrated in NIA and also in various stages of Latin (though not in modern Romance languages.

5 Conclusions

Regarding the grammaticalization paths of TAM-markers, the various developments in Indo-Aryan do not neatly mirror areal sub-classification, neither in meaning nor in form. Two pure tense markers, *-l-* for past and *-b-* for future, were the Eastern outcome of the grammaticalization of the past passive participle and the gerund, respectively. Except for the Marathi *-l-* past, none of these occur in Western languages in these functions. As for the meaning, the development of definite past out of the old resultative participle is omnipresent, whereas the development of the gerund into future is limited to the Eastern group, and its development into a modal predicate (*-av-*) is restricted to Marathi. Both developments contrast strikingly with the Latin-Romance data. In the latter languages, the perfect of comparable origin remained tense-wise a perfect (only recently encroaching on the definite past domain in spoken French). This is because of

the copula being present right from the initial step, but even more because of the resilience of the old aorist. The omnipresence of the "have" future based on a modal predication also differs from the restricted development of future in IA, even though it is of comparable origin: the resilience of the old future in the West IA accounts for its limited development in IA as a whole, whereas the Latin future disappeared in all Romance languages. Similarly, the development of the gerund form into an infinitive was apparently inhibited by the persistence of the old infinitive throughout Romance, a form which was lost. By contrast, the loss of the infinitive in all IA languages, except Marathi, allowed for an extensive development of the gerund into verbal noun and infinitive. The blocking effect of existing forms in the same function has obviously been decisive, and the co-existence of old and new forms, for instance for IA infinitives, usually corresponds to a specialization (i.e. one of them behaving more like a noun, the other more like a verb). As for the various alignment shifts, which show an areal distribution in IA, in contrast to the Romance uniform realignment in nominative patterns, they can be better understood by taking into account all the correlated patterns. Labelling the original nominal sentence with instrumental agent "pre-ergative", because of its further developments in the West, has been shown to be misleading. It is also misleading as it masks relevant correlations with both intransitive predicates and modal-future meanings which prevailed throughout IA up to the 15th century. As to why the Eastern non-nominative perfect realigned with nominative agent (and not the Western one), a possible answer lies in the parallel realignment of future clauses. I consider this explanation more plausible than contact with Dravidian as suggested by Chatterji (1926). Once the future meaning wiped out the modal one, the upgrading of the agent to a nominative subject is expected. Note in this respect that ergative alignment in the future is not common in languages with nominative alignment in the present. Since both past and future constructions pattern alike, both were simultaneously realigned, unlike what we find in West IA. The formerly instrumental subject of obligation sentences shifting to dative may be correlated with the dative subject sentences. The latter construction gained ground in the 15th-17th century with experiential predicates, and was extended to obligation constructions, once the formal link with the old form in -*av*- was lost. The old case pattern prevailed only in Marathi and Gujarati, which kept the -*av*- predicate and an -*av*- infinitive, respectively. Finally, one may wonder why the two forms studied changed in meaning and category so dissimilarly: whereas the PPP became a definite past everywhere, the gerund grammaticalized into future, deontic modality, verbal noun and infinitive. In the case of the former, as a resultative participle it

initially foregrounded the result of an action, while the implicature of the underlying process was later conventionalized. The latter form originally simply conveyed verbal meaning, there is no such unique implicature associated. Thus, the construction shows different developments depending on its function: it is interpreted as an intended action if used predicatively (Eastern IA languages), or as a verbal noun if used non-predicatively. Here again, Marathi is an exception: it is the only regional language not having (had) the non-predicative use, and the only Western IA language having maintained the predicative modal meaning, which calls for an explanation. Emphasizing its conservative character by pointing to the preservation of the old infinitive and of the three-gender system does not make much sense, in view of the neighbouring Gujarati also maintaining three genders. Social factors may provide an explanation, given the historical, self-claimed cultural Marathi specificity. The texts forming the active basis of this cultural heritage should be analysed with this in mind, in the same way as the bulk of literature decisive for the formation of Hindi/Urdu, another exception since they just lost the -*v*- form.

References

Adams, James Noel. 2011. Late Latin. In J. Clackson (ed.), *A companion to the Latin language*, 257–283. Chapter 16. London: Wiley-Blackwell.

Adams, James Noel. 2013. *Social variations and the Latin language*. Cambridge: Cambridge University Press.

Beníšek, Michael. 2010. The quest for a Romani infinitive. *Romani Studies* 5-20(1). 47–86.

Benveniste, Emile. 1952. La construction passive du parfait transitif. In Emile Benveniste (ed.), *Problèmes de linguistique générale*, vol. 1, 176–186. re-edited in 1966. Paris: Gallimard.

Benveniste, Emile. 1960. *Etre* et *avoir* dans leurs fonctions linguistiques. In Emile Benveniste (ed.), *Problèmes de linguistique générale*, vol. 1, 187–207. Paris: Gallimard.

Benveniste, Emile. 1966[1952]. La construction passive du parfait transitif. In Emile Benveniste (ed.), *Problèmes de linguistique générale*, vol. 1, 176–186. Paris: Gallimard.

Benveniste, Emile. 1966[1965]. Les transformations des catégories grammaticales. In Emile Benveniste (ed.), *Problèmes de linguistique générale*, vol. 2, 127–136. Paris: Gallimard.

Bloch, Jules. 1906. *La phrase nominale en sanscrit*. Paris: Champion.

Bloch, Jules. 1970[1920]. *The formation of the Marathi language.* originally published in French *La formation de la langue marathe].* Delhi: Motilal Banarsidass.

Boretzky, Norbert. 1996. The 'new' infinitive in Romani. *Journal of the Gypsy Lore Society* 6(1). 1–51.

Bourova, Viara & Liliane Tasmowski. 2007. La préhistoire des futurs romans — ordre des constituants et sémantique. *Cahiers Chronos* 19. 25–41.

Breunis, André. 1990. *The nominal sentence in Sanskrit and Middle Indo-Aryan.* Leiden: Brill.

Bubenik, V. 1998. *A historical syntax of Late Indo-Aryan (Apabhraṃśa).* Amsterdam: Benjamins.

Bubenik, V. & Ch. Paranjape. 1996. Development of pronominal systems from Apabhramsha to New Indo-Aryan. *Indo-Iranian Journal* 39. 11–32.

Bybee, Joan L. 2003. Mechanisms of change in grammaticization: The role of frequency. In Brian D. Joseph & Richard D. Janda (eds.), *The handbook of historical linguistics*, 602–623. Oxford: Blackwell.

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.

Cardona, G. & B. Suthar. 2003. Gujarati. In G. Cardona & D. Jain (eds.), *The Indo-Aryan languages*, 658–697. London: Routledge.

Cardona, George. 1970. The Indo-Iranian construction *mana* (*mama*) *krtam*. *Language* 46. 1–12.

Chatak, Govind. 1966. *Madhyapahari ka bhashashastrîa adhyayan.* Delhi: Radhakrishna Press.

Chatterji, Sunita Kumar. 1986. *The evolution of Bengali language. Vol. 3: Revision and additions.* Delhi: Rupa.

Chatterji, Suniti Kumar. 1926. *The evolution of Bengali language.* Delhi: Rupa.

Dave, Trimbaklal N. 1935. *A study of the Gujarāti language in the 16th century (v. S.) with special reference to MS Balavabodha to Upadeśmala (Suri).* London: Royal Asiatic Society. http://www.sudoc.abes.fr/DB=2.1/SET=3/TTL=1/CLK?IKT=1016/&TRM.

Davison, Alice. 2002. Agreement features and projections of tense and aspect. In R. Singh (ed.), *Yearbook of South asian languages and linguistics*, 27–57. Delhi: Sage.

Debrunner, Albert. 1954. *Altindische Grammatik.* Vol. II, 2. Göttingen: Vandenhoek und Ruprecht.

Annie Montaut

Desai, Mohan. 1926. *Jain Gurjar Kavyo*. Bombay: Jain Svetambar Conference Office.

Dhongde, Ramesh V. & Kashi Wali. 2009. *Marathi*. London: Benjamins.

Ernout, Ernest & François Thomas. 1951. *Syntaxe Latine*. Paris: Klincksieck.

Goswami, G. C. & J. Tamuli. 2003. Assamese. In G. Cardona & D. Jain (eds.), *The Indo-Aryan languages*, 411–443. London: Routledge.

Grierson, G. A. 1903-1928. *Linguistic survey of India, Delhi (reprint)*. Delhi: Motilal Banarsidass.

Haspelmath, Martin & Andrea D. Sims. 2010. *Understanding morphology. 2nd edition*. London: Hodder Education.

Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

Hock, Hans Heinrich. 1992. *Studies in Sanskrit syntax*. Delhi: Motilal Banarsidass.

Hopper, Paul J. & Elizabeth C. Traugott. 2003. *Grammaticalization*. Cambridge: Cambridge University Press.

Jaiswal, Mahesh Prasad. 1962. *A linguistic study of Bundeli*. Leiden: Brill.

Jha, Subhadra. 1958. *The formation of the Maithili language*. Delhi: Manoharlal.

Jha, Subhadra. 1985[1958]. *The formation of the Maithili language*. Delhi: Manoharlal.

Kachru, Yamuna. 1987. Ergativity, subjecthood and topicality. *Lingua* 71. 223–238.

Kellogg, R. 1972[1875]. *A grammar of the Hindi language*. Delhi: Oriental Book Reprints.

Khokhlova, Liudmila. 2013. Obligational constructions in New Indo-Aryan languages of Western India. *Lingua Posnianiensis* IV(2). 91–107.

Kuryłowicz, Jerzy. 1931[1960]. *Les Temps composés du roman*. re-ed. In Esquisses linguistiques, 1960, 104–108. Kraków.

Kuryłowicz, Jerzy. 1965. The evolution of grammatical categories. *Diogène* 51. 51–71.

Malchukov, A. & A. Spencer. 2009. *The Oxford handbook of case*. Oxford: Oxford University Press.

Mistry, P. J. 1997. Objecthood and specificity in Gujarati. In J. Hill, P. J. Mistry & L. Campbell (eds.), *The life of language*, 425–442. Berlin: Mouton de Gruyter.

Montaut, Annie. 1996. L'évolution des systèmes perfectal et futur en indo-aryen occidental et oriental. *Journal Asiatique* 284.2. 325–360.

Montaut, Annie. 1997. Benveniste et Kuryłowicz: Deux méthodes, deux trouvailles sur le système aspecto-temporel. N° spécial: Emile Benveniste vingt ans après. C. Normand (ed.) *Linx* 9. 337–353. http://linx.revues.org/1080.

Montaut, Annie. 2004. Oblique main arguments in Hindi as localizing predications. In P. Bhaskararao & K. V. Subbarao (eds.), *Non-nominative subjects*, 33–56. Amsterdam: Benjamins.

Montaut, Annie. 2007. The evolution of the tense-aspect system in Hindi/Urdu: The status of the ergative alignment. In M. Butt & T. Holloway King (eds.), *Proceedings of the LFG06 Conference*. Stanford: CSLI. Online Publications.

Montaut, Annie. 2015. The verbal form V-*â* in Hindi/Urdu: An aorist with "aoristic meanings". In Zlatka Guentchéva (ed.), *Aspectuality and temporality. Empiric and theoretical issues.* 413–447. Amsterdam: Benjamins.

Montaut, Annie. 2016. Why the ergative case in modal (in)transitive sentences the historical evolution of aspect and modality in Indo-Aryan. In Eystein Dahl & Krysztof Stronski (eds.), *Indo-Aryan ergativity in typological and diachronic perspective* (TSL 112), 135–167. Amsterdam: Benjamins.

Pandharipande, Rasjeshari. 1997. *A grammar of the Marathi language.* London: Routledge.

Pandit, Prabodh Bechardas. 1976. *A study of the Gujarati language in the 14th century with special reference to a critical edition of Ṣaḍāvaśyakabālāvabodhavṛtti of Taruṇaprabha.* Bombay: Bharatiya Vidya Bhavan. http://www.sudoc.abes.fr/DB=2.1/SET=2/TTL=1/CLK?IKT=1016&TRM=.

Peterson, John. 1998. *Grammatical relations in Pali and the emergence of ergativity in Indo-Aryan.* München: Lincom Europa.

Pirejko, L. A. 1979. On the Genesis of ergative construction in Indo-Iranian. In F. Plank (ed.), *Ergativity: Towards a theory of grammatical relations*, 481–488. London: Academic Press.

Pischel, Richard. 1971[1900]. *A grammar of the Prakrit languages.* [originally publ. in German Grammatik der Prakrit-Sprachen, Strassburg, K. J. Trübner]. Delhi: Motilal Banarsidass.

Renou, Louis. 1952. *Grammaire de la langue védique.* Lyon: IAC.

Saxena, Ram Baburam. 1937. *Evolution of Awadhi. Delhi [1971 reprint].* Delhi: Motilal Banarsidass.

Seržant, Ilja A. 2012. The so-called possessive perfect in North Russian and the Circum-Baltic area. A diachronic and areal approach. *Lingua* 122. 356–385.

Snell, Rupert. 1991. *The Hindi classical tradition: A Braj Bhasha reader.* London: SOAS.

Strońksi, Kryzstof. 2010. Variations of ergativity patterns in Indo-Aryan. *Poznan Studies in Contemporary Linguistics* 46(2). 237–253.

Annie Montaut

Stroński, Kryzstof. 2014. On the syntax and semantics of the past participle and gerundive in early New-Aryan. Evidence from Pahari. *Folia Linguistica Historica* 35. 275–305.

Tagare, G. V. 1948. *Historical grammar of Apabhramsha*. Poona: Deccan College.

Tessitori, Luigi. 1914–1916. Notes on the grammar of the old Western rajasthani, with special reference to apabhramsha and to gujarati and marwari. *Indian Antiquary* 42–44.

Tiwari, U. N. 1966. *The development of Bhojpuri* (The Asiatic Society X). Calcutta: no publisher.

Tiwari, Udayan N. 1961. *Hindi Bhasha ka udgam aur uska vikas*. Prayag: Bharati Bhandar.

Touratier, Christian. 1994. *Syntaxe Latine*. Louvain: Peeters.

Trask, R. L. 1979. On the origins of ergativity. In Franz Plank (ed.), *Ergativity: Towards a theory of grammatical relations*, 385–404. London: Academic Press.

Trumpp, Ernst. 1872. *Grammar of the Sindhi language*. London-Leipzig.

Tunga, Sudamsu Shekhara. 1995. *Bengali and other dialects of South Assam*.

Wali, Kashi. 2004. Non-nominative subjects in Marathi. In P. Bhaskararao & K. V. Subbarao (eds.), *Non-nominative subjects*, 223–252. Amsterdam: Benjamins.

Chapter 7

On the grammaticalization of demonstratives in Hoocąk und other Siouan languages

Johannes Helmbrecht
University of Regensburg

The topic of the present paper is the description of the grammaticalization of a paradigm of adnominal demonstratives in Hoocąk out of a set of three verbs/auxiliaries of existence. This process has never been described for Hoocąk nor does the literature on grammaticalization mention this kind of grammaticalization path. Hoocąk (also known as Winnebago) is a Siouan language still spoken in Wisconsin. Hoocąk has two paradigms of demonstrative pronouns. The first paradigm of demonstratives goes back to Proto-Siouan. The second paradigm is an innovation in Hoocąk. Two bound deictic forms *=re* and *=ga* are systematically combined with the three positional verbs *nąk* 'sit', *ąk* 'lie' and *jee* 'stand' in order to form a new paradigm of demonstratives. These new demonstratives are different from the first and old paradigm grammatically and semantically. First, they are always adnominal demonstratives determining the head noun, while the old paradigm can be used as adnominal as well as pronominal demonstratives. Secondly, they appear only post-nominally, while the first paradigm is more variable occurring pre- and post-nominally. Thirdly, the new demonstratives classify the head noun according to proximate and distal and according to its spatial position (neutral, horizontal, vertical), while the old demonstratives distinguish only proximal and distal and are used to refer anaphorically to aforementioned participants, whole propositions and episodes in a narration. Fourthly, the new paradigm of demonstratives can be used as relativizer and subordinator, which is not possible with the old paradigm of demonstratives. The positional verbs themselves, on the other hand, go back to Proto-Siouan. They are grammaticalized in Hoocąk (and other Siouan languages) as classificatory auxiliaries of being 'be.sitting/ be.lying/ be.standing' and in combination with other verbs as continuative/ progressive markers. The grammaticalization processes that are observed in Hoocąk will then be compared to corresponding ones

Johannes Helmbrecht

in Siouan languages of other sub-branches of Siouan, in particular with the grammaticalizations of classificatory definite articles in Omaha-Ponca (Dhegiha). It will be shown that the positional verbs underwent a different grammaticalization path in this subbranch of Siouan.

1 Introduction

1.1 Grammaticalization of demonstratives

Demonstratives are the starting point for a remarkable variety of different grammaticalization processes leading to quite different grammatical forms.[1] Well-attested in many languages and language families is the development of demonstrative pronouns to third person pronouns and finally to argument indexing pronominal affixes. Another well-attested and often described grammaticalization process is the development of adnominal demonstratives to definite articles, specificity marker, noun class/ gender marker and finally noun marker. A summary of the grammaticalizations of demonstratives brought forward in the literature is given in Table 1.

All targets in Table 1 that are marked **bold** are grammaticalizations that can be found in Siouan languages and will be presented later at different places in the course of this paper. Where do – in turn - demonstratives come from? From which sources do demonstratives grammaticalize? It is obvious that not much is known about this question. A few answers are given in the literature:

i. According to Diessel (1999: 154), there is no evidence that deictic roots, on which the demonstratives are based, are grammaticalized from lexical sources. Instead they belong to the basic vocabulary of every language and often show an iconic relationship between the phonetic shape and the meaning (with regard to distance relationships). Diessel claims (with others) that the exophoric usage of demonstratives is basic.

ii. In Lehmann (1995[1982]: 37-38) it is suggested that deictic roots combine with categorial nouns in order to form new demonstratives (illustrated with examples from Japanese *ko-re* 'this one', *so-re* 'that one', and *a-re* 'yonder one' (Lehmann 1995[1982]: 38). One may also think of complex demonstrative pronouns like the ones found in Korean. Korean has three

---

[1] This paper is dedicated to the memory of Robert L. Rankin who passed away in February 2014. He was the leading scholar in comparative Siouan linguistics and much we know about Siouan languages today is based in one way or other on his research.

Table 1: Demonstratives as sources for various grammaticalizations (summarized from Diessel 1999: Chapter 4, and Heine & Kuteva 2002)

source(s)		target(s)	some references
demonstrative pronouns	>	**3rd person PRO** > clitic PRO > affix	Givón (1984: 353-360); Lehmann (1995[1982]: 39-42); Heine & Kuteva (2002: 112)
	>	relative pronouns	Lehmann (1984)
	>	complementizers	Harris & Campbell (1995: 287); Heine & Kuteva (2002: 106)
	>	subordinators (adverbial clauses)	Heine & Kuteva (2002: 114)
	>	sentence connectives	Diessel (1999: 125); Heine & Kuteva (2002: 108)
adnominal demonstratives	>	**definite articles** > specific/indefinite > noun class/ gender markers	Greenberg (1978); Lehmann (1995[1982]: 38, 55); Heine & Kuteva (2002: 109); and many others
	>	relative pronouns	Lehmann (1984: 378-383); Heine & Kuteva (2002: 113)
	>	linkers	Himmelmann (1997: 172-188)
	>	boundary markers of post-nominal relative clauses/ **relative particles**	Diessel (1999: 132)
	>	determinatives (demonstratives that function as head of a relative clause)	Quirk et al. (1972: 217)
	>	**specific indefinite articles**	Gundel, Hedberg & Zacharski (1993)
adverbial demonstratives	>	**temporal adverbs**	Diessel (1999: 139)
	>	directional preverbs	Lehmann (1995[1982]: 97-104)
identificational demonstratives	>	**non-verbal copulas** > focus markers	Diessel (1999: 147-148); Heine & Kuteva (2002: 108, 111)
	>	explietives	Traugott (1992: 216-219)

deictic particles that are used as determiners (cf. the paradigm in (1). If they are used as demonstrative pronouns, they have to be combined with a defective categorial noun such as *il* 'thing' in example (2).

(1) Korean (Sohn 1999: 295)
i = person/thing near speaker
ku = person/thing near hearer
ce = person/thing away from speaker and hearer

(2) Korean
[***ce   il-ul***]   *nwu-ka   mak-keyss-ni*
[that thing-ACC] who-NOM block-will-Q
'Who would be able to block **that**?' (Sohn 1999: 295)

iii. A third answer can be found in Heine & Kuteva (2002: 172/294). The authors provide a few examples from Hausa, Lingala, and Ngbaka that show that adverbial demonstratives such as 'here' and 'there' may become proximal and distal demonstratives ('this', 'that').

1.2 Goals of the paper

The goal of the present paper is to present an admittedly incomplete overview of the grammaticalizations of demonstratives in Hoocąk and other Siouan languages. A major role in these historical developments play posture verbs designating 'sitting', 'standing' and 'lying'; Siouanists call them "positional verbs" or just "positionals". More specifically, it will be shown:

i. that the Proto-Siouan posture verbs became aspect marking auxiliaries in Hoocąk and many other Siouan languages;

ii. that the aspect marking auxiliaries (continuative aspect) were combined with subordinating deictic particles in Hoocąk that grammaticalized to new adnominal demonstratives;

iii. that these "new" demonstratives preserved a noun classifying and aspect marking function, if used to subordinate clauses;

iv. that the advent of these forms caused a shift in the usage of the old Proto-Siouan demonstratives in Hoocąk; and finally

v. that the same Proto-Siouan positional verbs underwent a different gram-
maticalization path in Mandan;

vi. that the same Proto-Siouan positional verbs underwent a different gram-
maticalization path in Omaha-Ponca and other Dhegiha languages; there
they became classificatory articles probably without an intermediate step
of being demonstratives, and from these classificatory articles new demon-
stratives were developed.

1.3 Hoocąk and the Siouan languages

Hoocąk is a North American Indian language of the Siouan language family still
spoken at various places in Wisconsin. The Siouan language family consists of
about 17 languages that were originally spoken in a large area covering most
of the Great Plains expanding from the Southeast of the US to the Northwest
into Southern Canada. The genetic sub-classification of the Siouan languages is
generally considered as summarized in Figure 1 (cf. Rood 1979; Mithun 1999: 501;
Parks & Rankin 2001).

Although there is some disagreement about details of this reconstruction it
is uncontroversial that Hoocąk (also called Winnebago in the older literature)
and Chiwere (also called Iowa-Otoe-Missouria or Báxoje-Jíwere-Ñút'achi) are
a subgroup of the Central Siouan or Mississippi Valley Siouan languages. All
Siouan languages, except the Dakotan languages and perhaps Crow, are highly
endangered languages and are on the verge of extinction or are already extinct
(indicated by little crosses in Figure 1). It is estimated that there are less than 200
Native speakers of Hoocąk left, who are all older than 60 years of age.

1.4 The data

The data for this study are taken from grammatical descriptions and published
text sources preferably of the Siouan languages that are documented best. Since
the languages of the Southeastern Siouan branch (Ohio Valley Siouan) are ex-
tinct for a long time now the descriptive information is not as detailed as from
Lakhota and other Mississippi Valley Siouan languages, or from the Missouri
River branch of Siouan. The data for Hoocąk come from fieldnotes and texts
that were collected within the DOBES project of the documentation of Hoocąk
(2003-8). The historical-comparative data are mainly taken from the *Comparative*

Johannes Helmbrecht
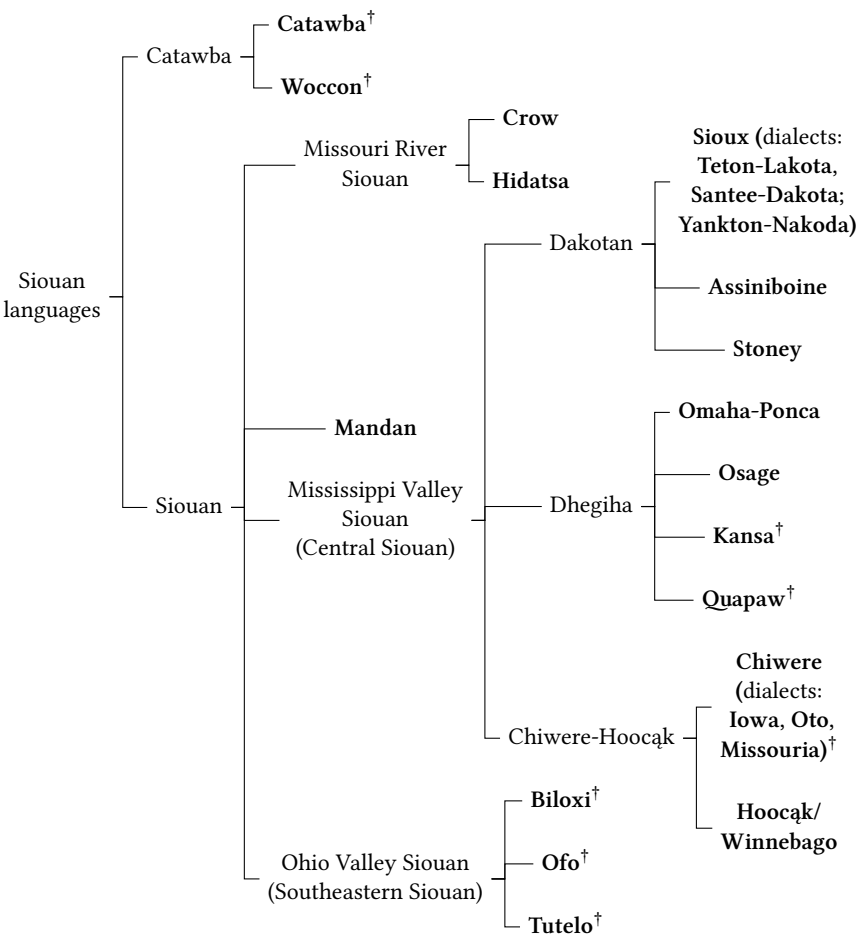
Figure 1: Genetic classification of the Siouan languages (cf. Rood 1979; Mithun 1999: 501; Parks & Rankin 2001)

Siouan Dictionary (Rankin et al. 2015).[2]

2 From positional verbs to aspect markers in Siouan

2.1 Positional verbs of Hoocąk

Hoocąk has a set of three so called positional verbs, which belong to the group of verbs of 'being/ existence'. These positionals designate the bodily posture of a human or animate subject, or the spatial orientation of an inanimate subject. See the forms and their meanings in Table 2.

Table 2: Hoocąk positionals (cf. Lipkind 1945: 45; Helmbrecht & Lehmann 2010: 26)

positional(s)	meaning
=nąk	'be (sitting position/ neutral position)'
=jee/=jąą	'be (standing position/ vertical)'
=(h)ak/=(h)ąk	'be (lying position/ horizontal)'

The positional verbs may be used as full verbs (as illustrated in (3), or as auxiliaries in combination with another full verb that is preceding it; cf the subsequent examples in (4) and (5). The positionals in Hoocąk have to be analyzed as enclitics, if they are used as auxiliaries.

(3) GMA007
    CW: waǧįǧį suucra   nųųpįį 'eegi hanąkwi
        waǧįǧį šuucra   nųųpi(w)i 'eegi hanąkwi
        waǧįǧį šuuc=ra   nųųpiwi 'eegi ha-**nąk**-
        ball   be.red=DEF two   here COLL-**POS.NTL**-PL
    CW: 'there are two red balls sitting here'

---

[2] See the website of the DOBES funding initiative of the Volkswagen Foundation (http://dobes.mpi.nl//) The glossed texts and audio and video files of the Hoocąk documentation project are stored in the digital archive of the Max-Planck-Institute for Psycholinguistics called "The Language Archive"; the corresponding URL is: http://dobes.mpi.nl/projects/hocank/. The website of the DOBES project "Documentation of the Hoocąk Language" led by Johannes Helmbrecht and Christian Lehmann at the University of Erfurt, Germany can be found under the following URL:http://www2.uni-erfurt.de/sprachwissenschaft/Vgl_-SW/Hocank/index_frames.html.

Johannes Helmbrecht

In (3), the positional *nąk* 'be.sitting' is used as sole verb in a predication of existence/ location. In this case, it classifies the referent of the subject NP, 'the two red balls', according to its inherent spatial orientation as sitting. In (4), *=nąk* 'be.sitting' is used as an auxiliary to the main verb *wee* 'talk' marking continuative aspect. In addition, the positional marks the spatial orientation of the actor/subject.

(4) CGF011

    CG: Virgilga   waanąkšąną:
    CG: Virgil-ga   wee=**nąk**=šąną:
    CG: Virgil-ᴘʀᴏᴘ talk(ꜱʙᴊ.3ꜱɢ)=**ꜱʙᴊ.3ꜱɢ.ᴘᴏꜱ.ɴᴛʟ**=ᴅᴇᴄʟ

    "Hakerekjane   'ee   waakšąną"
    ha-kere-kjane   'ee   wee=**ak**=šąną
    1ᴇ.ᴀ-go.back.there-ꜰᴜᴛ 3ᴇᴍᴘʜ talk-**ꜱʙᴊ.3ꜱɢ.ᴘᴏꜱ.ʜᴏʀ**=ᴅᴇᴄʟ

    BO: "Hakerekjane,   connection".
    BO: ha-kere-kjane,   connection
       1ᴇ.ᴀ-go.back.there-ꜰᴜᴛ, connection

    CG: Virgil **was saying**: "**he's saying** I'm going home".
    BO: I'm going home, connection.

There is a second positional in (4), *=ak* 'be.lying', which is an auxiliary to the main verb *wee* 'talk'. The positional in this construction indicates likewise the spatial orientation of the actor/ subject and continuative aspect. However, in this example, the actor is not lying, but horizontally moving. The way home is conceptualized as a long line lying on the surface; movement requires always the 'be.lying' or horizontal positional. The next example in (5) represents an instance of *=jee* 'be.standing'. This positional indicates that the actor/ subject of the 'telling' is in a vertical/ standing position. In addition, *=jee* marks continuative aspect.

(5) MOV024

    heeją́ga hopįnįsge   'eeja
    heejąga ho-pįį=nįįsge   'eeja
    now   ᴀᴘᴘʟ.ɪɴᴇꜱꜱ-be.good=ᴠᴀɢᴜᴇ there

    hagiregają,   Hank Tga
    ha-gii-ire=gają   Hank ᴛ-ga
    ᴄᴏʟʟ-arrive.back.there-ꜱʙᴊ.3ᴘʟ-ꜱᴇǫ Hank ᴛ-ᴘʀᴏᴘ

wokarakjeeną,	*hegų.*
wa-ho<ka->rak=**jee**=ną,	hegų
OBJ.3PL-<POSS.RFL->tell=**POS.VERT**=DECL	that.way

'and when they got back to good ground, Hank T **was telling** about himself:'

It has to be noted that the positional auxiliaries are not the only verbs of 'being'. There are four others (see Table 3) and one of these indicates continuative aspect in Hoocąk likewise.

Table 3: Other verbs of being/ existence in Hoocąk

verb of being	meaning	comment
nįhé	'to be/CONT'	this verb of being can be used to mark continuative aspect; cf. example (11)
heré	'to be'	copula/ auxiliary, never used to indicate aspect
'ųų	'to be/do'	never used as auxiliary with a full verb in order to mark continuative aspect, sometimes it marks rather a slight causative meaning, frequently combined with one of the positionals
wa'ų	'to be/ do'	similar as *'ųų*

2.2 Positionals of other Siouan languages

The positionals in Hoocąk in the previous section are Common Siouan. In all Siouan languages, at least traces of the positional verbs can be found. Compare the cognate forms as reconstructed by Rankin et al. (2015) in the *Comparative Siouan Dictionary* (cf. also Rankin 2004a) in Table 4.

In all Siouan languages - except the languages of the Dhegiha branch - the positionals are used as auxiliaries and often as markers that indicate continuative aspect. The Dhegiha languages lost the Proto-Siouan positional verbs. The positionals developed into classificatory definite articles in these languages, see §6.1 below.

Johannes Helmbrecht

Table 4: Positionals in Siouan (cf. Rankin 2004a, Rankin et al. 2015)

	SIT	LIE	STAND	STAND
Proto-Siouan	*rą́•-kE*	*wų́•kE*	*rahÉ*	*hą́(-kE)*
Crow	*da•čí*	*baačí*	-	*áahku*
Hidatsa	*rá•kE*	*wá•kE*	*rahÉ*	*háhku*
Mandan	*rąk*	*wąk*	*te*	*hąk*
Lakota	*yąká*	*yųká*	*he*	*hą*
Chiwere	*ną́ŋe*	*hą́ŋe*	*ǰe*	-
Hoocąk	*=ną́k*	*=(h)ąk*	*=jee*	*=jąą*
Biloxi	*ną́ki*	*mą́ki*	*ne*	*há́de*
Tutelo	*nąka*	*-mąki-*	*ne*	*-hą́k*

2.2.1 Crow positionals

In Crow, there is a set of six auxiliary verbs/ markers that indicate continuative aspect. Three of them are descendants of the Proto-Siouan forms marked in bold face in Table 5.

Table 5: Continuative markers in Crow (cf. Graczyk 2007: 305-309)

form	meaning	degree of coalescence with main verb	Proto-Siouan
datchí	'continue (by mouth)'	[V+AUX] one word	SIT
dawí	'continue in motion; begin to'	[V+AUX] one word	
(d)ahkú	'continue in activity, remain, dwell'	also independent verb	STAND
dachí	'remain voluntarily'	also independent verb	
baachí	'lie, remain involuntarily'	always independent	LIE
ilúu	'do repeatedly, continue'	always independent	

These auxiliaries are inflected for person/ number of the actor/ subject. The actor/ subject is obligatorily co-referential with the actor/ subject of the main verb.

144

If the main verb preceding the auxiliary does not form a single word with it, it obligatorily has a same subject marker with it (-*ak* SS). Crow as well as Hidatsa have developed a switch reference marking system. Otherwise, there is an additional continuative marker (-*a* CONT) between main verb and auxiliary. These auxiliaries behave differently with respect to the closeness of the coalescence with the main verb.

2.2.2 Mandan positionals

Mandan has four positional verbs with a stative meaning; all of them are descendants of the Proto-Siouan positionals. These positional verbs are used as full verbs designating the existence or being of an entity at some place. For instance, the bound form *te-* 'stand' is used to indicate the position of a village in the text in Mixco (not reproduced here; see Mixco 1997: 66, sentence 1). Three of the four positionals are used – in addition – as auxiliaries to indicate continuative aspect. They have a continuative marker -*æ* (CONT) with them and are translated by Mixco as 'abide:sitting/ standing/ lying'; cf. Table 6.

Table 6: Mandan positionals (Mixco 1997: 48f).
Mixco's (1997) analysis deviates somewhat from Kennard's (1936). Kennard takes the three positionals –*nąk* 'be.sitting', -*hąk* 'be.standing' and –*mąk* 'be.lying' as auxiliaries that indicate CONTinuative aspect if they are preceded by the continuative marker *ha-*. This marker is not mentioned in Mixco. Instead, Mixco postulates that the element -*æ* marks continuative aspect. Note also that [n] and [m] are taken as allophones of /r/ and /w/ before nasal vowels in Mandan by Mixco. There are no nasal consonants in the phoneme inventoory of Mandan.

stative verbs		continuative auxiliaries		Proto-Siouan
rąk	'sit'	*rąk-æ*	'abide:sitting'	SIT
hąk	'stand'	*hąk-æ*	'abide:standing'	STAND
wąk	'lie'	*wąk-æ*	'abide:lying'	LIE
te-	'stand'			STAND
		rųrįh	'exist.PL'	

Interestingly, the continuative LIE *wąk-æ* is by far the most frequently used aspect marker in the Mandan text that I examined. The posture meaning is neutralized in most of these usages. In addition, if the subject is plural, only the LIE continuative can be used bearing the regular plural marker; cf. the example in (6).

Johannes Helmbrecht

(6) Mandan

"ko- hú:- æ ki-rút-rį, **wą́:k- æ- kræ- oʔš!**"

"3sg-mother-sv MV-eat-ss **abide:lie-cont-pl-ind.male**"

é= he- ro:wąk- oʔš.

PV=say-narr.past-ind.male

'"They're eating their mother up!" he said.' ([Mixco 1997](): 69)

The combination of full verb plus auxiliary indicating continuative aspect is not as close as the one in Crow; the SS marker is not obligatory, often one finds a simultaneous (SIM) ending on the preceding verb. In addition to the aspect marking function of the positionals, Mandan has developed classificatory demonstratives on the basis of these positionals. They are not as firmly grammaticlized as in Hoocąk and differ from the Hoocąk ones in that the positional follows the deictic particle (cf. [Mixco 1997](): 42). I will discuss this construction below in §5.

2.2.3 **Teton-Lakota**, **Santee-Dakota**, **Yankton-Nakota positionals**

In Lakota, i.e. the Teton dialect of the Sioux language, there are likewise at least three verbs of 'being/existence' that are descendants of the Proto-Siouan positionals (cf. [Boas & Deloria 1941](): 126f); cf. Table 7.

Table 7: Positionals in Lakota (cf. [Boas & Deloria 1941](): 126)

form	meaning	Proto-Siouan
yąká	'to sit', 'be.sitting' (spherical objects, animals etc.)	SIT
=hą	'to stand', 'be standing' (long upright objects)	STAND
yųká	'to lie', 'be.lying' (mostly animate beings)	LIE

All three forms can be used as independent verbs of posture and of 'being'/ 'existence' in all three dialects of Siouan proper (Santee-Dakota, Teton-Lakota, and Yankton-Nakota). See an example from Lakota in (7). Further examples can be found in [Rankin (2004a)]() and [Barron & Serzisko (1982)]().

(7) Lakota

… *k'eyaṡ t'imá* **yąká-pi** *ki ʾátayaṡ wąwíc'ayakapiṡni ną* …
but in.the.tent **sit-they** the entirely they.did.not.see.them and

'but sitting in the tent they (the twins) did not see them, and…' (Deloria 1932: 193ff; Boas & Deloria 1941: 170)

However, in Lakhota, *=hą* has become a fully grammaticalized enclitic that marks continuative aspect (cf. Boas & Deloria 1941: 60f; also Ingham 2003: 31). As such it can no longer be inflected for person/ number of the subject/ actor; it can even be combined with one of the other positionals; see the example from the same text in (8). In this usage, *=hą* has lost completely its posture meaning.

(8) Lakota

*t'é**hą** yéṡ tąyą waáp'e **yąká-hą-pi** k'ų…*
a.long.time but well wait **sit-CONT-they** the.past

'but a long time they were waiting….' (Deloria 1932: 193ff; Boas & Deloria 1941: 170)

Note also, that *=hą* has become the basis of a variety of derivations such as time adverbials; compare *t'é**hą*** 'a long time' in (8); cf. Boas & Deloria (1941: 60f). The two other Siouan proper dialects, Santee-Dakota and Yankton-Nakota (cf. Figure 1) specialized the positional *yąká* 'be.sitting' as the neutral or general auxiliary in order to mark continuative aspect in case that the posture of the actor/ subject is not in focus or is unimportant (see Deloria 1932: 165). The same holds for Hoocąk: the 'be.sitting' positional *=nąk* is also used as the neutral unmarked continuative marker. This unmarked/ neutral usage of the 'be.sitting' positional *=nąk* is nicely reflected in the textual frequencies of this auxiliary in the entire DOBES corpus of Hoocąk; cf. Table 8.

Table 8: Absolute frequencies of the positionals in the DOBES corpus

form	gloss	frequency
=nąk	POS.NTL	n= 1286
=jee/=jąą	POS.VERT	n = 522
=(h)ąk/=(h)ak	POS.HOR	n = 167

The 'be.sitting' positional *=nąk* occurs as an auxiliary/verb twice as often in the corpus as the two others together.

Johannes Helmbrecht

2.2.4 Biloxi positionals

Biloxi is a Siouan language of the Southern branch. The cognate Biloxi positionals *nąki* 'sit', *mąki* 'lie', and *ne* 'stand' (see Table 4) are all used as classifiers in copula clauses that localize a non-human subject, and in verbal clauses with a complex predicate to mark continuative aspect; cf. (9).

(9) Biloxi
     Ayáⁿ xotká u-xě' **náñ̨ki**, xyihě' **náñ̨ki** [On'ṭi-yándi]
     tree hollow in-sit **be.sitting** growl **be.sitting** Bear-sbj
     'Bear was then in a hollow tree where he was growling.' (Dorsey & Swanton 1912: 16; sentence 10)

There is one peculiarity in Biloxi that other Siouan languages lack. Biloxi developed a gender classification of the possessum in possessive predications with positionals. The positional *nąki* 'sit' is used as a copula in possessive clauses that express possession of a female kin. The positional *mąki* 'lie' is used in turn to indicate that the possessum is a male kin; cf. Table 9 and an illustrative example in (10). There are numerous examples in Dorsey & Swanton (1912: 130; cf. also Kaufmann 2011) that illustrate this sex classification.

(10) Biloxi
      Ay-óⁿni   é   nañkí
      Your-mother   he/she sit(female.possessum)
      'You have a mother.' (Dorsey & Swanton 1912: 130)

Table 9: Biloxi positionals as copula in possessive clauses (cf. Kaufmann 2011)

positional	possessum	meaning	Proto-Siouan
nąki	female kin	'be.sitting'	SIT
mąki	male kin	'be.lying'	LIE
(h)ąde	singular; no classification	'to be', 'be.moving'	STAND
yukê	plural of (h)ąde; no classification	'to be', 'be.moving.PL'	

Interestingly, one of the Proto-Siouan positionals in Biloxi, *ne* STAND (cf. Table 4) seems to have developed into a demonstrative, cf. the enty in Kaufmann's

Biloxi dictionary (2011: 100); there are also examples in Dorsey & Swanton (1912: 117-167), where *ne* is used alternatively as a definite article and a demonstrative.

3 From positional auxiliaries to classificatory demonstratives

Synchronically, Hoocąk has two paradigms of demonstratives. The first paradigm is called here the "old" paradigm, since these forms can be traced back to Proto-Siouan as will be shown later. The second paradigm is called here the "new" paradigm, since it is a recent innovation in Hoocąk. The forms are composed of the positional auxiliaries plus a deictic particle distinguishing proximal (*-re*) and distal (*-ga*); (on the grammaticalization of these particles, see §4 below).

Both paradigms are frequently used in Hoocąk texts. In the subsequent sections (§3.1-§3.2) I will present a brief overview of the semantic, pragmatic and distributional properties of the forms of both paradigms. In §4 I will present some suggestions on the grammaticalization of the new paradigm and the effects on the usages of the "old" paradigm.

3.1 The "new" paradigm of adnominal demonstratives in Hoocąk

3.1.1 Morphosyntactic and semantic properties

The "new" demonstratives in Hoocąk are used exclusively as adnominal demonstratives. They always follow the head noun and occur in the same structural position as other determiners such as the definite and indefinite articles at the right edge of the NP; cf. the structural template of the lexical NP in Table 11.

The postnominal demonstratives classify the head noun according to the postural position of its referent and according to its distance from the reference point (proximal vs. distal). If the postural position of the referent is non-salient, the neutral demonstrative is chosen; cf. example 11. The postural position of the 'coal' in this utterance is not salient, hence the be.sitting/ neutral demonstrative has been chosen.

Johannes Helmbrecht

Table 10: Two paradigms of demonstratives in Hoocąk

form	meanings
"old" paradigm	
tée/ te'é	'this', 'here', 'now'
mée/ me'é	'this'
žée/ že'é	'that', 'there'
? ga'á	'that'
"new" paradigm	
=nąka (<=nąk-ga)	'that (sitting/ neutral position; distal)'
=nąąka (<=nąąk-ga)	'those (sitting/ neutral position; plural; distal)'
=nąągre (<=nąąk-re)	'these (sitting/ neutral position; proximal; plural)'
=nągre (<=nąk-re)	'this (sitting/ neutral position; proximal)'
=jeega (<=jee-ga)	'that (standing/ vertical position; distal)'
=jąąne (<=jąą-re) =jaane (<=jee-re)	'this (standing/ vertical position; proximal)'
=ąka (<=ąk-ga)	'that (lying/ horizontal position; distal)'
=agre =ągre (<=ąk-re)	'this (lying/ horizontal position; proximal)'

150

Table 11: Structure of the NP in Hoocąk. Elements in parentheses are optional.

(N)	(lexical modifier)	determiner	(quantifier)
- noun	- adjectival concepts	- definite article *=ra*; - indefinite article =hižą; - **"new" adnominal demonstratives** - Ø	- numerals - etc.

(11) BOF008

Hegų	*'ųų*	*hanįhaire,*		*hagoreižą*	*'ųųxįnį*
hegų	'ųų	ha-**nįhe**-ire		hagoreižą	'**ųųxįnį**
that.way	do/make	COLL-**be/PROG**-SBJ.3PL	sometime	**charcoal**

seepnąka	*tuussąną.*
seep=nąka	tuus=šąną
be.black=POS.NTL:DIST	take\1E.A=DECL

'They kept on going that way, **that coal** at some point I took it.'

In general, the postural classification of the referents is semantically motivated. Larger animals, for instance, usually are standing, hence the 'standing/ vertical' demonstrative is chosen in the utterance in (12).

(12) HOR064

Šųųkįgjeega		*šųųkxetera*	*haracap nąą'į*		*hegų.*
šųųk-įk=jeega		šųųkxete=ra	haracap	nąą'į	hegų
dog-DIM=POS.VERT:DIST		horse=DEF	taste	try(SBJ.3SG)	that.way

'That dog tried to bite the horse.'

The adnominal demonstratives are - like the definite article (=*ra*) - used to nominalize a clause. This is a general strategy in Hoocąk to indicate subordination. Relative clauses, for instance, usually require a nominalizing determiner such as the definite article or one of the "new" adnominal demonstratives; cf. an elicited example in (13). The new demonstrative classifies the head noun according to the posture; in addition, it still preserves a progressive meaning for the relative clause.

(13) (Phil Mike; elicited example)

Johannes Helmbrecht

> *wanį́ną   tuujágre*
> [wanį́=ra   [tuuc haa=**ágre**]ᵣₑₗₐₜᵢᵥₑ ₒₗₐᵤₛₑ]
> meat=DEF cooked
>
> '**this** meat **(lying/horizontal)** I am cooking now'

Example (14) illustrates that the "new" adnominal demonstratives are used as subordinators in general.

(14) BOF023

> *ʻEejaxjį   hegų   hąąp hitanįhąįja   hegųgają   hegų*
> ʻeejaxjį   hegų   hąąp hi-taanį-hą=hija   hegų=gają   hegų
> about.there that.way day   ORD-three-times-there that.way=SEQ that.way
>
> *žige   hįšjųwąk,   hegų   hegų   ʻeeja*
> žige   hį-šjųwą='ąk   hegų   hegų   **'eeja**
> again 1E.U-get.sleepy=POS.HOR that.way that.way **there**
>
> *hamįknąka,   žige   hanąąňegi,   ʻeeja wažą*
> **ha-mįįk=nąka**   žige   ha-nąą=regi   ʻeeja wažą
> **1E.A-lie.down=POS.NTL:DIST** again 1E.A-sleep=SIM/LOC there something
>
> *yaahąte.*
> hi<ha>hąte
> <1E.A>dream.of
>
> 'About on the third day I got sleepy again, **lying there** I went to sleep again, I dreamed again.'

From a semantic point of view it is interesting to see that the speaker chose the demonstrative of the neutral position, and not the one of the lying position, which one would expected.

3.1.2 Pragmatics of the adnominal demonstratives

NPs with one of the adnominal demonstratives mostly appear in texts, if the referent had already been introduced at some distance in the previous text; the demonstratives are used to refer back to an old or fainted topic. The following example illustrates this nicely.

(15) BOF035

> *Hiraijixjįgają   hegų   caaxšepjaane   žige   hižą*
> hira<gi>ji-xjį=gają   hegų   **caaxšep=jaane**   žige   hižą
> <APPL.BEN>reach-INTS=SEQ that.way **eagle=POS.VERT:PROX** again one

haǧepšąną.
haǧep=šąną
appear=DECL
'He was getting close and then **this eagle** appeared again.'

The eagle had been introduced a few clauses before the one in (15), and is then reintroduced by means of a NP with a proximal demonstrative. Since the eagle appeared up in the sky, its position is conceptualized here as vertical. There are also textual examples that illustrate that the proximal adnominal demonstrative can be used as a specific indefinite article like Colloquial English *this*. Compare the following utterance in (16). The "man" in this story is mentioned the first time; he is specific, but indefinite.

(16) TWI003
    Ciinąk kąnąkiregi   *ʻeeja ciiregi*   *hagoréižą*
    ciinąk kąnąk-ire=gi   ʻeeja cii-ire=gi   hagoreižą
    village place(OBJ.3SG)-SBJ.3PL=TOP there live-SBJ.3PL=TOP sometime
    hagoréižą wąąkjaané   *hinųkra*   *hakaráikižu*
    hagoreižą **wąąk=jaane**   hinųk=ra   ha<kara-kii>kižu
    sometime **man=POS.VERT:PROX** woman=DEF <POSS.RFL-RCP>be.together
    roogų́įňe.
    roogų-ire
    want-SBJ.3PL
    'Where they lived, **a man** and his wife wanted (something).' (lit. They placed a village, there they lived, once upon a time **this man** together with his wife, wanted something')

3.2 The "old" paradigm of Hoocąk demonstratives

3.2.1 Common Siouan origins

The paradigm of "old" demonstratives can be shown to be of Common Siouan origin; cf. the cognate forms in Table 12.

Johannes Helmbrecht

Table 12: Cognate sets of the Common Siouan demonstratives (Rankin et al. 2015)

	THIS (proximal)	THIS (proximal)	THAT (medial)	THAT (distal)
Proto-Siouan[3]	*Ree-[4]	*re-	*šee	*kaa
Crow				kaka
Hidatsa			še-'e	kaa
Mandan		re		ká-
Lakota	le-		še-	ka-
Chiwere	ǰe-		šé-'e	gá/ gá'e/ ká
Hoocąk	te-'e, tée		že-'é/ žée	=ga/ ʔga'a
Omaha-Ponca[5]		ðe	še	ka
Kansa[6]		ye, yé-che, yé-khe	še	ga
Osage[7]		ðe, ðee	še/ šee	ka/ kaa
Quapaw[8]		de	še	ká-khe
Biloxi[9]	ne-tka	de		ká-wa
Tutelo[10]	née	lèe		ka/ ko

The forms for Proto-Siouan that were reconstructed distinguish three grades of deictic distances.

- proximal/ close to speaker,

- medial/ close to hearer, and

- distal/ away from both speaker and hearer.

That the authors of the CSD reconstructed two different proximal demonstratives (*Ree-* and *re-*) is motivated by independent reasons (cf. Rankin, Carter & Jones 1998).

- The languages of Mississippi Valley Siouan all preserved the whole set of demonstratives; see the shaded lines in Table 12.

- Reflexes of the Proto-Siouan demonstratives are lacking in the Northwestern Siouan languages (Crow, Hidatsa and Mandan).

- The forms of Biloxi and Tutelo (both Ohio Valley) are less certain; these languages are not well documented.

- Interestingly, reflexes of the distal form can be found in all Siouan languages.

The Hoocąk forms that are of Common Siouan origin are given in Table 13 together with their function and meaning in contemporary Hoocąk.

The proximal and medial forms (*te'e* 'this' and *že'e* 'that') are obviously a composition of the Proto-Siouan deictic stem (*te-* and *že-*) plus a demonstrative pronoun *'ee* 'that (aforementioned)', which is likewise attested in all Siouan

---

[3] Cf. the *Comparative Siouan Dictionary* (Rankin et al. 2015).

[4] The capital *R* in **Ree* symbolizes a hypothetical cluster of a resonant /*r/ plus a laryngeal; cf. Rankin, Carter & Jones (1998). According to the autors of the CSD, there are independent reasons to postulate two different /r/ sounds.

[5] Cf. Boas (1907: 324-326); Koontz (1984: 138-142)

[6] Cf. Cumberland & Rankin (2012: 350f); the proximal form is only attested as an adverbial demonstrative, otherwise only in combination with one of the classifying definite articles.

[7] Cf. Quintero (2004: 359-368)

[8] Cf. Rankin (2002: ms); the medial and distal forms are attested only in combination with one of the classfying and definite articles in Quapaw.

[9] Cf. See Einaudi (1976: 69) for the proximate and medial form; see Kaufmann (2011: 77) for the distal form.

[10] Cf. Oliverio (1996: 155)

Johannes Helmbrecht

Table 13: Hoocąk demonstratives: the "old" paradigm and its grammat-icalizations

Hoocąk forms		meaning	Proto-Siouan	meaning
te'e/tee	< *te-* + *'ee*	'this', 'here', 'now'	**Ree-*	'this' (proximal)
me'e/ mee	< *me-+'ee*	'this', 'here'	?	
že'e/ *žee*	< *že-* + *'ee*	'that', 'there', 'then'	**šee*	'that' (medial)
=že/=še		QUOT		
ˀga'a	< *ga-* + *'ee*	'that'		
=ga		'that'; (bound enclitic form)		
=ga		proper name marker (anthroponyms, kinship terms as proper names);	**kaa*	'that' (distal)
=ga		sentence connector continuative ('and then');		
=ga	POS+ga	distal classfying adnominal demonstrative		
'ee		'thus', 'it', 'this', 'that', 'he', 'she' (aforementioned); always EMPH or in focus constructions (as a 3rd personal pronoun, free form);	**'ee*	'that' (aforementioned)

156

languages. This form is variably analyzed as a free pronoun or demonstrative pronoun refering back to somthing already mentioned (aforementioned) in discourse. In Hoocąk, it is not only used as an anaphoric pronoun, but also in focus constructions in order to express emphasis on a third person participant.

Semantically, both the proximal and medial demonstratives seem to have neutralized the deictic distance distinction almost completely; only in the adverbial uses the distinction between "close to speaker" and "far from speaker" is preserved.

The distal form *ga'a* 'that' is mentioned in older sources on Hoocąk (cf. Lipkind 1945: 52); however, there is not a single instance of this form in our DOBES corpus (which contains contemporary but also older texts from the beginning of the 20th century); the composition of this form is analog to the one of the proximal and medial forms attaching the anaphoric pronoun *'ee* 'that, etc.' to the distal demonstrative stem *ga-*. The vowel in turn is assimilated to the stem vowel (compare also the closely related Chiwere form *ga'e* 'that', where the vowel did not undergo this assimilation).

However, the distal deictic stem *=ga* developed different functions in Hoocąk: first, this demonstrative became an enclitic proper name marker that is used obligatorily with anthroponyms and with kinship terms, if they are used in third person reference function. Secondly, this distal demonstrative became a clause or sentence connecting element expressing temporal continuation. And thirdly, this form is used as a distal demonstrative. The latter is certainly not the major function of this demonstrative, there are only a handful instances of this usage in our DOBES corpus. However, the distal deictic stem *=ga* plays an important role in the formation of the new adnominal demonstratives with the positional auxiliaries; see below §4.

There is a second proximal demonstrative *me'e* 'this', which is mentioned in older sources (Lipkind 1945: 52) and occurs occasionally in our corpus. The origins of this /m/ initial form are unclear. This form cannot be traced back to one of the two Proto-Siouan proximal demonstratives on the basis of the known sound laws.

3.2.2 Morphosyntactic and semantic properties of the Hoocąk forms

The "old" demonstratives are used predominantly as demonstrative pronouns or as adverbial demonstratives ('here', 'there', 'now', 'then') in our texts. Sometimes they are also used as adnominal demonstratives, but these occurences are not frequent. If they function as adnominal demonstratives, they occur always postnominally or more specifically, at the right edge of the NP. This is probably

the Common Siouan order. The descendents of the Proto-Siouan forms in the other Siouan languages are all postnominally. However, the word order rules with respect to the "old" demonstratives have become less strict. Although these instances are rare in our corpus, the forms can also occur prenominally. If the demonstratives are used pronominally, they are used almost always endophorically, i.e they refer back anaphorically to a previously mentioned discourse participant, or they refer back to a whole proposition or episode of a narration; this is called here discource deictic reference. Note that this kind of discourse deictic reference is possible also with the demonstrative pronoun *'ee* 'he/she/it/this/that etc.' The "old" demonstratives may also be used as identificational demonstratives in non-verbal or copula clauses. Probably the most frequent use of these demonstratives in our text corpus is their use as adverbial demonstratives that are better translated as 'here' and 'there' or 'now' and 'then'. In these uses, the adverbial demonstrative refers to a previously mentioned situation or exophorically to the actual speech situation (proximal).

To summarize the findings:

- The major result of the grammaticalization of the "new" demonstratives is that there appeared a new paradigmatic distinction between **demonstrative pronouns** and **adnominal demonstratives**. The "old" demonstratives lost their usage as adnominal determiners (not entirely, though). This function has been taken over by the "new" paradigm of demonstratives.

- On the other hand, the "old" demonstratives are dominantly used pronominally in a variety of constructions and adverbially.

- In addition, it can be observed that the Proto-Siouan threefold proximal/ medial/ distal distinction has been bleached or even neutralized in the "old" Hoocąk demonstratives. There is no longer semantically a medial demonstrative, and the distal form *ga'a* 'that (distal)' has been lost entirely in this paradigm.

4 The grammaticalization of the Hoocąk adnominal demonstratives

As has been shown in Table 10, the adnominal demonstratives are historically a combination of the positional auxiliaries plus two deictic particles; *=re* for proximal and *=ga* for distal deixis. The different functions/ disributions of both deictic particles are summarized in Table 14.

Besides the occurence of these particles in combination with the positional auxiliaries, they are still used independently; *=re* (DEM.PROX) is quite frequent in our texts corpus, *=ga* (DEM.DIST) rather rare. If they are used independently, they usually nominalize a clause in order to indicate subordination. Recall that nominalization is a major strategy to form subordinate clauses in Hoocąk; cf. (17) for an illustrative example.

(17) MAP013

jaagu waac 'eeja hamįnągre   paaxų   nąga hegu   'eeja
jaagu **waac 'eeja ha-mįįnąk=re**   paaxų   nąga hegu   'eeja
what **boat there 1E.A-sit=DEM.PROX** pour\1E.A and   that.way there

waac 'eeja nąąjįp   nąga nįį   'eeja
waac 'eeja nąą<ha>jįp   nąga nįį   'eeja
boat there <1EA>tilt.with.the.foot and   water there

waakįnįpšąną
ho-ha-kįnįp=šąną
APPL.INESS-1E.A-fall.down=DECL

'whatever, I sat in the boat, I poured it out, and there I tipped over it (the boat), I fell in the water.' (lit. 'whatever, sitting in the boat there, I poured it out, and ....')

Table 14: Grammaticalization of the deictic particles/bound forms *=re* and *=ga*

form	target(s)
=re (proximal)[11]	proximal adnominal demonstrative (positional auxiliary + =re)
	nominalizer/ subordinator
	imperative marker (IMP)
	derivational means for time adverbials
=ga (distal)	distal adnominal demonstrative (positional auxiliary + =ga)
	nominalizer/ subordinator
	sentence connector (continuative)
	proper name marker

---

[11] The historical source of *=re* remains speculative. Perhaps it goes back to Proto-Siouan **ree*. I am grateful to Rory Larson, who indicated to me this possibility. As far as we know, such a historical developement would not violate known Siouan sound laws.

Johannes Helmbrecht

The clitic deictic particle *=re* (DEM.PROX) indicates subordination of the entire clause (in bold face), which otherwise could not be distinguished from a main clause with regard to its grammatical marking. Other determiners such as the definite article and the "new" adnominal demonstratives occur in the same structural slot with the same function, namely indicating subordination. It seems quite likely to me that the grammaticalization of the "new" demonstratives was mediated by the subordinating function of these particles. The deictic particles as subordinators always appear at the end of the subordinate clause, and if this subordinate clause contains a continuative marking positional, this positional always appears immediatley before the deicitic particle. At one point in the history of Hoocąk, the positional auxiliary lost its person/ number inflection for the subject/ actor of the subordinate clause and got fused with the nominalizing deicitc particle. Finally, this fused form extended its distribution and was generalized as a demonstrative determiner that could also occur with plain nouns in a NP. The grammaticalization of the "new" adnominal demonstratives, therefore, may have come about in three principal steps, cf. Table 15.

Table 15: Grammaticalization of classifying adnominal demonstratives in Hoocąk

step 1	relative clauses or subordinated clauses with a positional auxiliary (continuative aspect marking) are nominalized by *=re*/ *=ga*
step 2	the positional auxiliary + *=re*/ *=ga* are reanalyzed as a subordinating demonstrative
step 3	extension of the range of usages of the subordinating demonstratives, for instance as a determiner with a plain noun in a NP

A construction that represents the developement from step 1 to step 2 in Table 15 could be the following text example:

(18) MOV041

nįge	*paašihajawiga*			*'eeja (hąho) žegu*		
nįge	paašihajawiga			'eeja	hąho	žegų
nįge	**paaši=ha-jee-wi=ga**			'eeja	hąho	žeegų
where	**dance\1E.A=COLL-POS.VERT-PL=DEM.DIST**			there	INTJ	thus

160

howé	*hiperes*	*kįįjee(n)*
howé	hiperes	kįįjeeną
howe	hiperes	kįį=jee=ną
go.about	know(SBJ.3SG)	make.self-POS.VERT=DECL

'The place, **where we were dancing**, there he knows his way around.'

The positional auxiliary *=jee* 'be.standing' that marks continuative in the subordinate clause (given in bold face) is still inflected for the person/ number of the subject/ actor of the subordinate clause. The distal *=ga* is a nominalizing form marking subordination; once this inflection disappears, *=jee* and *=ga* are ready to be reanalysed as a single form. As was shown above, the "new" demonstrative retains the continuative aspect marking function (the auxiliary had) and in relative clauses the classificatory function. It makes also sense to interpret this construction as the starting point for the grammaticalization of *=ga* as a sentence connector indicating continuation.

5 Grammaticalization of classificatory demonstratives in Mandan

A different way to create classificatory demonstratives can be found in Mandan. As already mentioned, Mandan has three positional auxiliaries (cf. Table 6 above) that are used as full verbs in existential and locative clauses, and that are used as auxiliaries expressing continuative aspect when accompanied by a continuative marker. In addition these positionals combine with two demonstrative pronouns - *dɛ* 'this' and *ǫt* 'that' - in order to form classificatory demonstrative pronouns; cf. the forms in (19).

(19) Mandan (Kennard 1936: 28f)
dɛ-nąk
this-sitting
'this one (be.sitting)'
dɛ-hąk
this-standing
'this one (be.standing)'
dɛ-mąk
this-lying
'this one (be.lying)'
ǫt-nąk
that-sitting
'that one (be.sitting)'

Johannes Helmbrecht

> *q*t-hąk
> 'that one (be.standing)'
> *qt*-mąk
> 'that one (be.lying)'

Note that the proximal *dɛ* 'this' in Kennard is represented as *re* 'this' in more recent studies (cf. Mixco 1997: 42). Of the two demonstratives *dɛ* 'this'/ *qt* 'that', only *dɛ* 'this' can be traced back to Proto-Siouan. A similar combination of "old" demonstratives with the positionals as in Mandan does not exist in Hoocąk (I found only one example of this composition in the entire Hoocąk corpus). What is also interesting is that the order of forms in Mandan is different. The demonstrative form precedes the positional auxiliary. It is particular this property that suggests that a different scenario has to be assumed with respect to the grammaticalization of the classificatory demonstratives (with regard to posture) in Mandan. This question needs more research.

Another interesting difference between Hoocąk and Mandan is that Mandan, in addition, grammaticalized the positional auxiliaries to adnominal classificatory demonstratives without any combination with deictic particles. It is the plain forms of the positional auxiliaries that are used as demonstratives in the examples in (20) from Kennard.

(20)   Mandan (Kennard 1936: 28f)
        óti-hąk
        lodge-this.standing
        'this lodge'
        máta-mąk
        river-this.lying
        'this river'
        hárɛ-nąk
        cloud-this.sitting
        'this cloud'

It is difficult to think of a gramaticalization process that reanalyzes 'be' auxiliaries to proximal demonstratives without any support from deictic particles, and to the best of my knowledge such a process never has been described in the literature. This process is attested, however, only for the proximal deixis. For the distal deixis, the *qt* 'that' demonstrative has to be used. The positional auxiliaries do not occur in this function.

6 Omaha-Ponca (Dhegiha) made it differently

The grammaticalization of the positional verbs/ auxliaries in the Dhegiha sub-group of Siouan is remarkably different from that of the other Siouan languages and has been extensively investigated by several authors: cf. Ranklin 1977; Barron & Serzisko 1982; Rankin 2004a; Eschenberg 2005. This section strongly builds on the results of their research. I won't summarize these results *in toto* here for lack of space. Instead, I will select some of the grammaticalizations involving the Proto-Siouan positionals, classificatory demonstratives, and continuative aspect marking auxiliaries in this sub-branch of Siouan, in order to contrast them with Hoocąk. The following grammaticalizations of positionals and definite articles in Omaha-Ponca (OP) will be presented:

- from positionals to classificatory definite articles (§6.1);

- from classificatory definite articles to classificatory demonstratives (§6.2); and

- from classificatory definite articles/ copulas to continuative marking auxiliaries (§6.3).

It will be shown, in particular, that the Proto-Siouan positionals developed very differently in OP and the other Dhegiha languages compared to what has been discussed so far with regard to Hoocąk and some non-Dhegiha Siouan languages.

6.1 From positional verbs/auxiliaries to classificatory definite articles

All Dhegiha languages have developed remarkable paradigms of up to ten definite articles that - among other things - classify their nouns according to semantic features such as: animate vs. inanimate, agent, vs. non-agent, moving, standing, sitting, horizontal, vertical, round, scattered and singular and plural, cf. the summary in Figure 2 below.

Some of these definite articles are descendants of the Proto-Siouan posture verbs; they are marked bold in Figure 2; the others have presumably a different origin or their etymology is unclear (cf. Rankin 2004a: 209; Eschenberg 2005: 181ff). All classificatory definite articles in OP (cf. Figure 2) are multifunctional and occur in different constructions; for textual evidence and discussion, see Eschenberg (2005: 112-176):

Johannes Helmbrecht
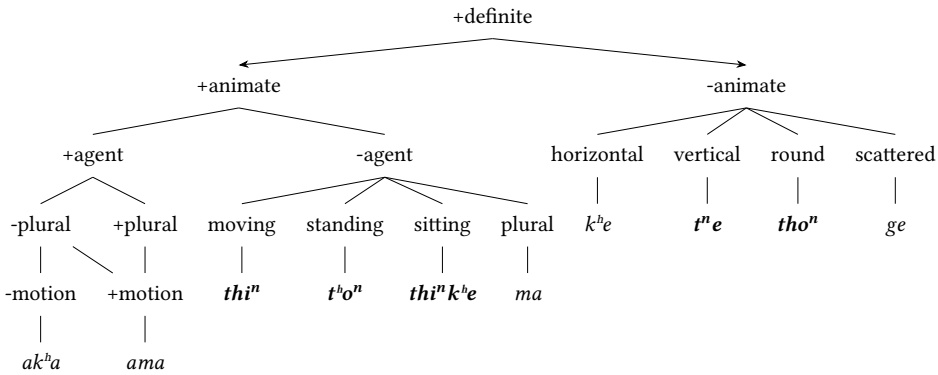
Figure 2: System of definite articles in Omaha-Ponca (cf. Koontz 1984: 144)

- they can be used as **relativizer**, some of them also as **general subordinator** (with a locative element attached to it);

- almost all classificatory definite articles in OP (except *ge* +def/-anim/ scattered(PL)) can be used as **auxiliaries indicating progressive aspect**;

- all definte articles occur as **copulas in existential and locatives clauses** of the type "it/this is.standing a dog";

- some of the definte articles in OP are used as sentence-final **evidential markers** indicating either direct evidence of the event by the speaker, inferential evidence, or hearsay;

- a few of the classifying definite articles can be used as **clause linking devices** (this holds for *t*ʰe +def./-animate/ vertical and *ge* +def./-animate/ scattered).

The question arises: how do positional auxiliaries become definite articles? Eschenberg (2005: 182-206) argues that the positional auxiliaries were reanalyzed as definite articles via their function as copulas in locative/ existential clauses. The definite articles in OP (all the forms in Figure 2) - no matter whether they come from Proto-Siouan positionals, or not – are all used as copula verbs in locative or existential clauses; cf. the illustrating examples in (21) and (22).

(21)   Omaha-Ponca

*Tizhebe **tʰe**.*   *Uthidoⁿ.*
Door   be.located.vertical 3SG.lock
'The door is positioned vertically. It is locked.' (Eschenberg 2005: 189)

(22) *Tizhebe-**tʰe**   uthidoⁿ.*
      Door-DEF.VERT 3SG.lock
      'The door is locked.' (Eschenberg 2005: 189)

In (21), the form *tʰe* 'be.located.vertical' is used as a copula in a clause of location/ existence. If the second predication "It is locked" becomes pragmatically more important than the existence of the "door", the copula may be reinterpreted as a determiner, which simply presupposes the existence of the "door"; this is illustrated in (22). No matter whether the historical scenario proposed by Eschenberg in her dissertation is correct or not, there is no evidence so far that the classificatory definite articles in OP developped from demonstratives (what one would expect). To the contrary, it is the classifying definite articles that became eventually classificatory demonstratives as will be briefly shown in the next section.

6.2 From classificatory definite articles to classificatory demonstratives

OP has preserved the "old" paradigm of Proto-Siouan demonstratives; see Table 12 above and Table 16.

These demonstratives are used variously as **demonstrative pronouns** and as **determiners**. In both usages they **cooccurred with the definite articles**, which led to the merging of the demonstrative stems plus a classificatory article. The results are grammaticalized **classificatory demonstratives** (cf. Rankin 2004b: 215).

(23) Omaha-Ponca (Dorsey 1890: 26/27)
      *Égaⁿ ðisan~´ga   mégaⁿ **še-kʰe**   ðizáiga hă,*
      so   your.younger.brother likewise **that-def.inanimate.lying** take.ye ?
      á-biamá.
      said.he-they say
      ' "Do you and your younger brother take that" he said.'

In (23), the demonstrative pronoun *še* 'that (close to hearer)' is combined with the inanimate definite classificatory article *kʰe* forming a classificatory demonstrative. This combination is phonologically one word. Other examples in the

Johannes Helmbrecht

Table 16: Omaha-Ponca "old" demonstratives

demonstrative pronouns/ determiners		demonstrative adverbs	
ðe	'this' (close to speaker)	*tu*	here
še	'that' (close to hearer),	*šu*	there
ka	'that' (remote, out of sight)	*ku*	yonder

text collection of J. O. Dorsey (1890) illustrate that these classificatory demonstratives may be formed with other definite classificatory articles (sitting/ lying/ standing) as well, and that they also may occur as determiners (see also Rankin 2004b: 215; Eschenberg 2005: 101f). The formation of classificatory demonstratives in OP resembles closely the forms in Mandan, see §5 above. The difference, however, is that in OP the deictic stems are combined with definite articles, while in Mandan these stems are combined with the positional auxiliaries. The order of elements in the classificatory demonstratives in OP is also different to the order of elements in Hoocąk, where the deictic particle follows the positional auxiliary.

6.3 From classificatory definite articles to continuative aspect marking auxiliaries

The last grammaticalization process in OP with regard to the positional verbs and classificatory definite articles that will be presented here is the development of auxiliaries that indicate continuative aspect. The Dhegiha languages had lost the Proto-Siouan positionals that were used as aspect marking auxiliaries in other Siouan languages. Instead, the Proto-Siouan positionals developed into classificatory definite articles. The next step in the historical process is that these classificatory definite articles developed into continuative marking auxiliaries. This can be concluded from the fact that the classificatory definte articles received a new verbal conjugation that differs from the inherited conjugation of the positionals. It is not fully clear how this process came about. Rankin argues that it was the Proto-Siouan sitting positional *rįk that inherited the aspect marking function in OP. This positional developed into different definite articles on the one hand, but was also combined with a bound verb *-he* 'to be in a place', which in turn could be conjugated. Later on, the other classificatory definite articles were analogically conjugated according to this model. Eschenberg contests this view. According to her, the starting point for the grammaticalization of the definite articles to continuative auxiliaries is their use as locational copulas; compare

the scenario in (24) and (25).

(24) Omaha-Ponca (Eschenberg 2005: 190)
[[*Tizhebe uthido^n*] *t^{h}e*].
Door   3SG.lock **be.located.vertical**
'The locked door is vertical.'

(25) [[*Tizhebe*] [*uthido^n*-*t^{h}e*]].
Door   3SG.lock-**be.located.vertical**
'The door is being locked.'

In (24), the copula modifies the whole NP "the locked door". This clause represents the original usage of the definite article **t^{h}e** 'the.inanimate.vertical', while in (25), the same form is used as an auxiliary and modifying only the preceding predicate. No matter, which hypothesis is correct, it remains that the continuative auxiliary in OP developed from the definite articles perhaps via their use as copulas and NOT from the old Proto-Siouan positional auxiliaries directly that were lost in Dhegiha eventually.

7 Conclusions

It has been shown that the Proto-Siouan posture verbs ("positionals" in Siouanist terminology) grammaticalized to aspect marking auxiliaries in many Siouan languages. As auxiliaries they underwent different degrees of coalescence with the full verb up to the point being an uninflected continuative aspect marker in Lakota. These aspect marking auxiliaries were ultimately lost in the Dhegiha languages. In Hoocąk, the positional auxiliaries grammaticalized to adnominal demonstratives via a subordinating construction and the coalescence with a deictic particle. The "new" adnominal demonstratives preserved a classificatory function; cf. (26).

(26) Grammaticalization of positionals to demonstratives in Hoocąk

posture verb	→	positional auxiliary/ continuative aspect marker	→	classificatory adnominal demonstrative	→	nominalizer _____ subordinator _____ relativizer

A different grammaticalization path towards demonstratives was found in Mandan. The positional auxiliaries combine with one of the "old" Proto-Siouan

Johannes Helmbrecht

demonstrative pronoun to form a "new" classsificatory demonstrative pronoun. The other path is the reanalysis of the positional auxiliary as an adnominal demonstrative; cf. (27).

(27)   Grammaticalization of positional to demonstratives in Mandan

| posture verbs | → | positional auxiliary/ continuative aspect marker | → | classificatory demonstrative pronoun of the type *dɛ-nąk* 'this one (be.sitting)' |
| | | | → | classificatory demonstrative determiner of the type *óti-hąk* 'this lodge' |
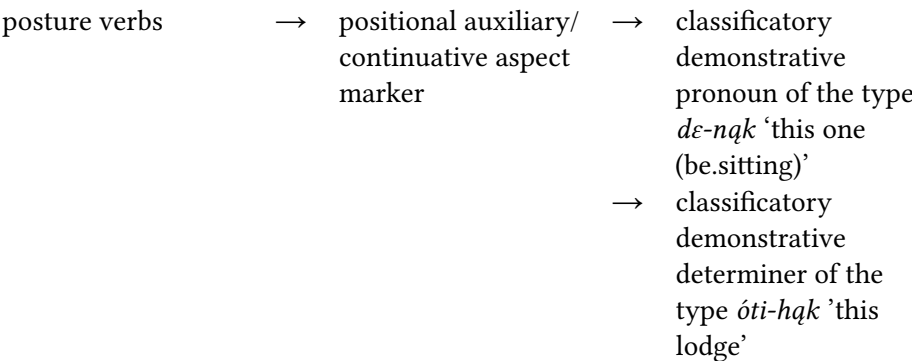
Even more different are the grammaticalization paths found in Omaha-Ponca representing the Dhegiha sub-branch of Siouan. Here the positional auxiliaries grammaticalized to classificatory definite articles perhaps via their use as copulas in existential/ locative clauses. No evidence was presented in the literature that this process was preceded by a stage in which the positionals were demonstratives. This is remarkable since it is commonly held that demonstratives become definite articles and not vice versa. Instead, classificatory demonstratives were grammticalized in OP from the definite articles by combining them with the "old" Proto-Siouan demonstratives. In addition, the classificatory definite articles developed into continuative marking auxiliaries; cf. (28).

(28)   Grammaticalization of positionals to demonstratives in Omaha-Ponca
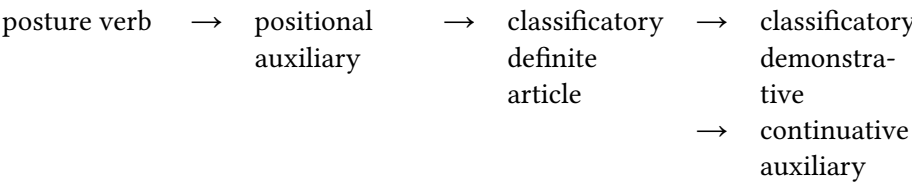
| posture verb | → | positional auxiliary | → | classificatory definite article | → | classificatory demonstrative |
| | | | | | → | continuative auxiliary |

Abbreviations

1, 2, 3,	first, second, third person	NOM	nominative
A	actor	PL	plural
ACC	accusative	POS.HOR	'be (lying/ horizontal position)'
APPL.INESS	inessive applicative prefix	POS.NTL	'be (sitting/ neutral position)'
COLL	collective marker	POS.VERT	'be (standing/ vertical position)'
CONT	continuative		
DECL	declarative	POSS.REFL	possessive reflexive
DEF	definite	PRO	pronoun
DEM	demonstrative	PROP	proper name marker
DIST	distal	PROX	proximal
DU	dual	Q	interrogative particle
E	exclusive		
EMPH	emphatic	SBJ	subject
FUT	future	SEQ	sequential
I	inclusive	SG	singular
IND.MALE	indicative with male address	SIM/LOC	simultaneous/locative
		SS	same subject marker
MV	middle voice	SV	stem vowel
N	noun	U	undergoer
NARR.PAST	narrative past tense	V	verb

References

Barron, Roger & Fritz Serzisko. 1982. Noun classifiers in the Siouan languages. In Hansjakob Seiler & Franz Joseph Stachowiak (eds.), *Apprehension. Das sprachliche Erfassen von Gegenständen, III: Die Techniken und ihr Zusammenhang in Einzelsprachen*, 85–105. Tübingen: Narr.

Boas, Franz. 1907. Notes on the Ponca grammar. *Papers of the International Congress of Americanists* 15(2). 317–337.

Boas, Franz & Ella Deloria. 1941. *Dakota grammar*. Washington, D.C.: Government Printing Office.

Cumberland, Linda A. & Robert L. Rankin (eds.). 2012. *Kaáⁿze Íe Wayáje. An Annotated Dictionary of Kaw (Kanza). Kaw-English/English-Kaw*. Kaw City, OK: Kanza Language Project of the Kaw Nation.

Deloria, Ella. 1932. *Dakota texts*. New York: Stechert.

Johannes Helmbrecht

Diessel, Holger. 1999. *Demonstratives: Form, function, and grammaticalization* (Typological studies in language 42). Amsterdam: Benjamins.

Dorsey, James & John Swanton. 1912. *A dictionary of the Biloxi and Ofo languages* (Bureau of American Ethnology, Bulletin 47). Washington, D.C.: no publisher.

Dorsey, James Owen. 1890. *The Ꞓegiha language* (Contributions to North American Ethnology VI). Washington: Government Printing Office.

Einaudi, Paula Ferris. 1976. *A grammar of Biloxi.* New York: Garland Publishing, Inc.

Eschenberg, Ardis. 2005. *The article system of Umoⁿhoⁿ (Omaha).* Buffalo, NY: University of Buffalo PhD thesis.

Givón, T. 1984. *Syntax. A Functional-Typological Introduction, Vol. 1.* Amsterdam: John Benjamins.

Graczyk, Randy. 2007. *A grammar of Crow.* Lincoln, NE: University of Nebraska Press.

Greenberg, Joseph H. 1978. How does a language acquire gender markers? *Universals of Human Language* 3. Greenberg et al. (ed.). 47–82.

Gundel, J. K., N. Hedberg & R. Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69. 274–307.

Harris, Alice & Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective.* Cambridge: Cambridge University Press.

Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization.* Cambridge: Cambridge University Press.

Helmbrecht, Johannes & Christian Lehmann (eds.). 2010. *Hocąk Teaching Materials, Vol I. Elements of Grammar, Learner's Dictionary.* Albany: SUNY Press.

Himmelmann, Nikolaus. 1997. *Deiktikon, Artikel, Nominalphrase: Zur Emergenz syntaktischer Struktur.* Tübingen: Niemeyer.

Ingham, Bruce. 2003. *Lakota* (Languages of the World/ Materials 426). München: LINCOM EUROPA.

Kaufmann, David V. 2011. *Tanêks-Tąyosą Kadakathi. Biloxi-English dictionary with English-Biloxi index.* Lawrence: The University of Kansas.

Kennard, Edward. 1936. Mandan grammar. *International Journal of American Linguistics* 9. 1–43.

Koontz, John E. 1984. *Preliminary Sketch of the Omaha-Ponka Language.* Doctoral dissertation. Boulder: University of Colorado.

Lehmann, Christian. 1984. *Der Relativsatz: Typologie seiner Strukturen, Theorie seiner Funktionen, Kompendium seiner Grammatik.* Tübingen: Narr.

Lehmann, Christian. 1995[1982]. *Thoughts on grammaticalization.* Munich: LINCOM EUROPA.

Lipkind, William. 1945. *Winnebago grammar*. New York: King's Crown Press.

Mithun, Marianne. 1999. *The languages of Native North America* (Cambridge Language Surveys). Cambridge: Cambridge University Press.

Mixco, Mauricio. 1997. *Mandan* (Languages of the World/ Materials 159). München: LINCOM EUROPA.

Oliverio, Giulia R. M. 1996. *A grammar and dictionary of Tutelo*. Doctoral dissertation. Lawrence: University of Kansas.

Parks, Douglas R. & Robert L. Rankin. 2001. Siouan languages. In Raymond J. DeMallie (ed.), *Handbook of North American Indians* (Plains 13), 94–114. Washington, D.C.: Smithsonian Institution.

Quintero, Carolyn. 2004. *Osage Grammar. (Studies in the Anthropology of North American Indians)*. Lincoln, NE: University of Nebraska Press.

Quirk, R., S. Greenbaum, G. Leech & J. Svartvik. 1972. *A grammar of Contemporary English*. London: Longman.

Rankin, Robert L. 2002. *Quapaw Dictionary*. Unpubl. manuscript. Lawrence.

Rankin, Robert L. 2004a. The history and development of Siouan positionals with special attention to polygrammaticalization in Dhegiha. *Sprachtypologie und Universalienforschung STUF* 57. 202–227.

Rankin, Robert L. 2004b. *An Ofo Grammar Sketch – based on materials collected by John R. Swanton and John R. Swanton's Ofo-English Dictionary*. Unpubl. manuscript. Lawrence: University of Kansas.

Rankin, Robert L., Richard T. Carter & A. Wesley Jones. 1998. Proto-Siouan phonology and grammar. In Xingzhong Li, Luis Lopez & Tom Stroik (eds.), *Papers from the 1997 Mid-America linguistics Conference*, 366–375. Columbia: University of Missouri-Columbia.

Rankin, Robert L., Richard T. Carter, A. Wesley Jones, John E. Koontz, David S. Rood & Iren Hartmann. 2015. *Comparative Siouan Dictionary*. Accessed on 2015-09-24. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://csd.clld.org.

Ranklin, Robert. 1977. From verb to auxiliary to noun classifier and definite article: Grammaticalization of the Siouan verbs 'sit,' 'stand,' 'lie.' In R. L. Brown, K. Houlihan, L. Hutchinson & A. MacLeish (eds.), *Proceedings of the 1976 Mid-America Linguistics Conference*, 273–283. Minneapolis, Minnesota: University of Minnesota.

Rood, David. 1979. Siouan. In Lyle Campbell & Marianne Mithun (eds.), *The languages of native America: Historical and comparative assessment*, 236–298. Austin: University of Texas Press.

Johannes Helmbrecht

Sohn, Ho-Min. 1999. *The Korean language.* Cambridge: Cambridge University Press.

Traugott, Elizabeth C. 1992. Syntax. In R. M. Hogg (ed.), *The Cambridge History of the English Language. Vol. 1: The Beginnings to 1066*, 168–289. Cambridge: Cambridge University Press.

Chapter 8

Grammaticalization of tense/aspect/ mood marking in Yucatec Maya

Christian Lehmann

University of Erfurt

Maybe the most pervasive among the changes analyzable as cases of grammaticalization in the languages of the Yucatecan branch of the Mayan stock is the formation of auxiliaries that allow finer tense/aspect/mood distinctions than the status suffixes inherited from Proto-Maya. It has been continually productive since colonial times. While this amounts to a replacement of the status system, it follows strictly language-internal patterns. And while the source constructions form a rather heterogeneous set, they converge onto a common TAM auxiliary pattern in Modern Yucatecan.

1 Introduction

This study is devoted to the grammaticalization of auxiliaries in Yucatec Maya, whose functional side is the formation of a complex tense/aspect/mood (TAM) system. In this, it aims at fulfilling several purposes at once. It is, in the first place, a contribution to a historical grammar of Yucatec Maya. To this end, it brings together a large set of data, contextualized in their historical situation. A side effect of this enterprise is a diachronic perspective on the system of present-day Yucatec Maya, which may, as usual, open an additional, viz. dynamic, dimension of understanding it. On the other hand, the analysis tries to systematize the facts in terms of a theory of grammaticalization so that they may become comparable with relevant facts of other languages. To secure understanding for the non-specialist, some elements of Mayan grammar will be explained in §3.

Some of the data used are actually in a diachronic relationship, viz. data from the history of Yucatec Maya. Most of the data of other Mayan languages belong

to recent stages of their evolution. Following established methodology, they will be projected onto the diachronic axis and be taken to represent stages of a development.

A word is necessary on the orthography. Yucatec Maya has had distinctive vowel length and tone at least for the period of its documented history, although it does not share tone with any of its sisters. Moreover, the glottal stop and /h/ are phonemes, and both can form a syllable coda. Since all of this is alien to Spanish, the orthography of the Colonial Yucatec Mayan sources hides important phonological information. These phonological properties have been marked consistently in the orthography only from the second half of the twentieth century on. For this reason and in order to facilitate diachronic comparison to the non-specialist, examples from Colonial Yucatec Maya are first quoted literally from the sources and then coupled with a representation in contemporary scientific orthography (which is, alas, not the one adopted currently by Mexican authorities; s. Lehmann 2015).

2 Prehistory and history of Yucatec Maya

The Mayan languages of today are spoken in a culture area called Meso-America. Some of the Mayan languages are so dissimilar that they may have branched off from the common stock as early as 2000 BC. The Yucatecan branch was the second to separate from the rest of the Mayan family. This took place during pre-classical times in terms of Mayan history, at the latest about 1000 BC. Both genetically and geographically, the closest neighbor is the branch of the Ch'olan languages, which are clearly mutually unintelligible with the Yucatecan languages. The Yucatecan languages are spoken on the peninsula of Yucatan and in more southern regions of the lowland in Belize, the Petén region of Guatemala and the Mexican state of Chiapas. The internal subdivision of this branch is relatively recent. It has the form shown in Figure 1.

Mopán on the one hand and the other Yucatecan varieties are hardly mutually intelligible and are commonly regarded as different languages. The latter three varieties do not differ more from each other than British and American English. The period of their separation does not exceed a few hundred years and is, thus, far shorter than the period of separation of the dialects of German, British English or Italian. They are mutually intelligible and should be regarded as dialects of one language rather than as distinct languages.

Mopán split off at the end of the first millennium AD. The Itzá people apparently emigrated from the peninsula to the Petén in the fourteenth century,
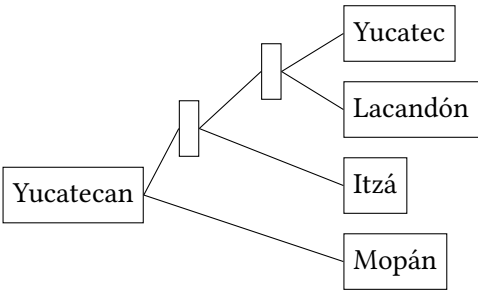
Figure 1: The Yucatecan branch of the Mayan languages

although keeping contact with Yucatec Mayas. The Lacandón people, too, are Mayas of Yucatán who retreated from the peninsula into the woods of Chiapas in order to avoid contact with the Mexican civilization. The closest relative of Yucatec is (Southern) Lacandón. It is a dialect that split off the main variety in the 18th cent. and preserves some archaic traits. The periods of the history of Yucatec Maya itself may be depicted as in Figure 2.
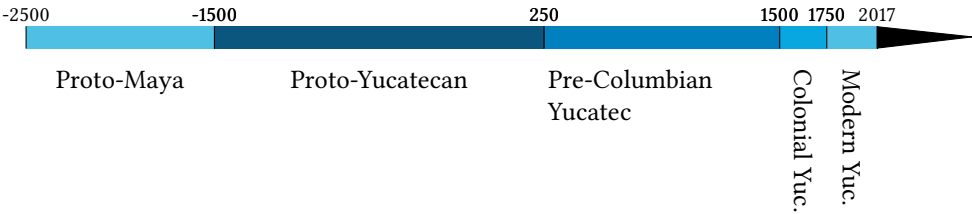
Figure 2: Periods of Yucatec language history

The inscriptions and codices of the Pre-Columbian Mayan culture span a period from roughly 250 to 1500 AD. They represent some Ch'olan language and are therefore relatively close to Pre-Columbian Yucatec. However, the glyphic writing as it has been deciphered up to now does not represent the morphology of the language very well, so that for our purposes, written documentation of the language starts with the Spanish conquest.

Yucatec Maya has been historically well attested since the early times of Spanish colonization.[1] This period of the language history is called Colonial Yucatec Maya, often also Classical Yucatec Maya. Apart from having a longer docu-

---

[1] While most of the hieroglyphic texts appear to represent the Ch'olan branch, one or another of the surviving codices, which probably stem from the fifteenth century AD, may be in Yucatec.

Christian Lehmann

mented history than most Amerindian languages, Yucatec Maya also boasts a set of early grammars and dictionaries as shown in Figure 3.
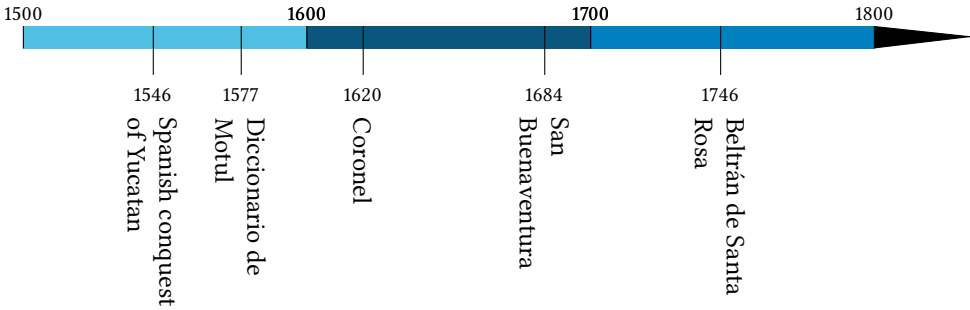
Figure 3: Colonial grammars and dictionaries of Yucatec Maya

The earliest source is the *Diccionario de Motul*[2], which some assume to be compiled around 1577.[3] The earliest grammars – and still among the earliest sources of data for Yucatec Maya – are Coronel 1620[4] and San Buenaventura 1684. In the course of the eighteenth century, Colonial Yucatec Maya passed into Modern Yucatec Maya (MYM). Beltrán 1746 is assumed to mark the transition between the two stages (Smailus 1989: 4).

Thus, the documented history of Yucatec Maya begins with colonial documents of the 16th century. Its prehistory is indirectly represented in Mayan hieroglyphic writing and may be accessed by internal reconstruction and historical comparison with cognate languages. Data from the other Yucatecan languages are from the second half of the 20th century. Lacandón preserves some archaic traits, lending thus additional support to reconstructions.

Given all this, reconstruction of Proto-Yucatecan is in a comparatively favor-

---

[2] The *Diccionario* or *Calepino de Motul* was first published in Martínez Hernández 1929. In the examples, it is referred to as *Motul*.

[3] Since its first published edition, the manuscript of the *Diccionario de Motul* has been attributed to Fray Antonio de Ciudad Real (1551-1617) and been dated to 1577. Now he may well be the author, the more so as he is known to have worked on Mayan language and culture until his death. However, he started living in Mérida only in 1573; and in 1577 he was 26 years old. Consequently, he either is not the author (but only a compiler of material gathered by others), or the year of completion must be much later. Hanks (2010: 164-168) discusses the problem extensively and essentially pleads agnostic.

[4] In quotations, I use Martínez Hernández 1929 for the page numbering, as it reproduces the pagination of the original edition; but I quote the text from the (more reliable, but unpaginated) online edition of http://www.famsi.org/reports/96072/coronelgmr.htm. (The critical edition of Coronel 1998b was not available to me).

able methodological situation. Not only can we reconstruct the diachrony by comparing four languages which are related closely enough to elucidate each other but different enough to provide variation which may be projected onto the diachronic axis. We also have 450 years of documented history in the case of Yucatec, which can substantiate or falsify our diachronic hypotheses to some extent. Thirdly, grammarians have described different stages of the history for the same period, sometimes even noting explicitly grammaticalization phenomena observable at their time. Under such circumstances, responsible diachronic analysis may reach back approx. 1,000 years, which is about the point where Proto-Yucatecan began to split up.

3 Typological sketch of Yucatecan languages

All Mayan languages are very much alike in their morphological and syntactic structure,[5] with some of the more principled differences being taken up below. The lexemes and the grammatical morphemes filling the structural slots are generally cognate within each of the subfamilies, while there are great differences among the subfamilies in this respect. Consequently, while the Yucatecan languages form a homogeneous group, this subgroup differs from other subgroups of the Mayan family chiefly in the individual lexical and grammatical morphemes and, to a lesser extent, in grammatical structure. We will here focus on the grammatical structure of the Yucatecan subfamily and mention deviations from Proto-Mayan *suo loco*.

Apart from numeral classifiers, the typologically notable features of the word-class system are limited to the subclassification of the major classes. Both nouns and verbs are subclassified according to relationality: absolute and relational nouns differ in morphology and syntax similarly as do intransitive and transitive verbs. If the valency of a stem includes a place for such an additional actant, then there is a pronominal index for it. If a clause lacks such an actant (no matter whether represented by an NP), the base must be derelationalized. And vice versa for an absolute or monovalent base. Moreover, besides pure verbs, there is a closed class of verboids which share all morphological and syntactic properties with verbs except that they do not inflect for status (§4.4) and therefore do not combine with an auxiliary (§4.5).

Mayan languages lack the category of case throughout. They do have a productive category of prepositions – most of them denominal in origin – but very few

---

[5] A recent typological overview of the Mayan family is in Grinevald & Peake 2012.

Christian Lehmann

primary prepositions; and the Yucatecan languages have only one fully grammaticalized preposition, *ti'* Loc, which marks the indirect object (as in 41 and 42b) below), local and other concrete relations.[6] Under these conditions, structural relations of modification are underdeveloped; the syntax is characterized by government. All dependency constructions are head-marking: indexes cross-reference the subject and direct object of a verb, the possessor in a nominal construction and the complement of a preposition. The index is obligatory, the nominal dependent is optional. The verb with its cross-reference indexes, possibly preceded by an auxiliary (s. §4.2), constitutes a full clause. No nominal or pronominal constituents are necessary.

Alignment of fundamental syntactic relations was ergative in Proto-Mayan. Some Mayan subfamilies have preserved this alignment to a large extent. The Yucatecan languages show traces of syntactic ergativity in focus constructions (Bricker 1981); but otherwise ergativity is restricted to a split in the index paradigm of the intransitive predicate conditioned by status, to which we return in §4.1.

The morphology is characterized by a medium degree of synthesis. Most affixes are suffixes. Most of the morphology is agglutinative; still, there are, especially in Yucatec Maya, several internal modifications. While declension is comparatively simple, verbs inflect for many conjugation categories. One of these must be singled out from start as it plays an important role in subsequent sections: The first morpheme after the (simple or derived) verb stem is a so-called status suffix, which comprises the subcategories of dependent status proper, aspect and mood. It is illustrated by the dependent incompletive suffix in (3) below. Word formation includes compounding and derivation, both in the nominal and in the verbal sphere. The entire verb derivation is based on transitivity: every stem is either transitive or intransitive; and this determines the allomorphy of conjugation categories, esp. of the status morphemes.

Mayan languages lack a copula.[7] The word order must have been left-branching in some remote pre-historic epoch. This is the environment in which the morphological categories marked by verb suffixes (s. §4.4), and possibly the phrase-initial nominal determiners and modifiers, too, originated. The proto-language then switched to right-branching syntax; proto-Maya was right-branching. To this day, Mayan languages are left-branching or juxtapositive only in the nominal syntax, as shown in Table 2; the rest of the syntax is right-branching, as detailed in Table 1.

---

[6] corresponding both etymologically and functionally to Ch'olan *tyi*

[7] Colonial Yucatec Maya features a suffix -*h* Cop, exemplified in (22), which verbalizes nominal predicates.

Table 1: Right-branching constructions

predicate	subject
verb	actant
verbal complex	adjunct
auxiliary	clause core
nominal group	relative clause
nominal group	nominal possessor
preposition	complement NP
conjunction	clause

Table 2: Non-right-branching constructions

short adverb	verb
adjective attribute	noun
numeral	numeral classifier
numeral complex	nominal group
determiner	nominal group

(The vague wording of the Table 2 heading reflects the fact that some dependency relations inside the NP (or DP) are less than clear.) One might add to Table 2 the clitic pronominal index preceding a verb or a possessed nominal and cross-referencing the subject or the possessor, resp. (i.e. the "Set A" index of §4.1).

Marked information structure provides for two sentence-initial positions to be occupied by main constituents, viz. the position of left-dislocated topical constituents and the focus position. The maximum configuration was dubbed LIPOC (language-independent preferred order of constituents) in Dik 1981:189ff and may be represented by Figure 4. (1) is an example.

[left-dislocated topic   [focus   extrafocal clause]]

Figure 4: Extended sentence structure

(1)   Modern Yucatec Maya
      le   chaan lak   he'l=a'   in   kiik
      DEM little   bowl PRSV=R1 A.1.SG elder.sister

Christian Lehmann

> *síih-mah-il*   *ten*
> give.as.present-PRF-DEP(B.3.SG)   me
> 'this little clay bowl, my elder sister gave it to me' (ACC_0039)

The left-dislocated constituent is marked by a referential enclitic, R1 in (1). The paradigm contains an element (R3) which functions as a topicalizer if the deixis is neutral.[8] The focus itself (*in kiik* in (1)) is not marked, but the extrafocal clause is marked by a dependent status suffix, *-il* in (1) (s. §4.4).

4 Verbal categories

In this section, we will pursue the fate of some categories in the functional domain of tense/aspect/mood in the Yucatecan languages. The starting point will be Colonial Yucatec Maya as documented in the sources enumerated in §2.

4.1 Pronominal indexes

All Mayan languages have at least three sets of personal pronominal formatives. All but one of these paradigms are clitic or bound and function as cross-reference indexes; the last is a set of independent personal pronouns. The main paradigms of bound indexes are called Set A and Set B in Mayan linguistics. The functions of the pronominal sets are as follows:

- indexes of Set A cross-reference the possessor of a nominal group and the actor of the transitive verb. Moreover, in the split-subject marking languages including those of the Yucatecan branch, they cross-reference the subject of an intransitive verb in some verbal statuses (§4.4). Thus, the syntactic function alignment based on the distribution of set A is accusative.

- indexes of Set B cross-reference the subject of a non-verbal clause and the undergoer of the transitive verb. In the split-subject marking languages, Set B also cross-references the intransitive subject in the complementary subset of statuses. Thus, the syntactic function alignment based on the distribution of Set B is ergative.

---

[8] The Yucatecan languages differ in the details. Itzá continues Pre-Columbian grammar in allowing the topicalizer *-e'/-eh* to follow – directly or at a distance – the deictic clitics (Hofling 1991:14f). Lacandón lacks the entire paradigm of referential clitics, including the topicalizer.

- The free pronouns are reinforced forms of Set B forms. They appear as the complement of a preposition, as left-dislocated topic and in focus position. Some languages including Yucatec Maya have enclitic variants which function as indirect object, as does *ten* in (1).

The labels 'Set A' and 'Set B' originate in the times of American structuralism. They are deliberately obscure and mnemonically unhelpful. We will nevertheless have to use them because the functions which might provide more practical labels are heterogeneous. At any rate, it may be helpful to bear in mind the following equivalences with more familiar labels of interlinear glossing:

- A = SBJ/POSS

- B = ABS.

Table 3 shows the Modern Yucatec forms of sets A and B. For 1st ps. pl., the exclusive form is given. All of these pronominal elements are free forms at the stage of Proto-Maya. The parenthesized glides are conditioned by a vowel-initial host of the pronominal index.

Table 3: Pronominal paradigms in Modern Yucatec Maya

		A	B
sg.	1	in (w-)	-en
	2	a (w-)	-ech
	3	u (y-)	∅
pl.	1	k(a)	-o'n
	2	a (w-)... -e'x	-e'x
	3	u (y-)...-o'b	-o'b

In all Mayan languages, the Set A index precedes the possessed nominal, cross-referencing the possessor. (2) provides representative examples of the indexes with verbs:

(2) Modern Yucatec Maya

    a. *h*   *bin-ech*
       PFV go (CMPL)-B.2.SG
       'you went'

Christian Lehmann

b. *t=u*   *t'an- ech*
    PFV=A.3 call(CMPL)-B.2.SG
    'he called you'

The examples are in the completive status, which triggers ergative marking in all Mayan languages. The Set A index immediately precedes the transitive verb. The Set B index is a suffix to the verb.

In the Yucatecan languages, Set A forms belong to a species of enclitics which are not banned from initial position. If they follow a word in the same phrase, they form a phonological unit with it. Since they syntactically depend on what they precede, they cliticize to what is, in grammatical terms, the wrong side. In the examples, clisis of Set A forms is marked by an equal sign (although some of the sources mistakenly write them as prefixes).

4.2 Verbal clause structure

Tense, aspect and mood are verbal categories and therefore possible only in verbal clauses. Other kinds of predicates have to be verbalized if these categories are to be specified. Therefore, we can narrow down the analysis to the verbal clause. With some simplification, the verbal complex has the structure shown in Figure 5. (3) is a transitive finite verbal complex.

verbal complex
finite verb
index A | verb stem   -status   -index B
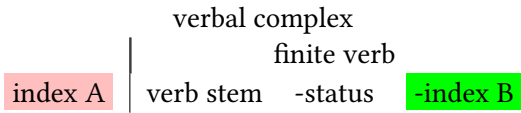
Figure 5: Transitive verbal complex

(3)   Colonial Yucatec Maya
    u ppaticech
    u   p'at-ik-ech
    A.3 leave-DEP.INCMPL-B.2.SG
    '(that) he leaves you' (Motul s.v. *Hun chilbac*)

The basic clause structure is 'predicate – subject'. If it is a verbal predicate, the verbal complex of Figure 5 comes first, then follow the free complements and adjuncts. The most elementary independent verbal clause at the stage of Colonial Yucatec consists of a verbal complex in completive status and its dependents, as in (4).

(4) Colonial Yucatec Maya
u kamah nicte in mehen

| *u* | *k'am-ah-*∅ | *nikte'* | *in* | *mehen* |

A.3 get-CMPL-B.3.SG flower A.1.SG son

'my son got the flower (i.e. got married)' (Motul s.v. *kamnicte*)

Already in Colonial Yucatec, many verbal clauses are introduced by a formative which codes tense, aspect or mood and which we will call an auxiliary (s. §4.5 for discussion of the appropriateness of this term). In Modern Yucatec, this is the default for independent verbal clauses. At this stage, the verbal complex with its dependents as illustrated by (4) only forms a **clause core**, while an independent declarative verbal clause generally (except in perfect status) requires an auxiliary in front of it. Figure 6 formalizes this construction. The second clause of (5) illustrates it with the recent past auxiliary.

verbal clause						
	verbal clause core					
		verbal complex				
			finite verb			
auxiliary	index A	verb stem	-status	-index B	dependents	referential clitic
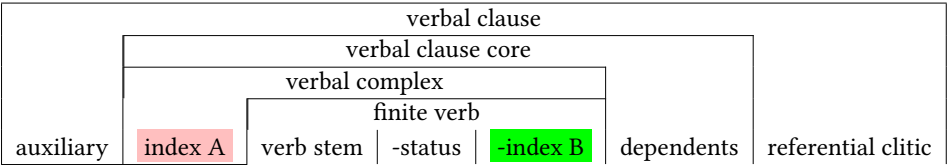
Figure 6: Verbal clause

(5) Modern Yucatec Maya

| *In* | *watan=e'* | *mina'n* | *way=e';* | *táant* | *= u* |

A.1.SG wife=TOP NEG.EXIST(B.3.SG) here=R3 REC.PST =A.3 go(INCMPL)=R3

bin=e'.

'My wife isn't here; she just left.' (BVS_05-01-36.2)

The last element in Figure 6 is the referential clitic conditioned by some of the auxiliaries, the recent past auxiliary being one of these.

4.3 Nominalization

Mayan languages generally lack an infinitive. The verb has a set of non-finite forms, some with nominal (incl. adjectival), some with adverbial function. Here we are concerned only with bare deverbal nouns, so-called action nouns, and with the processes which do no more than convert a verbal into a nominal constituent.

Christian Lehmann

From intransitive verb bases, action nouns are formed by two such processes. For agentive intransitive verbs, the verb stem also serves as an action noun stem, as in *óok'ot* 'dance' and *meyah* 'work'. For inactive intransitive verbs, an action noun, or rather a process noun, is formed by suffixing a morpheme *-Vl* to the verb root, where V is a copy of the root vowel, as in *wen-el* 'sleep (n.)' and *kóoh-ol* 'arrival'. Action nouns of intransitive bases are optionally possessed by their underlying subject, as in *in meyah* 'my work' and *u wenel* 'his sleep'. (6) provides examples of intransitive action nouns. (6a), with an agentive stem, lacks an index, while #b and #c show a Set A index in *genitivus subjectivus* function.

(6) Colonial Yucatec Maya

    a. ti canan
       ti' kanáan
       LOC watch

       'for watching' San Buenaventura (1684: 14v)

    b. et hazac ech ti in hanal
       ethas-ak-ech   ti' =in   han-al
       just.in.time-past-B.2.SG LOC =A.1.SG eat-DEP

       'you arrived just in time (to meet me) at having my meal' (Beltrán de Santa Rosa 1746: §299, p.132)

    c. in káti a benel
       in   k'áat-ih a ben-el
       A.1.SG want-CFP A.2 go-DEP

       'I want you to go' (Coronel 1998a: 51)

(7) Colonial Yucatec Maya

    a. v kin ocçah
       u k'iin ook-s-ah
       A.3 day enter-CAUS-INTROV

       '(it is) the sowing season' (Coronel 1998a: 56)

    b. in káti a cámbeçic in mehén
       in   k'áat-ih a kanbes-ik-∅   in   mehen
       A.1.SG want-CFP A.2 teach-DEP-B.3.SG A.1.SG son

       'I want you to teach my son' (Coronel 1998a: 50)

If the verbal base is transitive, there are two possibilities. The first consists in introverting the base, i.e. detransitivizing it by suppressing the direct object

position. Once this is done, the stem is nominalized like an agentive intransitive verb stem, which means that the introversive stem also serves as an action noun. Examples based on transitive roots are *xok* 'read' – *xook* (read\Introv) 'reading, study' and *k'ay* 'sing' – *k'aay* 'singing, song'. For derived transitive stems, introversion is marked by the suffix *-ah*: *kambes* 'teach' – *kambes-ah* (teach-Introv) 'teaching' (as in (74) below), *hets'kun* 'settle' – *hets'kunah* 'settlement'. Such a form also appears in (7a). The other possibility of nominalizing a transitive base consists in providing it with the dependent status suffix *-ik* and accompanying it by the Set A and Set B indexes for subject and object. This is shown in (7b).

The two nominalizing suffixes *-Vl* and *-ik* are glossed as dependent status in (6)f. They will become incompletive suffixes on their way to Modern Yucatec. The appearance of the Set A index in front of the nominalized verb is conditioned by rules of syntax which will not be detailed here. It suffices to note the following: In Modern Yucatec Maya, this element is missing (under coreference) from the purpose part of the motion-cum-purpose construction if its verb is intransitive, and occasionally also if it is transitive. This will be taken up in §4.8. In Lacandón, incompletive verbal complexes without a Set A index are widely used in nominalizations, as in (8).

(8) Lacandón
Ten ti'   met-ik   baalche',   Yum-eh.
I   LOC make-INCMPL honey.beer lord-VOC
'I am for making honey beer, my lord.' (Bruce S. 1974: 28)

The subordination of the nominalized verbal construction by the all-purpose preposition *ti'* illustrated by (6) and (8) deserves special attention. If the clause thus subordinated follows the main clause, it may be a purpose clause. This is still so in Modern Yucatec and Lacandón, witness (9)f.

(9) Modern Yucatec Maya
Meet   hum-p'éel léech ti' =k   léech-t-ik   le
make(IMP) one-CL.INAN trap LOC =A.1.PL trap-TRR-INCMPL DEM
haaleh=a'!
paca=R1
'Make a trap for us to trap this paca!' (RMC_1993)

(10) Lacandón
ts'a'   ten t=in   wil-ik
give(IMP) me LOC=A.1.SG see-INCMPL
'give it to me for me to see' (Bruce S. 1968: 63)

If, however, the subordinate clause precedes the main clause, the same preposition instead conveys simultaneity of the situation of the main clause with the background situation of the subordinate clause. This is illustrated by (11) with an intransitive nominalized verb. (12), with a transitive one, shows that this reading also occurs if the nominalized clause is postposed.

(11) Modern Yucatec Maya
  hach bin t=u t'úub-ul   k'iin=e' táan
  really QUOT LOC=A.3 submerge\DEAG-INCMPL sun/day=TOP PROG
  y-isíins-a'l   =u yatan yuum   ahaw
  A.3-bathe-INCMPL.PASS =A.3 wife master/father chief
  'Exactly at sunset, the chief's wife was washed' (HK'AN_502)

(12) Modern Yucatec Maya
  Ki'mak wáah bin y-óol yuum   ahaw t=u
  happy INT QUOT A.3-mind master/father chief LOC=A.3
  yil-ik!
  see-INCMPL(B.3.SG)
  'How happy was the king to see him!' (HK'AN_527)

We will meet this construction again at the genesis of the progressive aspect (§4.7.3).

4.4 Status

In all Mayan languages, the verb has a suffixal slot for a category called status, which comprises the subcategories of dependent status proper, aspect and mood. These suffixes belong to the earliest layer of the diachrony (they must antedate the introduction of right-branching word order in Proto-Mayan) and are completely grammaticalized. This implies that they mostly lack a clear semantic function and are instead conditioned by the construction. While the category of status itself and most of its subcategories are shared among Mayan languages, there is a great deal of heteromorphy among them, just as most statuses display a complicated allomorphy within each language.

All of verbal morphology and syntax depends on transitivity. Every verb stem is either transitive or intransitive, and this can only be changed by derivational means.[9] Transitivity is the major factor in conditioning allomorphy in status

---

[9] Already Beltrán de Santa Rosa (1746: §§107 and 150-158) is quite explicit about this (cf. (42) below), although his orthography represents neither tone nor the glottal stop, both of which play an important role in the morphological processes manipulating transitivity distinctions.

morphemes. The core of the paradigm of status morphemes for finite forms is shown in Table 4, which presents the forms in colonial orthography. For lack of relevance to our discussion, Table 4 omits the imperative, the perfect (only available for transitive verbs, anyway) and some intransitive conjugation classes. 'V' represents a copy of the root vowel; '/' and parentheses indicate allomorphy.

Table 4: Status conjugation of Colonial Yucatec Maya

status	stem class aspect/mood	intransitive basic	intransitive derived	transitive basic	transitive derived
plain	subjunctive	-Vc	-n-ac	-Vb	(-e)
	completive	(-i)	-n(-ah)(-i)		-ah
dependent	subjunctive		-ebal		-ic
	completive	-ci	-n-ici		-(i)ci/-i[10]
	incompletive		-Vl		-ic

Transitive finite forms are preceded by Set A clitics and followed by Set B suffixes as shown in Figure 5. Intransitive verbs, instead, take Set B suffixes in the plain forms, but Set A clitics in dependent forms. The finite verb forms in Table 5 illustrate the status conjugation of Table 4 for an intransitive and a transitive example verb.[11]

In the Yucatecan languages, aspect plays a more important role than tense. In Colonial Yucatec, there is one grammaticalized tense, the suffixal perfect (illustrated by (1) above). Past time is optionally marked by the adverb *cuchi* (i.e. *kuchih*) 'formerly' (Modern Yucatec Maya *ka'ch-il*), but is otherwise implied by most occurrences of the completive aspect (as in (4), which is essentially perfective.[12] Future is one of the senses of subjunctive status and optionally coded by auxiliaries which we will come to in subsequent subsections.

Dependent status is used in the extrafocal clause of a cleft-sentence (as in (18) below) and in certain complement clauses, examples of which may be seen in (47)f.b. Dependent status is, in fact, more frequent in the texts than plain status,

---

[10] The allomorph *-i* appears if the subject is the focus constituent of a cleft-construction.

[11] The sources do not provide examples for all persons, so that some of the forms entered in Table 5 are constructed by the grammarians' rules rather than primary data.

[12] Traditional terminology in Mayan linguistics designates as completive vs. incompletive what could also be called perfective vs. imperfective, were it not for the auxiliaries to be mentioned below, which go under the latter terms. S. Vinogradov 2016 for an attempt at semantically characterizing these two values of the status category.

Christian Lehmann

Table 5: Examples of finite verb complexes in Colonial Yucatec Maya

status	stem class aspect/mood	intransitive (basic)	transitive (derived)
plain	subjunctive	*cim-ic-en*	*in cambes-ech*
		'(that) I die'	'(that) I teach you'
	completive	*cim(-i)-en*	*in cambes-ah-ech*
		'I died'	'I taught you'
dependent	subjunctive	*in cim-ebal*	*in cambes-ic-ech*
		'(that) I may die'	'(that) I may teach you'
	completive	*in cim-ci*	*in cambes-ic-i-ech*
		'(that) I died'	'(that) I taught you'
	incompletive	*in cim-il*	*in cambes-ic-ech*
		'(that) I die'	'(that) I teach you'

especially in the incompletive. It appears every time that the full verb is preceded by another main constituent or by an auxiliary. Among the dependent statuses, the default is the incompletive. As a matter of fact, the incompletive dependent morphemes are nothing else than the nominalizers for intransitive and transitive verbs already reviewed in §4.3. These are the forms that we will meet most frequently in the periphrastic constructions to be analyzed below. The completive and subjunctive dependent forms involve a high degree of syncretism, hardly occur in the texts, and even the colonial grammarians are not sure about their form and function. Some of the forms fossilize, but the two subcategories themselves disappear as the status category reaches the stage of the modern Yucatecan languages. In other words, (apart from the perfect) the values of the status category in Modern Yucatec are 'subjunctive' and 'completive' (erstwhile: plain) and 'incompletive' (erstwhile: dependent).

There are more respects in which the paradigm of Table 4 is unstable. Its basic form, and the only form that a simple declarative sentence can be based on, is the plain completive.[13] All the other status forms occur in extended or complex or non-declarative sentences. The plain status obviously lacks the incompletive subcategory. This means that any kind of imperfective aspect – and as we shall see, much semantic differentiation is possible here – requires marking beyond

---

[13] It seems that Mayan languages are among those in which perfective aspect is the default aspect for verbal clauses.

the paradigm of Table 4, which entails complex constructions involving dependent statuses. The situation is similar in the other Mayan languages. All of them have an incompletive or imperfective aspect. There is, however, great heteromorphy; and mostly the syntactic conditions are as in the Yucatecan branch, viz. an auxiliary is needed in addition to the status morpheme (Vinogradov 2014).

Colonial grammars start the description of verbal morphology with a category called present which involves incompletive status. It will be analyzed extensively in §4.9. It is a rather complex periphrastic construction which is not at all basic to the system. It figures so prominently in the grammars essentially on account of a methodological mistake on the part of the grammarians (s. p. 34. The first to recognize this is Beltrán de Santa Rosa (1746: §§60, 172). He tentatively adduces as present a cleft-construction again containing the incompletive dependent status, which we must forego here.

The status paradigm is alive to this day, but given its high degree of grammaticalization, it is fragile. Several endings appear only *in pausa* and are syncopated otherwise (Beltrán de Santa Rosa 1746: §§135-147). Some of the allomorphy is utterly complicated, syncretistic and constantly exposed to variation. For instance, while the subjunctive of root transitives ends in *-Vb* for San Buenaventura 1684, Beltrán de Santa Rosa 1746: §112 says that this is now out of use, and the ending is *-e* (as it used to be for derived transitives).

4.5 Periphrastic aspects

There is a small set of syntagmatic positions at the left clause boundary, i.e. following any left-dislocated topic as shown in Figure 4 and immediately preceding the clause core. These positions may be plotted as in Figure 7:[14]

a	b	c	
Conjunction	Focus	Auxiliary	Verbal Clause Core
d			
Superordinate Predicate			
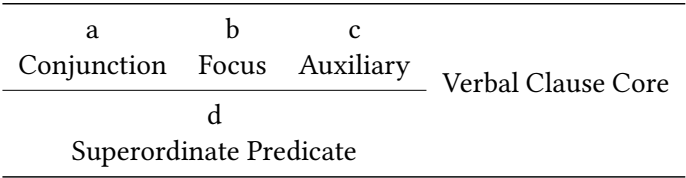
Figure 7: Clause-initial syntagmatic positions

---

[14] The left-dislocated topic of Figure 4 precedes (all the positions shown in) Figure 7. The rest of Figure 4 is a cleft-construction. However, a focused constituent may also precede a full clause, as shown in Figure 7.

Christian Lehmann

    a. The Conjunction slot may be occupied by conjunctions and other sentence-initial particles, as the conjunction in (16b) and the negator of (20b) and (41).

    b. The Focus slot may be occupied by focused constituents, as in (18).

    c. The Auxiliary slot may be occupied by grammaticalized auxiliaries, such as (28).

    d. Instead of all of this, a verbal clause core may be preceded by a superordinate predicate like the phase verb in (47), the modal verboid in (23) and one of the non-grammaticalized auxiliaries to be analyzed in §4.7. While the positional relation between any of the elements of #a – #c and the verbal clause core appears to be the same as the positional relation between such a superordinate predicate and the verbal clause core, the syntactic relation is different, since the superordinate predicate is not, of course, a constituent of the clause in question, but rather takes the clause core as a dependent, as shown in Figure 9 below.

Distributional relations between elements of the three classes shown in Figure 7 are complex, involving several conditions of mutual exclusion. In any case, none of the three slots is occupied obligatorily, and most frequently only one of them is occupied. As a consequence, any of the four kinds of elements mentioned in #a – #d may form a binary construction with an ensuing clause core. This is a structural pattern apparently inherited from Proto-Maya. It is an important presupposition for a reanalysis by which any such element may be reinterpreted as an auxiliary. As we will see, elements occupying slots #b – #d are, in fact, frequently so reanalyzed.[15]

Since the material ending up in the Auxiliary position of Figure 7 is so heterogeneous, its relation to the rest of the clause differs accordingly, and consequently the constructions with slot fillers of the four above kinds are syntactically different. The differences are reflected morphologically on the full verb, which depending on the construction is in the dependent incompletive, the completive or the subjunctive status. As we will be concentrating on such constructions in which the element in question gets grammaticalized to an auxiliary, the result is that the auxiliary conditions the status. Figure 8 takes up Figure 6 and in addition visualizes this dependency.

---

[15] In terms of Bisang 1991, esp. 511-513 and 535f, the auxiliary position of Figure 7 is an "attractor position", that is, a position which acts as a melting-pot for material recruited from different sources and grammaticalized in this position.
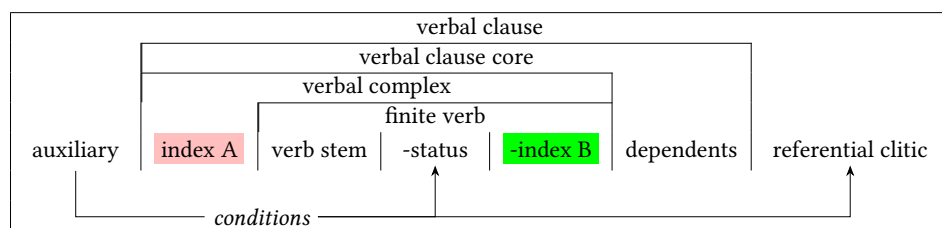
Figure 8: Syntagmatic relation between auxiliary and status

The first thing to be noted about Figure 8 is that the full verb is finite. This is a peculiarity of Yucatecan periphrastic constructions whose diachronic explanation will become clear in the following sections. As already shown in Figure 6, in the Yucatecan languages, the pronominal indexes do not combine with the auxiliary, but with the full verb. Thus, the auxiliary deserves its name only insofar as it carries tense/aspect/mood information. Person and number, however, are marked on the full verb, and consequently it is indeed finite. The discussion of the applicability of the auxiliary concept to this class of formatives will be taken up in §4.10.2.

There is in Yucatec a large variety of tenses, aspects and moods that are coded in the initial position of Figure 8.[16] None of the colonial grammars provides a systematic account of them. There are at least two reasons for this. Firstly, these grammars depend on the model of Latin grammar, which almost totally lacks auxiliaries, conjugation being essentially synthetic. Secondly, virtually none of the auxiliaries of Colonial Yucatec Maya is inherited and, thus, firmly entrenched in the system. While the clause-initial auxiliary is a Pan-Mayan category, practically all of the extant formatives of this category emerge at the time of the first colonial grammarians. With the exception of the auxiliary described in §4.9, none of the incipient auxiliaries made its way into their conjugation paradigms; instead, they throw those that they are aware of into the basket of particles. They do, however, use them in their examples.

The following subsections will pursue the grammaticalization of the subset of the tense/aspect/mood auxiliaries of Yucatec Maya shown in Table 6.

This is less than half of the auxiliaries actually in use. Among the ones missing from Table 6 are three past time auxiliaries (recent [illustrated by (5)], relative and remote past), the obligative, potential and volitive moods illustrated below in (23) and a commissive or assurative future. For a subset of these, the origin is unknown. None of the auxiliaries to be discussed here triggers the final ref-

---

[16] An extensive list of relevant markers appears in Briceño Chel 2006: ch. 1.2f.

Christian Lehmann

Table 6: Some Yucatec tense/aspect/mood auxiliaries

form	function	status conditioned
t-/h-	perfective	completive
k-	imperfective	incompletive
táan ts'o'k yan bíin	progressive terminative debitive/future predictive future	subjunctive
bin … ka'h	immediate future	incompletive/subjunctive

erential clitic mentioned in §4.2, so it will be left out of consideration. The last column of Table 6 indicates the status that the auxiliaries trigger on the full verb. By this criterion, there are four structural subclasses of auxiliaries and four different auxiliary constructions, each illustrated by one example in (13).

(13)   Modern Yucatec Maya

    a.   *h   lúub-en*
        PFV fall(CMPL)-B.1.SG
        'I fell'

    b.   *k=in   lúub-ul*
        IPFV=A.1.SG fall-INCMPL
        'I fall'

    c.   *bíin lúub-uk-en*
        FUT fall-SUBJ-B.1.SG
        'I will fall'

    d.   *bin   =in   ka'h lúub-ul*
        IMM.FUT =A.1.SG do   fall-INCMPL
        'I am going to fall'

From this presentation, it appears that the categories in question are coded twice, both by the introductory auxiliary and by the status morpheme. The question naturally arises why each auxiliary goes with a different status. This problem will be analyzed in the following subsections. We will see that all the auxiliation

constructions come about by grammaticalization, but that they originate from different sources.

Another difference between the statuses strikes the eye: Some of them have the intransitive subject represented by a Set A index, while others have it represented by a Set B index. This is the alignment split already mentioned in §3. Although it is not the main object of the ensuing analyses, these will nevertheless contribute to its understanding.

An item of methodology in the analysis of the grammaticalization of these auxiliaries is to be introduced here. At the point when an item is recruited to fill the clause-initial syntactic position, it is a word or even a phrase. Continuing grammaticalization then reduces auxiliaries to bound morphemes (illustrated by (13a) and (13b). There are two tests for the structural status of an auxiliary. First, as in many languages, the answer to a polar interrogative in Mayan involves repeating the main predicate with positive or negative polarity. From this we can derive a test to determine the main predicate of a sentence. In principle, in a configuration like Figure 8, either the auxiliary or the finite full verb may be the main predicate. The auxiliary, however, can be the main predicate only if it is a word. As we shall see, at the beginning of the process, the auxiliary does indeed constitute the answer to a polar question, while with advanced grammaticalization, this is no longer possible, and a short version of the verbal clause appears instead. The second test on the status of the auxiliary involves the placement of enclitic particles. Some of them occupy Wackernagel's position. They may therefore immediately follow the auxiliary if this is a word; and otherwise they must follow the full verb. One might think that the Set A indexes, which are enclitic to the auxiliary, already provide this test. However, these coalesce with the auxiliary once this forfeits its word status and therefore become useless for the test.

4.6 Auxiliation based on modification: from hodiernal past to perfective

As explained in §4.4 and illustrated by (4), the Colonial Yucatec Maya completive status is the only one that a simple independent declarative clause may be based on (i.e. without the need for an auxiliary).[17] This means, at the same time, that such clauses have little marking in comparison with all other tense/aspect/mood categories appearing in independent sentences. Moreover, the completive has zero allomorphs in several contexts. These may be the result of a phonological

---

[17] Of course, imperative sentences lack an auxiliary, too.

Christian Lehmann

process, viz. syncope of the vowels appearing in the completive line of Table 4 if this suffix is followed by a vowel; or else the overt allomorphs may be grammatically restricted to the position *in pausa*.[18] Thus, the transitive completive suffix of (14) and (20b) would be zero in informal speech (as it would be in a Modern Yucatec Maya version of these examples); and likewise the intransitive completive suffix appearing in (24) would normally be zero, as it is in (42) from Colonial Yucatec Maya, in (13a) from Modern Yucatec Maya and in (15).

(14) Colonial Yucatec Maya
      u chabtahon Dios
      u   ch'ab-t-ah-o'n   dios
      A.3 create-TRR-CMPL-B.1.PL god
      'god created us' (Motul s.v. *chab.tah.t*)

(15) Itzá
      Ka'   lub'(-ih)   ah tikin che'-eh   …
      then fall-CMPL(B.3.SG) M dry wood-TOP
      'Then the dry tree fell …' (Hofling 1991, 12:30)

Anyway, the result is that many completive verbal complexes occurring in texts reduce to verb stems provided with indexes. One might expect that such a formally weak category is ripe for reinforcement or renewal. This expectation will be only partially fulfilled.

In Colonial Yucatec Maya, the completive clause can be marked for hodiernal completive.[19] This is achieved by the particle *ti'* 'there' (or its prevocalic bound allomorph *t-*), which may start out in the Focus position of Figure 7, but anyhow ends up in the auxiliary position. (16) shows the simple plain completive for an intransitive (#a) and a transitive (#b) verb. The two parts form minimal pairs with the #a and #b sentences of (17), which show the hodiernal completive.

(16) Colonial Yucatec Maya

      a. Bini Fiscal ti yotoch ku,

---

[18] The completive endings are absent before a following vowel in Lacandón, too. Coon (2010: §3.3) reports similar facts about Ch'ol.

[19] It is hodiernal past according to Coronel 1998a: 41f and San Buenaventura 1684: 35r, although in Smailus 1989: 41 it is characterized as remote or anterior past. The treatment in Coronel is part of the section on dependent status. The first examples of hodiernal past in plain status are in San Buenaventura l.c.

8 Grammaticalization of tense/aspect/mood marking in Yucatec Maya

> *bin-ih*   *fiscal*   *ti' y-otoch*   *k'uh*
> go-CMPL(B.3.SG) inspector LOC A.3-house god
>
> 'The inspector went to the church'

b. ca vhaɔah palalob

> *káa* *=u*   *hats'-ah*   *paal-alo'b*
> CONJ =A.3 beat-CMPL child-PL
>
> 'and beat the children' (San Buenaventura 1684: 23r-v)

(17)   Colonial Yucatec Maya

a. ti bini padre

> *ti'*   *bin-ih*   *padre*
> HOD go-CMPL(B.3.SG) father
>
> 'the father (reverend) went today / has gone'

b. tin haɔah paal

> *t=in*   *hats'-ah*   *paal*
> HOD=A.1.SG beat-CMPL child
>
> 'I beat the child today / have beaten the child' (San Buenaventura
> 1684: 35r)

Two facts should be noted: First, the *ti'* functioning as auxiliary here is based on the word *ti'*, which is syntactically ambiguous between an adverb and a preposition. The adverb is a deictically neutral local demonstrative meaning 'there'. The preposition *ti'* LOC appears in (16a) and is seen to subordinate a nominalized verbal complex in (6a) and (8) (§3). The word occurs in both of these functions in (58) below. While the preposition governs the constituent following it and therefore presupposes dependent status on it if it is based on a verbal construction, the *ti'* presently at stake does not do this. The completive morph in the verbal clause core remains unaffected by the addition of the auxiliary in clause-initial position. Consequently, this auxiliary is based on the adverb, not on the preposition. The semantic shift from 'there' to HODIERNAL is obviously a metaphor from space to time. Second, the auxiliary is the same for intransitive and transitive verbs.[20]

The specification of hodiernal past is possible in dependent status, too:[21] the #a sentence of (18) illustrates simple completive, the #b sentence is its hodiernal

---

[20] In Ch'ol, the perfective auxiliary is *tsa'* (shortened to *tyi*) both for transitive and intransitive verbs.

[21] San Buenaventura 1684: 17r contends that the hodiernal past may trigger dependent status, and gives two examples of it. These are probably due to conditions as obtain in (18b).

Christian Lehmann

counterpart. Here, too, the completive morph is the same in both cases.[22]

(18) Colonial Yucatec Maya

    a. bal v chun a hásci?
      ba'l =u   chuun =a hats'-k-ih
      what =A.3 ground =A.2 beat-DEP-CMPL(B.3.SG)

      'why did you beat her?'

    b. bal v chun ta hásci?
      ba'l =u   chuun t=a   hats'-k-ih
      what =A.3 ground HOD=A.2 beat-DEP-CMPL(B.3.SG)

      'why have you beaten her' (Coronel 1998a: 42)

The hodiernal completive is already highly grammaticalized in Colonial Yucatec Maya.[23] Already in Coronel 1998a, some completive examples introduced by *ti'* are translated as simple past. For instance, (19) is translated as "Quien vino?"

(19) Colonial Yucatec Maya
    Macx ti tali?
    makx ti'   taal-ih
    who   HOD come-CMPL(B.3.SG)

    'Who has come?' (Coronel 1998a: 48)

In Beltrán de Santa Rosa's (1746) examples – e.g. §§264f *(t) luben* – the completive aspect appears variously with and without the aspect auxiliary *t-*, with the same Spanish translation *caí* 'I fell' and no comment on any semantic difference. In §36, he admits that, in front of intransitive verbs, the *t* is "semipronunciada", and establishes the variation taken up below. Apparently, the hodiernal component has disappeared, and what we now have is a perfective auxiliary, reduced to the phoneme *t*, as in (63) below, and therefore regularly univerbated with the following enclitic Set A index, as evidenced by (17b) and (18b). In Modern Yucatec, the perfective auxiliary has become obligatory with transitive verbs in completive status.

---

[22] Coronel (1998a: 41) postulates a contrast between dependent status suffixes for simple and hodiernal completive; but this finds no support elsewhere.

[23] In translating it into English, one has the choice of either rendering the specific semantics and consequently using *today* or else rendering the degree of grammaticity and thus using the perfect.

As for the tests for word status of this auxiliary, it cannot be host to an enclitic particle and cannot constitute the answer to a polar question. The latter may be inferred from (20), where the answer has to contain the full verb.

(20) Colonial Yucatec Maya

    a. ti kamchijnech ua. l. ta kamah ua a chij.
        ti'   k'am-chi'-n-ech   wáa o: t=a   k'amah wáa =a
        HOD get-mouth-CMPL-B.2.SG INT or: HOD=A.2 get-CMPL INT =A.2
        chi'
        mouth

        'Have you had breakfast?'

    b. Ma tin kamah in chi.
        ma t=in   k'am-ah =in   chi'
        NEG HOD=A.1.SG get-CMPL =A.1.SG mouth
        l. ma ti kamchijnen.
        o: ma' ti'   k'am-chi'-n-en
        or: NEG HOD get-mouth-CMPL-B.1.SG

        'I have not had breakfast' (Motul s.v. *kamchij*)

If *ti'* did start out in the Focus position of Figure 7, anyhow it has lost focus function by the start of the documented history of Yucatec Maya, witness such examples as (18b), where it follows the focus constituent. This is, then, the only auxiliary which has already lost word status at the stage of Colonial Yucatec and become a bound morpheme.

Intransitive completive verbs get a Set B index suffixed, as seen, *i.a.*, in (19). The monophonematic auxiliary therefore hits directly on the verb, which may start with a consonant, as in (63). Yucatec has a phonological rule which converts /t/ into /h/ in front of /t/. An extended version of this rule may have applied to the perfective auxiliary. At any rate, this auxiliary has an allomorph *h* with intransitive verbs. A preconsonantal /h/, however, generally disappears in Yucatecan. The *h* to be seen in (13a) is optional both in speaking and in writing, but is mostly absent, as it is in (15) and (16a). One may speculate that what manifests itself in such cases is an uninterrupted continuation of the plain completive of Colonial Yucatec Maya. This may be hard to settle. At any rate, since the hodiernal feature present at the beginning disappears, the result of the entire grammaticalization process is a weak reinforcement of the inherited completive status.

The picture of the Yucatecan languages with regard to this auxiliary is heterogeneous. Mopán shows no trace of a perfective auxiliary, which may reflect

Christian Lehmann

the original situation illustrated by (16). Lacandón has independent declarative clauses in completive status with and without an auxiliary. The latter is illustrated by (21) (from the epic style).

(21) Lacandón

    K=u   yen-s-ik   =u   yok lu'm Hachäkyum y-a'l-ah:
    IPFV=A.3 lower-CAUS-INCMPL =A.3 foot earth Hachäkyum A.3-say-CMPL

    'When Hachakyum set his foot onto the land, he said:' (Bruce S. 1968: 111
    ~ 1974: 19)

No process is known by which the perfective aspect auxiliary would reduce to zero in such a context. Consequently, this may be a functional opposition like the one illustrated by (16)f. In Itzá, the completive only appears to be used with the perfective auxiliary. In both of these latter languages, the distribution of the allomorphs is essentially the same as in Yucatec, except that the allomorph for intransitive verbs is always zero.[24]

The perfective is the only tense/aspect/mood auxiliary of the Yucatecan branch that cooccurs with completive status. The internal syntax of the hodiernal completive construction which is its source differs from all the other auxiliary constructions. The clause core does not depend on the auxiliary, but is, instead, modified by it. There are, of course, many more adverbs which occupy the focus position of Figure 7 and which, being mere modifiers, do not trigger any changes on the verb. However, in a language whose syntax is heavily based on government, a modifying construction is not a productive source for the grammaticalization of auxiliaries. The perfective remains a loner as regards both the source of the auxiliary and the status conditioned (or rather, conserved) by it on the verb. However, as we shall see, the more recent grammaticalization paths converge with it into a common paradigm.

4.7 Auxiliation based on complementation

4.7.1 Basics

Given that any dependents follow the verb, the subordinate clause follows the main clause. Of importance for complex syntax and especially for auxiliation is a kind of complex construction consisting of a main clause core and a complement

---

[24] Lacandón has a subordinator combining with completive aspect, viz. *kahin* 'when' (Bruce S. 1968: 100), corresponding to Yucatec (*le*) *ka'h*. While the Yucatec subordinator combines with the perfective auxiliary, the Lacandón one apparently does not.

clause core. The main predicate may be a nominal or verbal one. It is in any case monovalent and therefore has no dependents beside the complement clause. The latter functions as the subject of a verbal, and as the ("possessive") complement of a nominal main predicate. This presupposes its nominalization, and therefore it is in incompletive dependent status. Given the categorial polymorphy of the main predicate, this is simply categorized by its destination, viz. as an auxiliary to come, in Figure 9. This is construction #d of the set enumerated in §4.5 which shares a syntactic slot in front of the clause core. It is illustrated by (22).
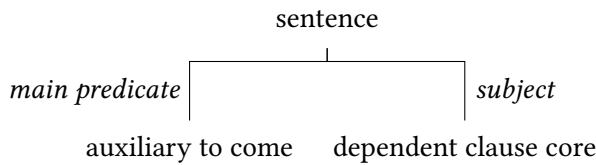
Figure 9: Subject complementation

(22) Colonial Yucatec Maya
      çebhi in canic maya than
      séeb-h-ih   =in   kan-ik   maaya t'aan
      fast=COP-CMPL =A.1.SG learn-INCMPL Maya   speech
      'I learnt Maya quickly' (lit.: 'it was quick that I learnt Maya') (Coronel
      1998a: 52)

From an SAE point of view, the full verb in the dependent clause core may appear to be the main predicate, which several SAE languages would modify by such peripheral concepts as the fastness of (22). A language like Maya, generally averse to modification, prefers the alternative of having the peripheral predicate govern the central predication (cf. Lehmann 1990 for this typological relationship). (23) illustrates the construction with modal verboids.

(23) Colonial Yucatec Maya

      a. v nah a benél
          u   nah   =a   ben-el
          A.3 decorum[25] =A.2 go-INCMPL
          'you ought to go' (Coronel 1998a: 69)

      b. Vchuc inbeelticlo

Christian Lehmann

> *uuchuk* =*in*   *beelt-ik*   =*lo'*
> possible =A.1.SG make-INCMPL =R2
> 'I can do that' (San Buenaventura 1684: 18v)

c. tac in xee
> *taak*   =*in*   *xeeh*
> prompted =A.1.SG vomit\INTROV(INCMPL)
> 'I have/want to vomit' (Beltrán de Santa Rosa 1746: §299, p.146)

As already indicated in §4.5, the complement construction resembles the cleft construction in having the main constituent in the same clause-initial position. An important difference between the two constructions consists in the fact that the subordinate clause of the former is just a nominalized clause. Its status marking is the incomplete dependent status, with non-past reference. The extrafocal clause, instead, may be in any dependent status and thus have any time reference.

As the following subsections will show, this construction is the model for a number of auxiliaries. The clause-initial slot attracts not only intransitive verbs, but also verboids, nouns and denominal adverbs. The construction, however, remains essentially the same: in all the constructions of §4.7, the clause core depends on the initial element.

4.7.2 From habitual to imperfective aspect

The inherited imperfective was renewed in Colonial Yucatec Maya.[26] At the beginning of this process, there is a set of words, apparently denominal in origin, which compete for the auxiliary position. Three of these appear in (24), listed as synonymous in the colonial grammar. The first is *lic(il)*, which has a variant *lac* and must be a root with the meaning 'this time span', although it is no longer found in the texts as such. The second of these auxiliaries is *tamuk*, a preposition and conjunction meaning 'during, while'. The third is *ualac* 'this time'. Both *lik* and *walak* survive in present-day Yucatec in a form adverbialized by the suffix *-il*.[27]

---

[25] lit. 'what befits you / your obligation', Span. *conviene*

[26] All Mayan languages have an imperfective auxiliary, but the forms are very different. For instance, Ch'ol has *muk'*, shortened to *mi*; Q'eqchi' has *nak-*; and so on. S. Vinogradov 2014.

[27] The form *licil* is treated extensively in Coronel 1998a, and on p. 46 he does assign it a habitual meaning. Otherwise, *licil* subordinates a clause similar in function to an oblique relative clause. Modern successors are Yucatec *ka'likil* 'at the time, while' and Itzá *kil* 'when' (Hofling 1991: 26). Acatec Maya has *chi* < *ki*.

(24) Colonial Yucatec Maya
cimçabi in yum
kim-s-a'b-ih   *=in*   *yuum*
die-CAUS-PASS-CMPL(B.3.SG) =A.1.SG master/father
tilic / tamuk / ti válac v hanál
ti' lik   */ tamuk' / ti' walak*   *=u han-al*
LOC this.span / while   / LOC this.time =A.3 eat-INCMPL
'my father was killed while eating' (Coronel 1998a: 57)

In Yucatec, the competition among the three formatives will be won by *lic*. The preposition *ti* subordinating it can already be omitted, as in (25).

(25) Colonial Yucatec Maya
lic u dzocol a hanal ca tacech uaye
lik   *=u ts'o'kol*   *=a han-al*   *káa tal-ak-ech*   *way=e'*
span =A.3 end-INCMPL =A.2 eat-INCMPL CONJ come-SUBJ-B.2.SG here=R3
'when you have eaten, you should come here' (Motul s.v. *ca₆*)

The clause introduced by *lic* may also be independent; then the originally temporal construction may have a habitual sense (cf. Coronel 1998a: 67), clearly visible in (26).

(26) Colonial Yucatec Maya
lic in uenel tamuk in hanal
lik   *=in*   *wen-el*   *tamuk' =in*   *han-al*
HAB =A.1.SG sleep-INCMPL while   =A.1.SG eat-INCMPL
'I usually fall asleep while eating' (Motul s.v. *lic₂*)

By further grammaticalization, the morpheme functions as a mere imperfective auxiliary, as in (27).

(27) Colonial Yucatec Maya
lic bin a haɔic a paalil tu men u tuz. –
lik=bin   *=a hats'-ik*   *=a paal-il*   *tumen =u tuus*
IPFV=QUOT =A.2 beat-INCMPL =A.2 child-REL because =A.3 lie\INTROV
'They say you (habitually) beat your boy because he lies.'
lic. lici.
lik(-ih)
IPFV-CFP
'Yes.' (Motul s.v. *lici lic*)

It may be noted that the two occurrences of the particle in (27) fulfill the conditions of the two tests for word status introduced in §4.5: the particle is, at this stage, syntactically independent. However, there already exists a shortened variant *c(i)*, apparently in free variation, as in the dialogue of (28):

(28)   Colonial Yucatec Maya

    a. bal ca uoktic?
       ba'l   k=a   wook'-t-ik
       what IPFV=A.2 weep-TRR-INCMPL
       'What are you crying for?'

    b. in kéban lic uoktic.
       in   k'eban lik   w-ook'-t-ik
       A.1.SG sin   IPFV A.1.SG-weep-TRR-INCMPL
       'It is for my sins that I am crying.' (Coronel 1998a: 67)

One and a half centuries later, *lic* is still found in the same contexts, as shown in (29)f.[28]

(29)   Colonial Yucatec Maya
     tilic   ú   tzicic   Dios Pedroe,   bin ú
     ti'-lik   =u tsik-ik   dios Pedro=e' bíin =u
     LOC-span =A.3 obey-INCMPL god Peter=R3 FUT =A.3
     chuc olt   dzabilah
     chuk-óol-t   ts'abilah
     attain-mind-TRR(SUBJ) grace
     'as long as Peter obeys god, he will attain grace' (Beltrán de Santa Rosa 1746: §261)

(30)   Colonial Yucatec Maya
     Lic ua ú hanal kohane? – Lic.
     lik   wáah   =u   han-al k'oha'n=e' lik
     IPFV INT =A.3 eat-INCMPL sick=R3 IPFV
     'Does the sick person eat? – He does.' (Beltrán de Santa Rosa 1746: §299, p.140)

---

[28] Beltrán de Santa Rosa (1746: §299, p.140) also mentions *liclili* (*likil-ili'*) with the meaning 'customarily, so it is always', which is a reinforcement of the same particle by the identifying suffix –*ili'*.

As (30) proves, at this stage, *lic* still stands both of the tests of syntactic independence. However, the status of its shortened variant *c(i)*, 'very common' according to Beltrán de Santa Rosa 1746: §101, is already ambivalent.[29] It can still serve as host to a following enclitic, as in the #a version of the variants offered in (31).

(31) Colonial Yucatec Maya

    a. ci bin in yacuntic
        ki   *bin*   =*in*   *yáakunt-ik*
        IPFV QUOT =A.1.SG love-DEP

    b. cin yacuntic bin
        k=in   *yáakunt-ik bin*
        IPFV=A.1.SG love-DEP   QUOT

    'it is said that I love him' (Beltrán de Santa Rosa 1746: §246)

On the other hand, the particle already optionally univerbates with the enclitic A index, as evidenced by the #b version (separate combinations of *ci in/a/u* in Beltrán de Santa Rosa 1746: §131. Beltrán uses the reduced auxiliary *c(i)* in his own examples when aspect is not at stake, thus, in order to choose unmarked aspect (as in (31) *et pass.*). This is already today's situation: The auxiliary only survives in its one-phoneme form *k*, obligatorily univerbates with the Set A index and carries aspectual information only in contrast with more specific auxiliaries.

Thus, the imperfective auxiliary becomes a bound monophonematic form just like the older perfective auxiliary seen in §4.6. The opposition between perfective and imperfective aspect emerges as a minimal one both in formal and in functional terms. It becomes the core of the extensive TAM auxiliary paradigm indicated in Table 6.

We come to the imperfective auxiliaries of the other Yucatecan languages. Both in Itzá and in Lacandón, imperfective aspect is marked by the same formative *k* as in Yucatec.[30] However, Lacandón shows more variation. On the one hand, the formative is optional (Bruce S. 1968: 62), imperfective aspect then being marked only by the incompletive status suffix, as in (32). Especially in Chan K'in Viejo's terse epic style, an incompletive verbal complex often constitutes an independent sentence, as in (33).

---

[29] Beltrán de Santa Rosa dedicates a section (95) to *lic*(*il*), attributing a habitual function to it, and another section (101) to *ci*, attributing present tense function to it, without noting any connection between the two.

[30] Its analysis as a future marker in Bruce S. 1968: 61 must be due to a confusion with the future subordinator *k*(*en*).

Christian Lehmann

(32) Lacandón
K'ayyum =u häts'-ik   Cham-Bol
K'ayyum =A.3 beat-INCMPL Chan-Bor
'Kayum beats Chan Bor' (Bruce S. 1968: 105)

(33) Lacandón
In   want-ik-ech   Yum-eh.
A.1.SG help-INCMPL-B.2.SG lord-VOC
'I (will) help you, my lord.' (Bruce S. 1968: 26)

The most plausible analysis of this construction is that the auxiliary has been reduced to zero.[31] This is, then, an example of complete grammaticalization within half a millennium.

On the other hand, there is a formative *k(ah)* which functions as a temporal conjunction. It may be illustrated by (21), repeated here as (34).

(34) Lacandón
K=u   yen-s-ik   =u yok lu'm Hachäkyum y-a'l-ah:
IPFV=A.3 lower-CAUS-INCMPL =A.3 foot earth Hachäkyum A.3-say-CMPL
'When Hachakyum set his foot onto the land, he said:' (Bruce S. 1968: 111
~ 1974: 19)

The initial *k* is glossed as 'imperfective'. It might as well be glossed as 'when'.[32] The Yucatecan languages have a rather large set of subordinating formatives which start with or at least contain a /k/. Occupying the position indicated in Figure 7 of §4.5, some of them allow a following auxiliary. Recall that the Colonial Yucatec Maya formative *lik*(*il*), which yields the Yucatec imperfective auxiliary, is first mostly found in temporal clauses. The exact relationship between the imperfective auxiliaries and these conjunctions remains to be sorted out.

In Mopán, the alternate auxiliary *walak* was chosen, which appears in (35).

(35) Mopán
walak =ti   ad-ik
HAB   =A.1.PL say-INCMPL
'we always say it' (Danziger 2011: 129)

---

[31] An alternative, and less plausible, account would be to assume that Lacandón uses the nominalized constructions of §4.3 as independent sentences, in which case the change would instantiate insubordination. Note that this is not analogous to the Lacandón use of the completive without auxiliary, discussed in §4.6, since the completive construction at its origin was independent without an auxiliary.

[32] This is actually the gloss provided by Bruce S. l.c.

As may be seen, this is less grammaticalized, both functionally and formally, than its original competitors in the sister languages.

4.7.3 Progressive aspect

The progressive itself is a Proto-Mayan category. In Colonial Yucatec Maya, it is based on the relational noun *tan* (*táan*),[33] illustrated in (36)f in its lexical meaning 'front, middle'.

(36)   Colonial Yucatec Maya
       tan cah
       táan   kah
       middle village
       '(in) the village center' (Beltrán de Santa Rosa 1746: §299, p.147)

(37)   Colonial Yucatec Maya
       tutan Dios
       t=u   táan dios
       LOC=A.3 front god
       'in front of god' (San Buenaventura 1684: 39v)

(37) shows the regular syntactic construction which is natural for a noun designating a spatial region, viz. preceded by a possessive Set A clitic[34] and governed by the default preposition *ti'* LOC. The same configuration is also at the source of its aspectual use. The full form *tután* is only mentioned in Coronel 1998a: 47, but not illustrated in the sources. The earliest evidence lacks the preposition. (38)f illustrate the incipient progressive function for intransitive and transitive verbs, resp. (38) is obviously a variant of (24).

(38)   Colonial Yucatec Maya
       vtán v hanál in yum,
       u   táan =u han-al   =in   yuum
       A.3 middle =A.3 eat-INCMPL =A.1.SG master/father
       ca cimçabi

---

[33] The progressive function of this morpheme may be inherited from Proto-Maya; some languages, including Kaqchiquel, have plausible cognates.

[34] The only Set A index ever attested in this construction is *u* A.3. This leads to the interpretation made explicit in the literal translation of (38) and to the gloss 'middle'. If the clitic could have been of first person, then the other meaning of *táan*, viz. 'front', would appear to underlie the construction: 'in front of me/us, P is happening'.

Christian Lehmann

> *káa   kim-s-a'b-ih*
> CONJ die-CAUS-PASS-CMPL(B.3.SG)
>
> 'my father was in the middle of eating when he was killed' or: 'while my father was eating, he was killed' (Coronel 1998a: 57)

(39) Colonial Yucatec Maya
Vtan incambecic paal,
u   táan   =in   kambes-ik   paal
A.3 middle =A.1.SG teach-DEP.INCMPL child
ca xolhi tu pix.
káa   xol-hih   t=u   píix
CONJ kneel-CMPL(B.3.SG) LOC=A.3 knee

'While I was teaching the child, he knelt down.' (San Buenaventura 1684: 9Br)

The original construction with the subordinating *ti'* and its further evolution are, at any rate, completely analogous to the imperfective *ti' lik* seen in (24): It follows the pattern of Figure 9, where the full verb of the complement clause is in the incompletive dependent status. Initially, the new auxiliary is typically used in complex sentences, where the progressive clause provides the background for the event of the main clause, as clearly shown by (38)f. However, and again like the imperfective, the progressive also appears in monoclausal sentences as (40)f. (41) features, already at Coronel's time, a further reduced form of the auxiliary, where the original possessive clitic preceding *táan* is no longer there.[35]

(40) Colonial Yucatec Maya
U tan in beeltic
u   táan   =in   beel-t-ik
A.3 PROG =A.1.SG make-TRR-DEP.INCMPL

'I am (in the middle of) doing it' (San Buenaventura 1684: 37r)

(41) Colonial Yucatec Maya
ma tan a túbul ten
ma' táan   =a   tu'b-ul   ten
NEG PROG =A.2 escape-INCMPL me

'I am not going to forget you' (Coronel 1998a: 34)

---

[35] Since Colonial Yucatec Maya, there has been a complex form *ma'táan* of the negator *ma'*, which according to Coronel (1998a: 83) triggers the incompletive of intransitive and the subjunctive of transitive verbs. It is certainly present in (41). It is not clear whether it contains the morpheme *táan* presently at stake.

Beltrán de Santa Rosa (1746: §261) includes *utan* in the list of particles adopted from his predecessors, but in his own examples he only uses the reduced form *tan*. Seeking to render the Spanish progressive ("gerundio") in Maya, he offers, among other alternatives, the pair of examples in (42), which illustrates, at the same time, the morphological correlates of the transitivity contrast:

(42)   Colonial Yucatec Maya

    a.   tan in tzeec, ca lub kuna
        táan =in   tse'k   káa lúub   k'u-nah
        PROG =A.1.SG preach(INCMPL) CONJ fall(CMPL) god-house

        'I was preaching, there the church collapsed'

    b.   tan in tzeectic ú than Dios tiob,
        táan =in   tse'k-t-ik   =u   t'aan dios ti'-o'b
        PROG =A.1.SG preach-TRR-INCMPL =A.3 word god LOC-3.PL
        ca cim Joan
        káa kim   Juan
        CONJ die(CMPL) John

        'I was preaching god's word to them, there John died' (Beltrán de Santa Rosa 1746: §262)

As may be seen, this is now just a progressive aspect. (43) illustrates the test on susceptibility of serving as the host to a clitic particle, with positive result for contemporary Yucatec Maya.

(43)   Modern Yucatec Maya
    Táan wáah =a   bin?
    PROG INT =A.2 go(INCMPL)
    'Are you going (leaving)?' (Hnazario_406)

In its further development, and again in analogy with the development of the imperfective auxiliary as illustrated by (31b) above, the progressive auxiliary coalesces with the Set A index which regularly follows it. The full form of the auxiliary survives essentially in writing and, in the oral mode, in cases like (43). The coalescence is a process in two phases. At first, the product of the merger of *táan* with the three singular indexes *in, a, u* is *tíin, táan, túun*, as illustrated by (44).

(44)   Modern Yucatec Maya

Christian Lehmann

> *Túun   tsikbal.*
> PROG:A.3 tell(INCMPL)
> 'He was talking.' (Monforte, Dzul & Gutiérrez Bravo 2011: 48)

This is, however, just a transitional stage rarely represented in writing. In the end, these forms are shortened to *tin, ta, tu* (cf. Briceño Chel 2006: 24f), as in (45).

(45)   Modern Yucatec Maya
       T=u   sáas-tal
       PROG=A3 dawn-FIENT.INCMPL

       káa   h   téek   líik'   y-ich   hun-túul   le   peek'=o'
       CONJ PFV for.a.moment rise(CMPL) A.3-eye one-CL.AN DEM dog=R2
       'It was dawning when one of the dogs suddenly rose his glance'
       (hts'oon_310.1)

In the syntactic configuration illustrated by (45), the progressive clause specifies a situation holding in the background simultaneously with the event described by the following clause. This is functionally equivalent with the combination described in §4.3 (cf. 6b and 11f), where a nominalized clause subordinated by *ti'* serves as background information for the main clause. In fact, since the products of the merger of the preposition and of the progressive auxiliary with the following Set A index are homonymous, the two constructions are not easily distinguished. It may be assumed that the (much older) model of the nominalized construction played a role in the rather radical reduction of the auxiliary complex.

By the same token, the reduced variant of the progressive auxiliary becomes homonymous with the perfective auxiliary. The two aspects, however, do not thereby become homonymous, since the progressive conditions incompletive status, while the perfective conditions completive status; and these two are distinct for all verbs (cf. Lehmann 2014, §3.4.2). This convergence of two aspectual auxiliaries constitutes an important contribution to the maintenance of the status category, which otherwise might have been grammaticalized to zero (cf. §4.4).

It remains to note that the progressive with *tan* is a Pan-Yucatecan construction; s. Bruce S. 1968: 93, 97 for Lacandón, Hofling 1991: 30 for Itzá and Danziger 2011: 125 for Mopán. In Itzá and Lacandón, the reduced forms are as the above-mentioned intermediate forms of Yucatec (44). The full form *tan in wilik* is in free variation with the reduced form of (46) (Bruce S. 1968: 61, 97).

(46) Lacandón
      tin wilik
      tan=in   wil-ik
      PROG=A.1.SG see-INCMPL
      'I am seeing it' (Bruce S. 1968: 34)

Thus, the progressive auxiliary becomes a bound monophonematic form just like the imperfective auxiliary seen in §4.7.2.

4.7.4 Terminative aspect

The first constituent of Figure 9 is filled by a noun in the cases reviewed in the two preceding sections. This is, however, not the most fertile grammaticalization path for auxiliaries. Apart from modal verboids, the most important subclass of intransitive predicators to fill this position are phase verbs. The central Yucatec phase verbs are *ho'p'* 'start' and a set of verbs including *ch'en*, *ts'o'k*, *haw*, *nik* all meaning 'end'. They are normally impersonal (s. already Coronel 1998a: 34f). Personal use is possible with a few of them, but does not generate auxiliaries. In the impersonal construction, actancy is coded on the dependent verb; with some marginal exceptions, there is in Yucatecan no "raising".[36] (47) and (48) illustrate the construction for *ho'p'* 'start' and *ts'o'k* 'finish', respectively. Whether or not the main clause is clefted (#a vs. #b examples), the dependent verb is in the incompletive dependent status.[37]

(47) Colonial Yucatec Maya

      a. hoppi in beeltic
          ho'p'-ih   =in   beel-t-ik
          start-CMPL(B.3.SG) =A.1.SG make-TRR-DEP.INCMPL

          'I have begun to do it' (Coronel 1998a: 53)

      b. çamal v hoppol in ɔibtic
          sáamal   =u   ho'p'-ol   =in   ts'íib-t-ik
          tomorrow =A.3 start-INCMPL =A.1.SG write-TRR-DEP.INCMPL

          'tomorrow I will start writing it' (Coronel 1998a: 35)

---

[36] Ch'ol has the same construction; s. Aulie & Aulie 1998: 239. According to Coon 2010: §5.2 the Ch'ol auxiliaries which trigger incompletive status do allow raising of the absolutive enclitic.

[37] Smailus (1989: 89) claims it to be in the subjunctive. However, although crucial evidence, with an intransitive dependent verb, appears to be rare, Coronel 1998a: 35 does have *maytoh ts'o'kok in menyali'* 'I have not yet finished working', with *menyal* in the incompletive.

Christian Lehmann

(48)   Colonial Yucatec Maya

    a.   ɔoci incanic
       ts'o'k-ih   *=in*   *kan-ik*
       end-CMPL(B.3.SG) =A.1.SG learn-DEP.INCMPL
       'I finished learning / have learnt it' (San Buenaventura 1684: 17r)

    b.   çamal v ɔócol in canic
       sáamal   *=u* *ts'o'k-ol*   *=in*   *kan-ik*
       tomorrow =A.3 end-INCMPL =A.1.SG learn-DEP.INCMPL
       'tomorrow I will finish learning it' (Coronel 1998a: 35)

The phase verb *ts'o'k* 'finish' shown in (48) combines with aspect auxiliaries just like any full lexical verb, e.g. in (25). It continues this life up to the present day. In (49), it regularly goes into the subjunctive required by the construction, and only the translation suggests its auxiliary function.

(49)   Modern Yucatec Maya
    le   *kéen*   *ts'o'k-ok*   *=u* *pa't-al=e'*
    DEM when.IMPF finish-SUBJ =A.3 form\PASS-INCMPL=TOP
    k=u   *ts'a'bal*   *píib*   …
    IMPF=A.3 put/give:INCMPL.PASS underground.oven
    'When they have been formed, they are put into the earth-oven …'
    (chaak_028)

(50)   Modern Yucatec Maya
    beey túun ts'o'k-ol   *=u* *kuxtal le*   *p'us-o'b=o'*
    thus then finish-INCMPL   =A.3 life   DEM hunchback-PL=R2
    'This then was the end of the life of the P'uz.' (chem_ppuzoob_011)

(50) displays a symptom of grammaticalization: the phase verb is in the incompletive, but it lacks both the introductory imperfective auxiliary and the Set A index. This suggests that even in the construction at hand, where the main clause comprises more than just the phase verb, the latter fulfills an auxiliary function, with the form *kuxtal* in its subject not just being an abstract noun, but rather the verbal predicate of the dependent clause core (a case of the zero nominalization described in §4.3).

This grammaticalization process starts in the colonial period. The seventeenth century grammars adduce the phase verbs *ɔoc* 'finish' and *hopp* 'begin' only in order to mention their regular impersonal or personal construction as illustrated

by (25) and (47)f above. It is in the eighteenth century that the ongoing grammaticalization of the third person completive form *ts'o'k*[38] could no longer escape a critical linguist's ear. Beltrán, writing his grammar in Mérida in 1742, observes the expansion of the use of *ts'o'k* as auxiliary in vogue at his time (§§85f), notes that it is a partial competitor to the (firmly established) perfective, quotes some periphrastic forms which are actually in use up to today and opposes violently both to this fashion and to the idea that *ts'o'k* means 'already' (which it does in its function as terminative auxiliary). His verdict is that the perfective is formed without auxiliary or "better" with the auxiliary *t-* (of §4.6 above), while *ts'o'k* means 'finish' and nothing else.

The form of this verb which occupies the clause-initial position, becoming, thus, a component of the grammaticalization path, is the completive form triggered by perfective aspect, as in (51) (where *meyah* – just like *kuxtal* in (50) – can be an intransitive verb with the zero allomorph of the incompletive or an abstract noun).

(51) Modern Yucatec Maya
      (h) ts'o'k   *=in*   *meyah*
      PFV finish(CMPL:B.3.SG) =A.1.SG   work
      'my work ended = I finished working = I have already worked'
      (Briceño Chel 2000b: 84)

In the sequel, the perfective auxiliary is omitted. In fact, by the evidence of (48), the grammaticalization of *ts'o'k* probably started at a time when the completive alone could make a perfective clause. Otherwise, however, the new auxiliary can maintain its full form even in the colloquial style. It passes the two tests on word status up to the present day, as evidenced by (52).

(52) Modern Yucatec Maya

      a. *Ts'o'k*   *wáah =in*   *bo'l-t-ik*   *=in*   *p'aax?*
         finish(CMPL:B.3.SG)   INT =A.1.SG pay-TRR-INCMPL   =A.1.SG   debt
         'Have I paid my debt?'

      b. *Ma' ts'o'k-ok=i'.*
         NEG finish-SUBJ=NEGF
         'No (you haven't).' (hnazario_375f)

---

[38] The grammaticalization of *ho'p'* to an auxiliary will not be described here. Both in Yucatec and in Itzá (Hofling 1991: 105), it is common in narratives and reports to mark a new situation.

Christian Lehmann

There is, however, a reduced form in addition to the full form, although not as widely used as the reduced form of the progressive auxiliary. The auxiliary is then reduced to its initial consonant and coalesces with the Set A clitic, as shown by (53) (cf. Briceño Chel 2000b: 87f).

(53) Modern Yucatec Maya
      ts'=in   *w-a'l-ik*   *te'x*
      TERM=A.1.SG 0-say-INCMPL you.all
      'I have told it to you' (muuch_340)

The terminative is a kind of perfect and therefore in partial competition with the inherited suffixal perfect. They share the semantic component that the situation designated is finished at topic time. Their semantic difference lies in the implication of this fact. The Yucatec perfect implies that the agent has the result of his action at his disposal, while the terminative focuses on the crossing of the end boundary of the situation, which may be counter to expectations.[39]

Like the progressive, *ts'o'k* is a Pan-Yucatecan auxiliary. Its Lacandón form is *ts'ok*;[40] (54) is an example.

(54) Lacandón
      Ts'ok *=u* *me(n)t-i(k)* *k'ax,* …
      TERM =A.3 make-INCMPL woods
      'He had made the woods, …' (Bruce S. 1974: 24)

Likewise in Itzá, *tz'o'k* is used in terminative function, as shown by (55):

(55) Itzá
      Tz'o'k-i(h)   *=u* *man ka'-p'eel*   *k'in,* …
      TERM-CMPL(B.3.SG) =A.3 pass two-CL.INAN day
      'Two days had passed, …' (Hofling 2006, 12:39)

Besides this, Itzá has grammaticalized another phase verb to a terminative auxiliary, viz. the verb *ho'm* (Hofling 1991:25, 65), whose original meaning is 'wane, abate'.

As an aside, it may be mentioned that the phase verb *ts'o'k* in the imperfective aspect is also the grammaticalization source of a paratactic conjunction that is very widely used in the colloquial register of Modern Yucatec Maya, as witnessed by the monotonous repetition in (56).

---

[39] Terminative aspect is incompatible with a temporal adverb in the same clause; s. Briceño Chel 2000b: 82f

[40] According to Bruce S. 1968: 81, 93, 99 the function is immediate past.

(56) Modern Yucatec Maya

    K=u   *ts'o'k-ol=e'*   *k=in*   *p'o'-ik;*
    IPFV=A.3 finish-INCMPL=TOP   IPFV=A.1.SG wash-INCMPL

    'Then I wash it;'

    k=u   *ts'o'k-ol=e'*
    IPFV=A.3 finish-INCMPL=TOP

    k=in   *ts'a'-ik*   *t=eh*   *k'áak'=o'* ...
    IPFV=A.1.SG put/give-INCMPL   LOC=DEM fire=R2

    'then I put it on fire ...' (chakwaah_03f)

The phrase *ku ts'o'kole'* is commonly reduced to *ts'o'(h)le'*, the loss of the auxiliary complex being due to grammaticalization, while the shrinking of the verb form follows regular phonological processes.

4.7.5 From existential via debitive to future tense

The existential predicate in the Yucatecan languages during their entire documented history is the intransitive verboid *yaan*, illustrated by (57).

(57) Colonial Yucatec Maya

    yan cutz

    yaan *kuts*

    EXIST turkey

    'there are turkeys' (Beltrán de Santa Rosa 1746: §199)

Apart from predicating sheer existence, *yaan* is also the locational copula, as in (58).

(58) Colonial Yucatec Maya

    tij yan ti yotoch

    ti'   *yaan ti'*   *y-otoch*

    there EXIST LOC A.3-home

    'there he is at his home' (San Buenaventura 1684: 35v)

Furthermore, the canonical construction coding ascription of possession is obtained by substituting a possessed nominal for the central actant of *yaan*, as in (59).[41]

---

[41] Interestingly, Beltrán de Santa Rosa 1746: §199 makes the not unreasonable claim that the verboid *yaan* lacks the first and second persons in the existential and possessive uses. However, the first example offered by the *Diccionario de Motul* s.v. *yan* features just the second person in the existential use.

Christian Lehmann

(59) Modern Yucatec Maya
      yaan =in   nah-il
      EXIST =A.1.SG house-REL
      'I have got a house' (muuch_274)

Once a nominalized verbal complex is substituted for the possessum of the ascription of possession, a debitive construction results. Just as the possessum is ascribed to its possessor in (59), so the obligation is ascribed to the actor of the nominalized verbal complex in (60).

(60) Modern Yucatec Maya
      ba'l=e'   *yan =a   bo'l-t-ik-en*
      however=TOP DEB =A.2 pay-TRR-INCMPL-B.1.SG
      'however, you must pay me' (hala'ch_084)

This use is not found in the colonial sources and is documented only in the modern Yucatecan languages. In Itzá, the construction is the same as in Yucatec (Hofling 1991: 25). In Lacandón, the dependent clause core is introduced by the subordinator *ti'*, as shown by (61).

(61) Lacandón
      yan ti' =a   kaxt-ik   =u   hel
      DEB LOC =A.2 search-INCMPL =A.3 replacive
      'you have to look for another one' (Bruce S. 1968: 81)

The most recent development, first documented in the 20th century oral register, is a pure future without debitive connotations, as in (62), where the speaker articulates what he thinks will certainly happen.

(62) Modern Yucatec Maya
      yan =u   kaxt-ik-ech   =a   taatah
      DEB =A.3 search-INCMPL-B.2.SG =A.2 father
      'your father will search you' (hnazario_402.1)

This construction is currently ousting the (much older) predictive future (§4.8), which gets pushed back into the formal register.

4.8 Auxiliation based on motion cum purpose: predictive future

The motion-cum-purpose construction is a regular syntactic construction in the Yucatecan branch. It is a complex clause core starting with an oriented motion

verb followed by a verbal clause core in the subjunctive, the latter coding the purpose. As long as nothing precedes the motion verb, the core verb is in plain status subjunctive, as in (63).[42]

(63) Colonial Yucatec Maya
      t binén in cimez uacax
      t   bin-en   =in   kim-es   wakax
      PFV go(CMPL)-B.1.SG =A.1.SG die-CAUS(SUBJ) cow
      'I went to kill cows' (Beltrán de Santa Rosa 1746, §110)

The central verbs of oriented motion ('go', 'come', 'pass') become irregular in their conjugation on their way to Modern Yucatec. Specifically, they lose the *-Vl* suffix which marks their nominalization and would be expected in their incompletive status (see also (81) below). Moreover, the verb *ben* 'go' becomes *bin* in Yucatec, while in the other Yucatecan languages it becomes *bel*. The changed forms appear both with their lexical meaning 'go' and as auxiliaries.

The motion-cum-purpose construction with *bin* as motion verb is grammaticalized to a future in the Yucatecan branch. Coronel (1998a) already calls it "futuro" and provides examples of it. Beltrán de Santa Rosa (1746: §299, p. 128) lists *bin* as 'partícula de futuro', giving examples (64)f for the intransitive and transitive construction, resp. (29 is another example; see Table 4 for the allomorphs).

(64) Colonial Yucatec Maya
      bin bolnacén dzedzetàc
      bíin bo'l-nak-en   ts'e'ts'etak
      FUT pay-SUBJ-B.1.SG little.by.little
      'I shall pay little by little' (Beltrán de Santa Rosa 1746: §299, p. 149)

(65) Colonial Yucatec Maya
      caix u tancoch in hanale,
      kayx   =u   táankoch =in   haanal=e'
      although =A.3 half   =A.1.SG meal=R3
      bin in ziib tech
      bíin =in   síih-ib   tech
      FUT =A.1.SG present-SUBJ(B.3.SG) you
      'although it is half of my meal, I'll give it to you' (Beltrán de Santa Rosa 1746: §299, p.129)

---

[42] If the future clause as introduced by *bin* is an extrafocal clause, as in (67) and (69), the full verb goes into dependent = incomplete status.

Christian Lehmann

The core verb keeps the subjunctive of the source construction.[43] The motion verb complex has been reduced to the root of the motion verb. This becomes impersonal like all the other auxiliaries and, in Yucatec and Lacandón, undergoes an idiosyncratic phonological change: the vowel of the auxiliary *bin* (not of the lexical verb!) is lengthened and gets high tone in Yucatec. This may be due to analogy with the progressive auxiliary *táan*, but may also be regarded as the expression counterpart of the grammatical change. At any rate, the impersonalization and morphological impoverishment of the auxiliary comes under paradigmaticization and may be ascribed to analogical pressure from the older auxiliation constructions analyzed in §4.7.[44] (66) illustrates the construction for both an intransitive and a transitive verb.

(66) Modern Yucatec Maya
      Bíin suu-nak   yéetel bíin =in   wil-eh.
      FUT return-SUBJ(B.3.SG) and   FUT =A.1.SG see-SUBJ(B.3.SG)
      'He will come back and I will see him.' (xipaal_032)

This future construction finds its place in the tense/aspect/mood paradigm at the side of three other futures, viz. the debitive future (§4.7.5), the immediate future (§4.9) and an assurative future not analyzed here. It does not become an immediate future, as so many futures based on the motion-cum-purpose construction do in other languages. Neither does it contrast with the immediate future on the time axis, as can be inferred from examples like (65). Instead, it bears a feature of neutral, objective prediction, which may be related to the impersonality of its auxiliary and which opposes it to the other three futures. Since this semantic component matters less in what is going to happen next, time reference is often to a remote future. But this is only a favorable circumstance, not a condition for the appropriateness of a prediction.

We find the predictive future at an intermediate stage of grammaticalization. On the one hand, the reduction process mentioned above proves that it is grammaticalized to some extent already at the stage of Colonial Yucatec Maya. (67) provides evidence in the same sense, as it shows that the construction is compatible with an additional, preceding focus constituent.

(67) Colonial Yucatec Maya

---

[43] In Modern Yucatec Maya, the motion-cum-purpose construction itself diverges from its source by having the intransitive verb in the incomplete instead of the subjunctive status.

[44] Ch'orti' (a Ch'olan language, thus closely affiliated to Yucatecan) has the same impersonal construction with an etymologically unrelated verb meaning 'go'.

bay bin v cíbic Dios teex
bay bíin =u kib-ik Dios te'x
thus go =A.3 do-DEP.INCMPL god you.PL

'thus will god do with you' (Coronel 1998a: 72 = San Buenaventura 1684: 24r)

On the other hand, the predictive future auxiliary stands the clitic placement test to this day:

(68) Modern Yucatec Maya
bíin wáah p'áat-ak-en hun-p'éel k'iin he'bix-ech=a'
FUT INT stay-SUBJ-B.1.SG one-CL.INAN sun/day ever:how-B.2.SG=R1

'will I become like you one day?' (xipaal_092)

The predictive future construction is, again, Pan-Yucatecan. Lacandón conserves a variant of it which is structurally identical to the motion-cum-purpose construction, to be seen in (69).

(69) Lacandón
way k=u bin p'at-al t=in meyah
here IPFV=A.3 go stay-INCMPL LOC=A.1.SG work

'it will stay here for my work' (Bruce S. 1974: 42)

However, it also has the reduced auxiliary construction like Yucatec, as in (70).

(70) Lacandón
b'ihn a-kihn-s-∅-een
bíin =a kíin-s-en
FUT =A.2 die-CAUS(SUBJ)-B.1.SG

'you will kill me' (Bergqvist 2011: 247)

Itzá again has the full motion-cum-purpose construction with future function, to be seen in (71):

(71) Itzá
way=e' k=in b'el =in pak'-t-eech
here=R3 IPFV=A.1.SG go =A.1.SG wait-TRR(SUBJ)-B.2.SG

'here I'm going to await you' (Hofling 1991, 15:126)

The origin of the predictive future construction is the motion-cum-purpose construction. It differs from the other tense/aspect/mood auxiliaries analyzed in

Christian Lehmann

§4.6f in that the emerging marker – the verb 'go' grammaticalized to a future marker – does not originally occupy the clause-initial position described at the beginning of §4.5 and instead is the remnant of a complete superordinate clause. However, the canonical model for an auxiliary construction is Figure 6: the auxiliary is monomorphematic, impersonal and occupies the clause-initial position. In its grammaticalization, the motion-cum-purpose construction is forced into the Procrustean bed of the verbal clause expanded by an initial position, which is the template for the auxiliary construction. This is, thus, a clear example of grammaticalization guided by analogy.

4.9 Auxiliation based on focused progressive: immediate future

As noted in §4.5, the clause-initial position is a melting-pot for constituents of very different kinds, among them the focus. We now come to an auxiliation strategy originating in a focus construction, more specifically, in a verb-focus construction. From there, we get to the immediate future in two steps: First, on the basis of the verb 'go' in focus, a focused progressive is formed. Second, this strategy applies to the 'go' verb of the motion-cum-purpose construction to form the immediate future of its purpose component.

Putting the lexical main verb of a clause into its focus position requires filling the gap that it leaves in the extrafocal clause by a verb meaning 'do'.[45] For this purpose, Colonial Yucatec Maya used a verb *cib* 'do' which is totally irregular and defective. Table 7 presents the forms adduced in Coronel 1998a: 71f.

Table 7: Partial paradigm of Colonial Yucatec Maya *cib* 'do'

category	form
[fossilized status]	cah
completive	cibah
subjunctive	cib (not *cibib*!)
incompletive dependent	cibic

Already Beltrán de Santa Rosa (1746: §§209f) doubts this paradigm and contends that the verb is defective, being reduced to a "present" form *cah*. The verb is rarely found in a simple transitive clause to code the meaning 'do, make';[46]

---

[45] S. Lehmann 2008, §4.3 for a comprehensive account of the underlying information structure and the Yucatec development.

[46] One of the rare examples is (67) above, featuring dependent incompletive status.

the lexicon offers other verbs with this meaning. Instead, it is used almost exclusively in focus constructions. A relatively straightforward one appears in (72).

(72) Colonial Yucatec Maya
balamil u cah pedro
balam-il =u *ka'h Pedro*
tiger-ADVR =A.3 do Peter
'Peter makes the tiger / Peter is like a tiger' (lit.: 'tiger-like is what Peter does'; Motul s.v. *cah₃*)

At the stage of Colonial Yucatec, the verb is indispensable as a pro-verb in the verb focus construction. The paradigm shown in Table 7 is illustrated by (73).

(73) Colonial Yucatec Maya

  a. hanál v cah
    han-al =u *ka'h*
    eat-INCMPL =A.3 do
    'he is eating'

  b. hanál v cibah
    han-al =u *kib-ah*
    eat-INCMPL =A.3 do-CMPL
    'he was eating'

  c. hanal bin v cib
    han-al *bíin* =u *kib*
    eat-INCMPL go =A.3 do(SUBJ)
    'he is going to eat' (Coronel 1998a: 71; cf. San Buenaventura 1684: 23v)

  d. lúbul tu cibah
    lúub-ul *t=u* *kib-ah*
    fall-INCMPL HOD=A.3 do-CMPL
    'he fell (earlier today)' (Coronel 1998a: 71)

As suggested by the translations of (73)a-c, the same construction functions as a progressive in Colonial Yucatec Maya. As a matter of fact, it figures much more prominently in colonial grammars than the simpler progressive with the auxiliary *táan* (§4.7.3). All of them start their account of the conjugation with the periphrastic construction based on *ka'h*, calling it the "presente". (74) completes the example series with a transitive verb.

Christian Lehmann

(74)　Colonial Yucatec Maya
　　　cámbeçah in cah ti pálalob
　　　kambes-ah 　　　　 *=in* 　　 *ka'h ti'* 　 *paal-alo'b*
　　　teach-INTROV(INCMPL) =A.1.SG do 　 LOC child-PL
　　　'I am teaching the children' (Coronel 1998a: 72)

While all of the examples (72) – (74) are focus constructions, there are a number of peculiarities. First, if these were standard cleft sentences, the pro-verb of the extrafocal clause would have to be in dependent status. While this is hard to know for the irregular forms *ka'h* (73a) and *bíin* (73c), the forms of (73b) and d appear to be forms of the plain status. Second, while any constituent can be focused without its form being thereby affected in any way, a finite verb cannot; it must be nominalized. Therefore, the focused verbs in (73)f show the nominalizing suffixes introduced in §4.3, identical with incompletive (dependent) status. Third, the process is relatively unproblematic with intransitive verbs, as in (73), as their only actant is identical with the subject of *ka'h* and may thus safely be suppressed by the nominalization. Things are more complicated with transitive focused verbs, as in (74). The purpose of the verb-focus construction is to put the verb into focus. Consequently, its dependents remain in the extrafocal clause. Therefore, the verb is detransitivized before it is nominalized. The internal syntax of the extrafocal clause is adapted, too: what was the direct object of the focused verb becomes a prepositional object (Beltrán de Santa Rosa 1746: §172). The verb focus construction is, consequently, marked with plurivalent verbs.

The progressive aspect views what the verb designates as an ongoing situation that the referent of the subject is in. Consequently, the functional locus of the progressive aspect is in intransitive verbs.[47] The verb focus construction is therefore well suited to get grammaticalized into a progressive aspect.[48] The resulting construction may be dubbed **focused progressive** (as in Lehmann 2008). Two symptoms of the grammaticalization of the focused progressive construction in Colonial Yucatec Maya will be mentioned: First, its susceptibility to nominalization by coercion, i.e. by having it depend on the preposition *ti'*, as in (75).

(75)　Colonial Yucatec Maya
　　　ti cimil in cah

---

[47] Evidence for this is provided, *inter alia*, by the documented history of the evolution of the progressive aspect in English and in substandard German; see Lehmann 1991: section 3.2.

[48] The progressive aspect of other languages has a similar origin; cf., e.g., Güldemann 2003 for Bantu.

> *ti' kim-il   =in   ka'h*
> LOC die-INCMPL =A.1.SG do
> 'at/by my being ill' (Coronel 1998a: 58)

Second, since the action feature of the basic meaning of *kib* is lost, it combines even with passive verbs, as in (76):

(76) Colonial Yucatec Maya
      tzicil in cah
      tsi'k-il   =in   ka'h
      obey\PASS-INCMPL =A.1.SG do
      'I am (being) obeyed' (San Buenaventura 1684: 11v)

Modern Yucatec Maya has a verb-focus construction, too, but it is not as central to the conjugation paradigm as the focused progressive appears to be in the grammars of Colonial Yucatec Maya. This has two totally unrelated reasons. The first is that the Colonial Yucatec Maya construction is much more grammaticalized than is the Modern Yucatec Maya verb focus construction, which was renewed with the lexical verb *beet/meent* 'make' (seen in (40) above). The modern counterpart to (73d) would consequently be (77).

(77) Modern Yucatec Maya
      lúub-ul   t=u   meet-ah
      fall-INCMPL   PFV=A.3 make-CMPL
      'fall was what he did' (~ 'all of a sudden, he fell')

The Colonial Yucatec Maya construction is clearly a kind of progressive aspect, which the Modern Yucatec Maya construction is not; it is rather more of a thetic construction fit for all-new-utterances. The second reason for its prominence in the colonial grammars is a methodological one: The category is not nearly as frequent in the texts as it is in the grammars. The explanation is not hard to find: The grammarians needed to fill up the conjugation paradigms presupposed by Latin grammar (Hanks 2010:214f). If one looks for a present tense in Colonial Yucatec Maya, the closest analog would appear to be the imperfective aspect described in §4.7.2. This, however, originates in complex sentences, whereas here an isolated verb form was needed. In a decontextualized sentence reduced to a finite verb, all of the emphasis is on the finite verb. Which provokes a verb-focus construction.

On its way into the modern Yucatecan languages, the pro-verb *cib* is further fossilized; only the form *cah/ka'h* occurs in a couple of contexts. This is ousted

from its function as a pro-verb in verb-focus constructions by the lexical verb *beet/meent* illustrated in (77). *Ka'h* survives in this function only in the formulaic pattern illustrated by (78).

(78) Modern Yucatec Maya
Chéen uk'ul =u ka'h.
only drink:INTROV(INCMPL) =A.3 do
'drinking is all he does / he only drinks (all the time)' (Briceño Chel 1998: 77)

Neither is the focused progressive with *ka'h* further grammaticalized to a plain progressive. As we have seen in §4.7.3, the progressive construction which gets established involves a different auxiliary. Instead, verb focusing is applied to the motion-cum-purpose construction analyzed in §4.8. What is put into focus position is the verb *benel/binel/bin* 'go', while the purpose part of the construction is left behind in the extrafocal clause core. The resultant specific construction is, thus, a merger of the focused progressive with the motion-cum-purpose construction. (79)f illustrate it with an intransitive and transitive full verb, resp.

(79) Colonial Yucatec Maya
benel in cah ti hanal
ben-el =in ka'h ti' han-al
go-INCMPL =A.1.SG do LOC eat-INCMPL
'I am going to eat' (Coronel 1998a: 50)

(80) Colonial Yucatec Maya
Binel in cah incambez palalob.
bin-el =in ka'h =in kan-bes paal-alo'b
go-INCMPL =A.1.SG do =A.1.SG learn-CAUS(SUBJ) child-PL
'I am going to teach the children.' (San Buenaventura 1684: 9Br)

As already observed above (fn. 43), on its way to Modern Yucatec Maya, the motion-cum-purpose construction develops an asymmetry conditioned by the transitivity of the full verb which persists into the focused progressive: a transitive full verb (80) is in the subjunctive motivated by the motion-cum-purpose construction, while the status of an intransitive full verb (79) is the incompletive, which is diachronically the pure nominalized form (§4.3). This is in consonance with the latter being subordinated by the preposition *ti*.[49] Again at the stage of

---

[49] The documentary situation is such that this latter change appears earlier in the focused progressive than in the motion-cum-purpose construction proper.

Colonial Yucatec Maya, the binary contrast between *bin* 'go' and *tal* 'come' is yet maintained in their grammaticalization, as proved by (81). Observe, by the way, the third person on the pro-verb, obviously in analogy to the third person in the phase verb construction of §4.7.1.

(81) Colonial Yucatec Maya

  a. tal(el) v cah in botic in ppax
      tal(-el)   *=u*   *ka'h* *=in*   *bo't-ik*   *=in*   *p'aax*
      come-INCMPL =A.3 do   =A.1.SG pay-INCMPL =A.1.SG debt
      'I would like to pay my debt' (Coronel 1998a: 69)

Further reduction of the paradigm, however, leads to the consequence that the only verb possible in the Modern Yucatec Maya focused motion-cum-purpose construction is *bin*, and the construction only survives in the modern immediate future, illustrated by the intransitive and transitive sentences of (82).

(82) Modern Yucatec Maya

  a. *bin*   *=in*   *ka'h* *xíimbal*   *ti'* *le*   *chaan kaah* *…=e'*
      IMM.FUT =A.1.SG do   walk(INCMPL) LOC DEM little   village =R3
      'I am going to walk to that little village' (hts'on_016)

  b. *bin*   *=in*   *ka'h* *=in*   *xíimba-t*   *yuum*   *ahaw*
      IMM.FUT =A.1.SG do   =A.1.SG walk-TRR(SUBJ) master/father chief
      'I am going to visit the chief' (hts'on_020)

The preposition *ti'* no longer shows up in this construction in Modern Yucatec Maya. And as in the focused progressive (76), the full verb does not need to be an agentive verb, as shown by (83)f.

(83) Modern Yucatec Maya
      bin   *=in*   *ka'h* *kíim-il*
      IMM.FUT =A.1.SG do   die-INCMPL
      'I am going to die' (FCP 395)

(84) Modern Yucatec Maya
      bin   *=u*   *ka'h-o'b* *suut*   *ba'ba'l-il-o'b*
      IMM.FUT =A.3 do-3.PL turn\INTROV demon-ADVR-PL
      'they were becoming demons' (hnazario_415.5)

Christian Lehmann

By desemanticization, the semantic component of motion has disappeared, and what remains is only the direct tie between present topic time and future event time. *Bin … ka'h* is now a complex auxiliary with the value of immediate future.

(85) can serve for the clitic placement test:

(85)   Modern Yucatec Maya
        Behe'la'=e' bin   =in   ka'h wáah túun =in   kíins-ech?
        today=R3   IMM.FUT =A.1.SG do   INT   then =A.1.SG kill-B.2.SG
        'And now I shall kill you?' (hk'an_610)

It shows that – in contrast with the *bíin* of the predictive future – the first component of the discontinuous auxiliary cannot be host to a clitic, but the second component can. This is in consonance with the reduction processes to be analyzed in a moment and argues for the structural unity of the discontinuous auxiliary.

The structure of this auxiliation is peculiar within the grammar of Yucatec Maya in several respects. First, this is the only auxiliary which conditions different statuses on the full verb depending on the latter's transitivity, as is shown by (82). This is a reflection of the blending of two different constructions at its origin: The subjunctive on the transitive verb is a reflection of the motion-cum-purpose construction, which requires this status for the purpose verb. The incompletive on the intransitive verb is its nominalized form, which in turn is required by the preposition which originally governed this verbal core. It only remains to find out why the intransitive morphology reflects the verb-focus construction, while the transitive morphology reflects the motion-cum-purpose construction.

Secondly, *bin … ka'h* is the only discontinuous auxiliary of the language. What is more, the real auxiliary in the construction is the component *ka'h*. This, however, does not occupy the clause-initial position taken by all the other auxiliaries of the language. This position is, instead, taken by a verb which has the role of a full verb in the source construction. Thirdly, while *bin* is impersonal like all the other auxiliaries, *ka'h* is the only one with personal inflection. As a consequence, with transitive full verbs, the subject is cross-referenced twice (Briceño Chel 1998: 82), as is apparent from examples like (82b). There is, consequently, much redundancy in this auxiliation. In the colloquial register of Modern Yucatec, the full forms are rarely used. They are normally reduced in phonologically irregular ways, and there is currently much variation in this respect. Briceño Chel (1998: 82, 2000:88f, 2006: §§1.2, 1.3) notes the fusion of *bin in/a/u ka'h* into

nika'h/naka'h/nuka'h, as in (86a). If the full verb is transitive and therefore preceded by a Set A index, the *ka'h* of the auxiliary coalesces with it, as in #b.

(86) Modern Yucatec Maya

    a. *Ni-ka'h*   *meyah t=in*   *kool.*
       IMM.FUT\A.1.SG-do work LOC=A.1.SG milpa
       'I am going to work on my cornfield.' (Briceño Chel 2000a: 88)

    b. *Ni-k=in*   *hant*   *bak'*
       IMM.FUT\A.1.SG-do=A.1.SG eat:TRR(SUBJ) meat
       'I am going to eat meat' (Briceño Chel 2000a: 99)

Other idiosyncratic mergers occur in a variant of the construction in which the *ka'h* component takes Set B indexes. Using this variant with a transitive verb leads to cross-referencing the subject three times. The reduction processes applied in this context disguise this to a certain extent. Thus, the first syllable of the complex auxiliary in (87) contains the vowel of the 1st ps. sing. set A clitic.

(87) Modern Yucatec Maya
    mi-ka'h-en   *=in*   *wa'l*   *te'x*   …
    IMM.FUT\A.1.SG-do-B.1.SG =A.1.SG say(SUBJ) you.PL
    'I'm going to tell you …' (FCP_043)

However, contractions with clitics of other persons may also contain an *i*, so that the interim result of these changes is an auxiliary which takes Set B suffixes to cross-reference the subject of the clause core. In cases like (87) it leads to doubling, quite untypical of the language. The only comment one may make on the situation is that before a construction becomes a fixed grammaticalized inflected form, much variation occurs.

The other Yucatecan languages, too, have developed an immediate future on the basis of a focused progressive of the motion-cum-purpose construction involving their cognates of Yucatec *bin* 'go'. (88) shows the focused progressive with the defective pro-verb in Lacandón, which here already assumes an imminent future function (Bruce S. 1968:80, 101):

(88) Lacandón
    ok'ol   *=u* *kah*
    weep-INCMPL =A.3 do
    'he is about to cry' (Bruce S. 1968: 80)

Christian Lehmann

Applying this to the motion verb of the motion-cum-purpose construction already illustrated by (69) yields the Lacandón immediate future. Just as in Yucatec, reduction of the immediate future construction involves merger of the Set A index preceding the transitive full verb with the auxiliary *kah* immediately preceding it. Thus, *kah=in/a/u* yields *kin/ka/ku* (Bruce S. 1968: 95, 101), as in (89) (where *kah* must be a variant of *k=a* [do=A.2]) and (90a).

(89) Lacandón
      Bin =a   kah päy-e   lu'um-o',
      go =A.2 do   carry-SUBJ earthling-PL
      'You are going to take the earthlings with you,' (Bruce S. 1968: 76)

As an alternative to the construction of (69), an intransitive purpose clause may be introduced by the preposition *ti*, as in (90b). This may be seen as a direct continuation of the Colonial construction represented by (79) and is furthermore in analogy with the debitive construction illustrated by (61).

(90) Lacandón
      a. *bin   =in   k=in   wuk'-ik*
          IMM.FUT =A.1.SG do=A.1.SG drink-INCMPL
          'I am going to drink it'
      b. *bin   =in   kah t=in   wuk'-ul*
          IMM.FUT =A.1.SG do LOC=A.1.SG drink-INCMPL
          'I am going to drink' (Bruce S. 1968: 101)

In (91) from Itzá, the verb *b'el* 'go' is the full verb occupying the focus position in a simple verb-focus construction.

(91) Itzá
      (B'el)   =u   ka'a ich =u   kool.
      go(INCMPL) =A.3 do/go in =A.3 milpa
      'He is going to his cornfield.' (Briceño Chel 2000a: 90f)

If *b'el* is the motion verb of a motion-cum-purpose construction, an intransitive verb in the purpose clause is subordinated by the preposition *ti*, as in (92), while a transitive verb, as in (93), is in the subjunctive.

(92) Itzá

8 Grammaticalization of tense/aspect/mood marking in Yucatec Maya

(B'el) =*u*   *ka'a*   *ti*   *han-al.*
go   =A.3 do/go LOC eat-INCMPL
'He is going to eat.' (Bruce S. 1968: 91)

(93) Itzá
U-ka'ah   =*u*   *b'et-eh*   =*u-yotoch*
A.3-do/go =A.3 make-SUBJ =A.3-home
'He is going to make his home' (Hofling 1991, 1:5)

The peculiarity here is that since occurrence of the defective verb *ka'a* is all but limited to the construction with *b'el* in focus,[50] it assumes the sense of 'go' by syntagmatically mediated coding (Lehmann 2014). Consequently, *b'el* becomes redundant and may be omitted. This is true not only for the immediate future developed from the motion-cum-purpose construction, but also for the simple verb-focus construction of (91).[51]

The facts of Mopán, finally, are similar. (94) illustrates the simple verb-focus construction.

(94) Mopán
T'an   =*in-ka'aj.*
speak =A.1.SG-do
'I am speaking.' (Hofling 2011: 154)

(95) shows the immediate future construction with an intransitive full verb in the second person (cf. Hofling 2011:153). The #a and #b examples represent the full and reduced variants, resp. The same relationship holds between (96a) and (b), where the pronominal enclitic preceding the transitive verb is involved in the contraction, too. As may be seen, contraction of the auxiliary with the Set A index works similarly as in the Yucatec (86). Moreover, the intransitive verb of (95) is in the incompletive status and subordinated by *ti*, while the transitive verb of (96) is in the subjunctive.

(95) Mopán

a. *Bel*   =*a*   *ka'a ti*   *wäy-el.*
   go(INCMPL) =A.2 do   LOC sleep-INCMPL
   'You are going to sleep.'

---

[50] Hofling 1991:17 does present an example with *ka'a* as the main verb meaning 'do'. A similar construction in Ch'ol employs the cognate verb *cha'l* 'do' (Coon 2010, §3.1).

[51] This is mentioned in Briceño Chel 2000a, but not in Hofling 1991.

Christian Lehmann

   b. *B=a-ka'a ti wäy-el.*
     go=A.2-do LOC sleep-INCMPL
     'You're going to sleep.'

(96) Mopán

   a. *Bel   =in   ka'a =in   koykin   =a   nene'e*
     IMM.FUT =A.1.SG do   =A.1.SG lay.down(SUBJ) =DEM baby
     'I am going to lay the baby to sleep'

   b. *B=i(n)-k=in   koykin   =a   nene'e*
     IMM.FUT=A.1.SG-do=A.1.SG lay.down(SUBJ) =DEM baby
     'I'm going to lay the baby to sleep' (Briceño Chel 2000a: 95)

The languages of the Yucatecan branch share all the essential properties of the immediate future auxiliation: the discontinuous auxiliary, the multiple cross-reference to the subject and the asymmetry of status marking of the full verb conditioned by its transitivity, which reflects the contamination of two different syntactic constructions operative at the origin of this auxiliation. All four languages reduce this complex auxiliary construction; but as the processes operative here are not phonologically regular, they also differ among the languages.

The grammaticalization of the construction is a process in two main phases:

  a. verb focus construction > focused progressive

  b. focused progressive of auxiliary 'go' > (simple) immediate future.

More in detail, the following minimal steps compose the process:

- The motion verb *bin* 'go' is semantically bleached; the movement component disappears.

- The incompletive or subjunctive verb remaining in the extrafocal clause is reinterpreted as the main verb.

- The internal structure of the complex '*bin* set_A_index *ka'h*' is blurred. By being forced into the Procrustean bed of the initial position, it is reanalyzed as a discontinuous immediate future auxiliary with internal inflection.

- The whole sentence ceases to be complex; it is reinterpreted as a single clause.

- Whatever may have remained of the focal emphasis on the initial verb vanishes; the construction becomes open to different information structures that may be superimposed.

The model of this complex reanalysis is the structure of the simple fully finite clause of Figure 6, in which the initial auxiliary combines with the enclitic subject pronoun and is followed by the verbal complex (as, e.g., in (17b)). The result of the change conforms to that model to the extent possible for a discontinuous auxiliary.

4.10 Auxiliation in Yucatecan languages

The inherited suffixal system, where a minimum aspect system is coded as part of the status category, is renewed, in the period from Proto-Yucatecan to Modern Yucatec, by a large paradigm of aspectual auxiliaries. The sources of these auxiliaries are of different categories and form different syntactic constructions with the clause core. This explains the different status categories that they condition on the full verb. Conditioning them, they render them largely redundant. The new categories mark relatively fine distinctions not only of aspectual, but also of temporal and modal categories.

4.10.1 Syntactic relations

The new set of auxiliaries is structurally completely different from the inherited suffixal status-aspect-mood system. Since it owes its origin essentially to grammaticalization, it is based on syntactic rules operative at the time of its formation. There are four syntactic constructions at work:

a. an adverb modifying the verbal clause (core) following it and leaving its status marking intact

b. complementation, where a relational noun, an impersonal phase verb or modal verboid takes a verbal clause core in the dependent (subsequently incompletive) status as its complement

c. the motion cum purpose construction, where a verb of directed motion is followed by its purpose complement, represented by a verbal clause core in the subjunctive

Christian Lehmann

d. the verb-focus construction, which puts the main verb of the clause into focus position, leaving behind in the extrafocal clause a pro-verb with all the dependents of the focused verb.

The primary structural division of this set contrasts construction #a with constructions #b – #d. Construction #a is mono-clausal from the beginning. The auxiliary to-be bears a modifying relation to the clause core, which is syntactically independent. Constructions #b – #d transcend the simple clause; #b and #c are biclausal, #d is clefted. In these, the auxiliary to-be constitutes the main clause, while the clause core depends on it. As a consequence, auxiliation strategy #a leaves the syntactic relations in the clause core intact, while strategies #b – #d require some degree of nominalization of the clause core.

This difference has consequences for the configuration of basic syntactic relations in the clause core. These do not concern the transitive subject. Since Proto-Maya, this has been cross-referenced in all Mayan languages by the same Set A indexes which also cross-reference the possessor. This produces the ergative pattern of alignment shown by the cross-reference indexes. Since it appears primarily in completive status, which is semantically perfective, one may plausibly assume that assignment of possessive marking to the transitive subject stems, in its turn, from a pre-historic nominalization process. Be that as it may, the subordination of the clause core with auxiliation strategies #b – #d again requires nominalization of the clause core. Since the underlying transitive subject is already marked by a possessive relation, the intransitive subject now remains to be affected. This is why, in all tenses and aspects except the perfective, and also excepting subjunctive mood, it is marked by Set A indexes. The result is a rather peculiar form of aspect-conditioned split subject marking, which occurs in intransitive, not in transitive clauses.[52]

4.10.2 Grammaticalization of the auxiliary

Although the four constructions are clearly distinct, they share a clause-initial position which becomes the melting-pot for the aspectual and modal formatives recruited from different sources. Paths #a, #c and #d have been followed only once each in history; path #b has been the most prolific one.

Since the process of renewal and grammaticalization of auxiliaries has not finished, the paradigm is open and heterogeneous both in functional and in struc-

---

[52] If the analysis proposed in Coon 2010, §6 is accepted for Yucatecan, the indexing pattern would be ergative throughout, because what looks like accusative marking in almost all aspects actually occurs in subordinate clauses.

tural terms. In contemporary Yucatec, while all of the auxiliaries occupy the same structural position, the older ones are bound while the more recent ones are independent. And although several of them stem from verbs, they share the property of leaving conjugation categories to the full verb while remaining uninflected themselves. This is true with the single exception of the immediate future auxiliary, which is idiosyncratic in many respects.

The grammaticalization of auxiliaries evidences a process of clause union: it shrinks an original biclausal construction into a monoclausal one. This is perhaps clearer in Mayan languages, with their preference for verb-initial position and for impersonal constructions, than in many other languages. The many auxiliary constructions of the Yucatecan languages occupy all conceivable positions on a continuum from a complex sentence consisting of a matrix and a complement clause down to a one-clause sentence. Once the matrix predicate in initial position has been grammaticalized to an auxiliary, one might think that the construction is monoclausal. However, a simple test like the form of the answer to a polar question reveals that the auxiliary keeps being the main predicate.[53] Only after the auxiliary coalesces with the subject cross-reference index is it an irremovable part of a unitary clause.

The coalescence of the auxiliary with the following enclitic subject index is especially interesting. In SAE languages, the auxiliary is an element that hosts the conjugation categories of a finite verb, the most important of these being person and number. These are just the categories that the Yucatecan auxiliary lacks. Instead of denying it auxiliary status on these grounds,[54] it is intriguing to observe that, as a consequence of purely phonological enclisis, it coalesces with the subject indexes which syntactically accompany the following full verb, ending up as a morphologically complex form which codes not only tense, aspect and mood, but also person and number like an SAE auxiliary. However, the morphology – or maybe rather, the phonology – here is treacherous and not transparent to the syntax: even if merged with the preceding auxiliary, the pronominal index clearly forms a syntactic constituent not with it, but with the following verb, as shown in Figure 6 and proved by numerous examples like (4), (6c) and (32)f.

Although according to available descriptions, the complex of auxiliary plus

---

[53] While it is clear that the non-completive status suffixes are explained diachronically by the syntactic relation between the auxiliary and the full verb complex, Coon 2010, §2.3.3 insists that non-completive status clauses in Ch'ol are synchronically subordinate to the auxiliary. The imperfective and progressive auxiliaries, which are at stake here, do have a few more verbal properties than the Yucatecan imperfective auxiliary.

[54] Andrade (1955: §4.13) has a rather extensive discussion on the applicability of this term to the formatives in question.

Set A index is prefixed to the full verb in other Mayan languages, this has not happened in the Yucatecan languages. First of all, the enclitic status of the Set A index does not favor its univerbation with the material following it. Moreover, given the configuration 'set_A_index X', neither X nor this binary configuration is categorially uniform, since X may either be the head of this syntagma or may be a modifier of a head which is yet to follow (an adjective in a noun phrase or an adverb in a verbal complex). Consequently, although the auxiliary forms a phonological complex with the Set A index in many cases, there are syntactic obstacles to the univerbation of this complex with the verb of the following verbal complex.[55]

The grammaticalization of TAM in Yucatecan languages is a clear example of convergence of grammaticalization paths starting from different sources. The convergence is fostered, if not forced, by a rather rigid syntactic framework that a clause must fit in: First, an element that has scope over a verbal clause core must precede it. Although there are three distinct structural positions preceding a verbal clause core, their neutralization and merger into only one position is already predestined by the structure of Figure 7. Second, all of the operators that may occupy this position are impersonal. With these two constants to begin with, practically the only variable is the syntactic relation between the initial element and the clause core. This then determines the status to be chosen on the full verb. Since this variation in status is conditioned rather than informative, it could, in principle, be leveled out with ongoing grammaticalization. However, phonological reduction has rendered a subset of aspect auxiliaries homonymous. These aspects can then only be distinguished by the different status categories that they condition. This, in turn, prevents, for the time being, the disappearance of the status category.

The methodological lesson from the above for synchronic grammatical description is the following. Although all of the auxiliaries occupy the same structural position immediately preceding the clause core and although we are dealing with periphrastic constructions, a description which aims to account for the status forms of the full verb which accompany the diverse initial aspectual "particles" has to make explicit the syntactic relations between the initial element and the clause core. This, in turn, is facilitated if the grammaticalization source of these

---

[55] Some analysts (e.g. Hofling 1991:25, Pye 2009:266) claim the aspect auxiliaries to be prefixes. They are definitely not, in none of the Yucatecan languages. Hofling (1991: 37) keeps this analysis up by declaring the adverbs which may occur between the pronominal clitic and the verb to be "incorporated into the verb". This, however, is not so, witness the conjugation shown by verbs preceded by such adverbs: the stems do not become complex by this combination, which shows that it is a syntactic construction.

elements is taken into account.

5 Conclusion

While many of the grammatical formatives in the Mayan languages are etymologically unrelated, their functional categories and their structural properties are often identical. For instance, Yucatecan and Ch'olan languages share a large portion of the system of TAM auxiliaries; and these appear in the same structural position in all of these languages. What is more: they share particular aspects such as the perfective, imperfective, progressive etc.; but the morphemes appearing in these functions are unrelated. One must infer from this picture that the Mayan languages have been very conservative, over the millennia, as to their grammatical structure, and have limited themselves to renewing the formatives from time to time.

In view of the fact that grammaticalization is again and again hawked as a process of linguistic change, one must emphasize again and again that it is a process of linguistic variation both on the synchronic and on the diachronic axes. Moreover, history is always more complicated than diachrony: Variants that succeed each other on a dimension of grammaticalization co-occur synchronically, both within one language and across sister languages. And what would be a unitary source of a grammaticalized construction if one had to reconstruct it, with consideration of historical data turns out to be a set of variants and competing constructions that contributed in shaping the construction in question.

Christian Lehmann

Abbreviations

A	possessive/subject function	IPFV	imperfective
ADVR	adverbializer	LOC	locative
AN	animate	M	masculine
B	absolute function	NEG	negator
CAUS	causative	NEGF	negator, final part
CFP	clause final particle	PASS	passive
CL	classifier	PFV	perfective
CMPL	completive	PL	plural
CONJ	conjunction	PRF	perfect
COP	copula	PROG	progressive
DEAG	deagentive	PRSV	presentative
DEB	debitive	QUOT	quotative
DEM	demonstrative	rc	referential clitic
DEP	dependent status	R1	clitic of 1st ps. deixis
EXIST	existent	R2	clitic of 2nd ps. deixis
FUT	future	R3	non-deictic referential clitic
HAB	habitual	REC.PST	recent past
HOD	hodiernal past	SG	singular
IMM	immediate (future)	SUBJ	subjunctive
INAN	inanimate	TAM	tense/aspect/mood
INCMPL	incompletive	TERM	terminative
INT	interrogative	TOP	topic
INTROV	introversive	TRR	transitivizer
		VOC	vocative

References

Andrade, Manuel J. 1955. *A grammar of Modern Yucatec*. Microfilm Collection of Manuscripts on Middle American Cultural Anthropology, 41. Chicago: University of Chicago Library.

Aulie, H. Wilbur & Evelyn W. de Aulie. 1998. *Diccionario ch'ol de Tumbalá, Chiapas, con variaciones dialectales de Tila y Sabanilla* (Serie de vocabularios y diccionarios indígenas Mariano Silva y Aceves 121). Emily F. Scharfe de Stairs (1996 re-editing). Coyoacán, D.F.: Instituto Lingüístico de Verano.

Beltrán de Santa Rosa, Pedro. 1746. *Arte del idioma maya, reducido a sucinctas reglas, y semilexicon yucateco*. México D.F.: Viuda de d.J. Bernardo de Hogal.

Bergqvist, Henrik. 2011. Agentivity and status in Yukatekan languages. In Heriberto Avelino (ed.), *New perspectives in Mayan linguistics*, 242–257. Newcastle upon Tyne: Cambridge Scholars.

Bisang, Walter. 1991. Verb serialization, grammaticalization and attractor positions in Chinese, Hmong, Vietnamese, Thai and Khmer. In Hansjakob Seiler & Waldfried Premper (eds.), *Partizipation*, 509–562. Tübingen: G. Narr.

Briceño Chel, Fidencio. 1998. La gramaticalización del verbo *bin* ("ir") en el maya yucateco. In Zarina Estrada Fernández (ed.), *Memorias del IV encuentro internacional de lingüística en el noroeste. 4 volúmenes* (Serie Lingüística 1), 69–93. México D.F.: Instituto Nacional de Antropología e Historia.

Briceño Chel, Fidencio. 2000a. La gramaticalización del verbo "ir" en las lenguas mayas de la rama yukateka. In Zarina Estrada Fernández (ed.), *Memorias del V encuentro internacional de lingüística en el noroeste. 3 volúmenes*, vol. 1, 85–101. Hermosillo, México: UniSon.

Briceño Chel, Fidencio. 2000b. La gramaticalización del verbo 'terminar' (*ts'o'okol*) en maya yucateco. *Lingüística Mexicana* 1. 79–91.

Briceño Chel, Fidencio. 2006. *Los verbos del maya yucateco actual. Investigación, clasificación y sistemas conjugacionales*. México D.F.: Instituto Nacional de Lenguas Indígenas.

Bricker, Victoria R. 1981. The source of the ergative split in Yucatec Maya. *Journal of Mayan Linguistics* 2(2). 83–127.

Bruce S., Roberto D. 1968. *Gramática del lacandón* (Departamento de Investigaciones Antropológicas 21). México D.F.: Instituto Nacional de Antropología e Historia.

Bruce S., Roberto D. 1974. *El libro de Chan K'in* (Colección Científica Lingüística 12). México D.F.: Instituto Nacional de Antropologia e Historia, Departamento de Lingüística.

Coon, Jessica. 2010. Rethinking split ergativity in Chol. *International Journal of American Linguistics* 76(2). 207–253.

Coronel, Juan. 1998a. *Arte en lengua de maya*. Repr.: Martínez Hernández (ed.) 1929: 2-55. México D.F.: D. Garrido.

Coronel, Juan. 1998b. *Arte en lengua de maya. Edición de René Acuña* (Fuentes para el estudio de la cultura maya 14). México D.F.: Universidad Nacional Autónoma de México, Instituto de Investigaciones Filológicas.

Danziger, Eve. 2011. Once more with feeling: A forbidden performance of the Great Speech of the Mopan Maya. *Anthropological Quarterly* 84(1). 121–140.

Christian Lehmann

de Ciudad Real, Antonio. 1929. *Gran diccionario o Calepino de la lengua maya de Yucatán.* unpublished; first printed in Martínez Hernández (1929). Motul, Yucatán: Convento Franciscano.

Dik, Simon C. 1981. *Functional grammar* (Publications in Languages Sciences 7). Dordrecht & Cinnaminson: Foris.

Grinevald, Colette & Marc Peake. 2012. Ergativity in Mayan languages: A functional-typological approach. In Katharina Haude & G. Authier (eds.), *Ergativity, valency and voice*, 15–50. Berlin: Mouton de Gruyter.

Güldemann, Tom. 2003. Present progressive vis-à-vis predication focus in Bantu. A verbal category between semantics and pragmatics. *Studies in Language* 27. 323–360.

Hanks, William F. 2010. *Converting words. Maya in the age of the cross.* Berkeley etc.: University of California Press.

Hofling, Charles A. 1991. *Itzá Maya texts with a grammatical overview.* Salt Lake City: University of Utah Press.

Hofling, Charles A. 2006. A sketch of the history of the verbal complex in Yukatekan Mayan languages. *International Journal of American Linguistics* 72(3). 367–396.

Hofling, Charles A. 2011. Voice and auxiliaries in Mopan Maya. In Heriberto Avelino (ed.), *New perspectives in Mayan linguistics*, 144–159. Newcastle upon Tyne: Cambridge Scholars.

Lehmann, Christian. 1990. Strategien der Situationsperspektion. *Sprachwissenschaft* 16. 1–26.

Lehmann, Christian. 1991. Grammaticalization and related changes in contemporary German. In Elizabeth C. Traugott & Bernd Heine (eds.), *Approaches to grammaticalization. 2 vols.* (Typological Studies in Language 19), 493–535. Amsterdam & Philadelphia: J. Benjamins.

Lehmann, Christian. 2008. Information structure and grammaticalization. In Elena Seoane & María José López Couso (eds.), *Theoretical and empirical issues in grammaticalization* (Typological Studies in Language 77), 207–229. Amsterdam & Philadelphia: J. Benjamins.

Lehmann, Christian. 2014. Indirect coding. In Franz Rainer, Francesco Gardani, Hans Christian Luschützky & Wolfgang U. Dressler (eds.), *Morphology and meaning. Selected papers from the 15th international morphology meeting, vienna, february 2012* (Current Issues in Linguistic Theory 327), 97–124. Amsterdam & Philadelphia: J. Benjamins.

Lehmann, Christian. 2015. *Variación y normalización de la lengua maya, note = Conferencia magistral en el II Coloquio UNAM en la Península "El Maya Entre la*

variación y la estandarización". Centro Peninsular en Humanidades y en Ciencias Sociales (CEPHCIS) de la Universidad Nacional Autónoma de México (UNAM), Mérida, Yucatán, del 3 al 5 de noviembre de 2015. Mérida, Yucatán.

Monforte, Jorge, Lázaro Dzul & Rodrigo Gutiérrez Bravo. 2011. *Narraciones mayas.* México D.F.: Instituto Nacional de Lenguas Indígenas.

Pye, Clifton. 2009. Cycles of complementation in the Mayan languages. In Elly van Gelderen (ed.), *Cyclical change* (Linguistics Today 146), 265–284. Amsterdam & Philadelphia: J.Benjamins.

San Buenaventura, Gabriel de. 1684. *Arte de la lengua maya.* Pagination in this edition is in disorder, which is why references contain doubled page numbers and specification of recto (r) and verso (v). México D.F.: Viuda de B. Calderón.

Smailus, Ortwin. 1989. *Gramática del maya yucateco colonial* (Wayasbah publication 9). Hamburg: Wayasbah.

Vinogradov, Igor. 2014. La categoría de incompletivo en las lenguas mayas: Un estudio comparativo de variación semántica. *Studii de gramatică contrastivă* 21. 65–82.

Vinogradov, Igor. 2016. Acerca de la semántica del completivo/incompletivo en las lenguas mayas. *Cuadernos de Lingüística de El Colegio de México* 3(1). 1–44.

Chapter 9

Diachrony and typology of Slavic aspect: What does morphology tell us?

Björn Wiemer

University of Mainz

Ilja A. Seržant

Leipzig University

In this article we consider the Slavic perfective/imperfective opposition, a well-known example of viewpoint aspect which establishes a classificatory grammatical category by means of stem derivation. Although Slavic languages are not unique in having developed a classificatory aspect system, a survey of such systems shows that the Slavic perfective/imperfective opposition is a particularly rare subcase of such systems, first of all because it combines prefixing with suffixing patterns of derivation. We therefore explore the morphology involved, tracing its development from Proto-Indo-European into Early Slavic. The emergence of Slavic aspect is atypical for grammatical categories, and it deviates considerably from mainstream instances of grammaticalization in many respects. We show that there is a strong tendency (i) towards abandonment of highly lexically conditioned and versatile suffix choices in Proto-Indo-European and in Common Slavic, which led to fewer and more transparent suffixes, and (ii) towards concatenation, away from originally non-concatenative (fusional) schemata. Furthermore, we compare Slavic with some other Indo-European languages and inquire as to why in Europe no other Indo-European group beyond Slavic went so far as to productively exploit newly developed prefixes (or verb particles) merely for use as aspectual modifiers of stems and to combine them with a (partially inherited, partially remodelled) stock of suffixes to yield a classificatory aspect system. The Slavic system, thus, appears quite unique not only from a typological point of view, but also in diachronic-genealogical terms. Based on this background, amplified by some inner-Slavic biases in the productivity of patterns of stem derivation, we pose the provocative question as to whether the rise and consolidation of the stem-derivational perfective/imperfective opposition in Slavic was favoured by direct and indirect contacts

Björn Wiemer & Ilja A. Seržant

with Uralic (Finno-Ugric) and Altaic (Turkic) populations at different periods since at least the time of the Great Migrations.

1 Introduction

The Slavic aspect opposition of perfective/imperfective verbs is based on productive patterns of stem derivation involving both prefixes and suffixes. In general, the Slavic perfective/imperfective opposition does not belong among standard examples of grammaticalization; it can be captured by parameters as formulated in Lehmann (1995[1982]) or in Heine & Kuteva (2002) only to a very limited extent.[1] The main reason for this is that the morphological inventory involved does not originate in lexical items: suffixes have been created on the basis of the inherited Proto-Indo-European (henceforth: PIE) suffixes by various morphological reanalyses; in turn, the rise of prefixes from lexical items in principle corresponds to standard examples of grammaticalization, but this rise clearly predates the emergence of aspect. It can, thus, by no means be considered as a sufficient condition of grammaticalization, it supplies only one among many premises. Therefore, the ultimately lexical origin of prefixes should not be overestimated as a factor in the evolution of the Slavic aspect system. Semantic bleaching and morphological coalescence with verb stems prove to be well-attested processes, particularly in other Indo-European (IE) languages of Europe in which no aspect system has developed (see §5). If compared to PIE and Common Slavic, the new morphological patterns have become more transparent and, hence, less fusional. Although these patterns, to some extent, built upon an older system, the rise of Slavic aspect does not provide a counterexample against grammaticalization. On the one hand, it involved new morphemes (the prefixes); on the other hand, it consisted in a reduction of inherited patterns and a redistribution of suffixes (see §3). Thus, old and new techniques of affixation were combined; these processes involved only a minimal amount of inherited "material", but in their sum they led to a decrease of morphophonological opacity. This is rather atypical for grammaticalization. However, the gist of the story of Slavic aspect consists in properties that have knit together derivational patterns into a system. Regular and transparent processes of stem derivation have established a binary opposition of verb stems, called perfective and imperfective, which tend toward complementary functional distribution. Verb stems become divided according to an increasing amount of

---

[1] Nor could this sort of grammatical opposition be captured by an alternative proposal to give a unified account of grammaticalization phenomena as the conventionalization of discursive secondariness (cf. Boye & Harder 2012).

grammatical contexts, starting in the domain of actionality. That is, consistent patterns of stem derivation have led to a *classificatory category*: grammatical functions (imperfective vs. perfective) are indicated not by different forms (inflections) applying to one verb stem, but by the choice of different, though derivationally related, stems which, as it were, share a common paradigm of grammatical functions (see §2 and §4). Since a choice between perfective and imperfective stems is inevitable even in non-finite forms, the perfective/imperfective distinction has more and more become interwoven not only with other verbal categories (first of all tense and voice), but with virtually any sort of functional distinction on clause level and even beyond. In assessing the diachronic changes that led to the rise of the Slavic aspect opposition and which allowed it to be consolidated, we think it is essential to distinguish the morphological make-up (i.e. the derivational patterns) from the inventory of functions, such as [± limitation], single vs. repeated situation, volition- vs. cognition-oriented modal or illocutionary functions. Both, the derivational patterns and the function inventories, are necessary for grammatical aspect to arise and to strengthen its place in the grammatical system. However, in this paper we will treat aspect functions, beyond the core functions (see §2), at a minimum, since the continuous differentiation of functions associated with the choice of perfective vs. imperfective verb stems is part of the aspect story at comparatively recent stages. In fact, inner-Slavic differentiation has taken place primarily for those functions that are less motivated semantically by aspectual features, or actionality, proper (Wiemer 2008; 2015). These more recent stages build on an already established system of stem derivation whose basic architecture is unitary for all Slavic languages. We will here be concerned with this basic architecture and its rise. Furthermore, the uniformity with which this basic system of combined prefixation and suffixation applies to the core grammar of all Slavic languages is probably the reason why the perfective/imperfective opposition has long been considered as a hallmark of the Slavic group as a whole, even if classificatory aspect systems are not as typologically unique as they have often been considered to be (Arkadiev & Shluinsky 2015; see §4). Slavic aspect stands out for another reason as well: it is one of the few innovations in grammar which has affected the entire group, along with the rise of the pronominal declension of adjectives, the be-perfect based on the so-called *l*-participles, or the imperfect. The latter, however, together with the aorist, declined early and did not survive in most of the Slavic languages, whereas the perfective/imperfective opposition has not only survived, but has even been continually strengthened. In its basic morphological shape and its basic functions

Björn Wiemer & Ilja A. Seržant

the PFV:IPFV opposition is common to almost all Slavic languages,[2] while, for instance, other TAM-grams developed later and are much more differentiated (also in form), such as the future, the new perfects, as well as minor grams like the absentive in Czech.

Thus, the primary aim of this article is to trace back the *morphological* conditions which were necessary for the Slavic aspect system to arise. These conditions are partially shared by other ancient IE languages (with even some cognate morphemes). However, beside Slavic no other IE group has developed both prefixation and suffixation to the extent witnessed in Slavic, with its remarkably stable and consistent system and its pervasive impact on the entire grammar.

Therefore, as a secondary aim, we wish to ask how this system compares to the background of other IE languages of Europe, for which comparable conditions were inherited from IE predecessors, but which eventually did not develop a classificatory aspect system like the Slavic languages (with Lithuanian being an exception to some extent; see §5.4).

Since the core system of classificatory aspect is identical for all Slavic languages, it must have been established in Common Slavic times, i.e. prior to an increase in dialectal differences that would be an obstacle in the spread of innovations from different parts of the Slavic-speaking world. Common Slavic is assumed for approx. 3rd-7th c. AD (Andersen 1985; Holzer 2014: 1126f.), and this is the period of the Great Migrations, which must have affected early Slavs and their neighbors (e.g. Goths, Balts, but also some Altaic populations). Thus, one feels justified in seeking external factors that might have been favorable for the germs of an aspect system based on stem derivation to evolve into a consistent system (see §6).

The article is structured as follows. We start with a condensed presentation of the modern Slavic aspect system (§2). In §3, we present a diachronic account of the morphological inventory involved and the main functional changes it underwent in Common Slavic and the early documented stages of Slavic. In §4, our aim is to establish the peculiarities of the Slavic system on a broader typological background such as the consistent combination of prefixation *and* suffixation. While §4 is a synchronic comparison, §5 is devoted to a survey of the functions of stem derivation (mainly those resulting from preverbation) in some neighbouring, old IE languages. Moreover, since this survey provides ground for the

---

[2] From the functional point of view, two Slavic minority languages outside today's coherent territory of the Slavic speaking world are somewhat exceptional, each in different respects: colloquial Upper Sorbian (in Saxony, Germany) and Molisean Slavic (in southern Italy). We may however neglect the peculiarities of their tense-aspect systems since they do not substantially change the line of the entire argument developed here.

assumption that some complementary factors probably favoured the evolution of Slavic aspect, we sketch considerations concerning contacts with speakers of non-IE languages as an additional factor. This is done in §6, where we also formulate conclusions and give an outlook for further research. For reasons of space and perspicuity we refrain from giving glosses whenever full information about the morphology (beyond the aspect of the verb) is not mandatory. Aspect, of course, will always be indicated. If no other glossing is supplied we will mark it with upper-case small capitals (PFV, IPFV).

2 A sketch of the contemporary system

Our sketch of the contemporary core system starts with terminological clarifications and the basic functional oppositions before we explain the patterns of stem derivation. We will gloss over many details and have to simplify some points, but hope that examples sufficiently illustrate the crucial issues.

2.1 Actionality types and aspect as an operanda—operator relationship

The term *actionality* refers to basic situation types, or eventualities, which can be denoted by predicative units, first of all verbs. The basic eventuality types are event, process and state. Events are situations that are conceived of as holistic entities (e.g., *open a door*, *put on one's coat*, *take a glimpse*, *buy a book*). If the lexical concepts denoting these situations imply an internal structure this structure is out of consideration. Thus, for instance, the concepts denoted by *open (a door)* and *buy (a book)* may be internally quite complex and consist of different subevents that, as it were, prepare the event OPEN or BUY, but for the lexical units this complexity is, by default, out of focus. This default may be cancelled, though (see §2.2). By contrast, the eventuality behind *take a glimpse* is usually conceived of as instantaneous and the meaning of this expression does not comprise any accompanying events (e.g., *raise one's head*). Let us comment on the two other basic situation types. Processes are dynamic and can be subdivided into phases (e.g., *walk in the park*, *watch TV*, *deliberate*, *work*), whereas states do not have phases since they are not dynamic (e.g., *sleep*, *love*, *cost*, *hold in esteem*, *live*). Processes and states can be limited. Events, processes and even states can be made subject to repetition (e.g., *Every morning he smoke a cigarette / watched TV*; *Every second year she was pregnant*). Actionality features (and their alternations, see §2.2) exist for lexical concepts in any language, prior to, and independently from,

aspect as a grammatical category (which a language may have or not).

Aspect, in turn, operates on actionality features; the relation between action-ality types and aspect can, thus, be captured as a relation between operanda (lexical or clausal categories)[3] and operators (grammatical categories). For an overview cf. Sasse (2002). Actional features inherent to lexemes or phrases have also been subsumed under the term *lexical aspect* and been opposed to *viewpoint aspect*, which amounts to a system of operators making up aspect as a gram-matical category. The terminological distinction of lexical vs. viewpoint aspect was introduced by Smith (1991). Misleading as either of these terms is (Johanson 2000: 28; Plungjan 2011: 405f.), it has become quite commonplace in aspectol-ogy and related domains of (formal) semantics, and it is roughly equivalent to the operandum—operator distinction. In what follows, *aspect* will always mean viewpoint aspect, and we use the latter term only if we want to emphasize a con-trast with actionality (or *lexical aspect*). The adjective *aspectual* will be used to mean either actionality features or distinctions conveyed by the choice of aspect. This corresponds to commonplace practice, but we are confident that each time this adjective is used the context will make it clear whether we are talking about lexical features or properties of the aspect opposition. The term *event* will relate only to the specific actionality type defined above and further discussed in this section; the generic term for all actionality types is *eventuality* or *situation (type)*.

2.2 Aspect as grammaticalized marking of (un)boundedness

What is essential about the grammatical character of aspect? Trivially, for any grammatical category it is necessary to find regular distinctions of form that al-low stable functional oppositions to be predicted with a high degree of reliability. The other side of the same coin is that distinctions between forms become more and more compulsory. Moreover, if these distinctions have become established sufficiently firmly, they begin to be associated more reliably with distinctions in functional domains that are only remotely related to core features, that is to fea-tures which originally motivated the given grammatical opposition. As concerns aspect, a systematic variation in the morphology of verbs (or verb phrases) must show highly consistent correspondences with recurrent oppositions of functions relevant for actionality. This is to tell the least: as indicative of the continuing

---

[3] It may be argued that eventualities are not properties of lexemes (verbs, adjectives), but of verb phrases. Likewise, Vendler's known categories (achievements, accomplishments, activities, states) are similar to the eventualities named above, but they are essentially clausal features. Anyway, these different levels all constitute operanda for aspect, therefore they do not affect the fundamental point being made here.

9 Diachrony and typology of Slavic aspect: What does morphology tell us?

grammaticalization of aspect distinctions one may regard that morphological distinctions, once established in the domain of actionality, encroach into other functional domains, such as modality, and that these distinctions become increasingly constrained by other verbal categories (see below and §2.5).

In Slavic, core features for which the PFV:IPFV opposition proves highly sensitive are ([± limitation], a.k.a. *boundedness* and [± singular situation]). However, the formal opposition that serves to distinguish these core features has also expanded into modal and illocutionary domains. The latter ones will be illustrated below, let us first discuss the core domain. Compare, for instance, Russian:

(1) a. *Par-a   (dolgo)   **ljubova-l-a-s**$^{\text{IPFV}}$ zaxod-om solnc-a.*
        pair[F]-NOM.SG long.time admire-PST-SG.F-RFL sun.set-INS sun-GEN
        'A long time the pair **was admiring** the sunset.'

    b. *Para   **po-ljubova-l-a-s**$^{\text{PFV}}$ zaxod-om solnc-a (i*
        pair[F]-NOM.SG PFX-admire-PST-SG.F-RFL sun.set-INS sun-GEN and
        uš-l-a$^{\text{PFV}}$).
        go.away-PST-SG.F
        'The pair **admired** the sunset (for some time) and (then) went away.'

(2) a. *Petj-a   **čita-l**-∅$^{\text{IPFV}}$ roman-∅.*
        PN[M]-NOM read-PST-SG.M novel-ACC.SG
        'Peter **read / was reading** a novel.'

    b. *Petj-a   **pro-čita-l**-∅$^{\text{PFV}}$roman-∅.*
        PN[M]-NOM PFX-read-PST-SG.M   novel-ACC.SG
        'Peter **finished reading** a/the novel.'

(3) a. *Včera   Katj-a   **kupi-l-a**$^{\text{PFV}}$ sebe   nov-oe plat'-e.*
        yesterday PN[F]-NOM buy-PST-SG.F RFL.DAT new   dress[N]-ACC.SG
        'Yesterday Katja **bought** herself a new dress.'

    b. *Každ-yj mesjac-∅   Katj-a   **po-kupa-l-a**$^{\text{IPFV}}$ sebe*
        every   month[M]-ACC.SG PN[F]-NOM PFX-buy-PST-SG.F RFL.DAT
        nov-oe plat'-e.
        new   dress[N]-ACC.SG
        'Every month Katja **bought** herself a new dress.'

From the lexical point of view, the imperfective verb *ljubovat'sja* 'admire, feast one's eyes upon' denotes a process with no inherent endpoint; so does its perfective counterpart, and this shared lexical feature is what unites both verbs which

Björn Wiemer & Ilja A. Seržant

are also related to each other by a derivational affix (see below). We can speak of an aspectual pair. But while imperfective *ljubovat'sja* in (1a) is used to mark the process as unlimited, i.e. in accordance with this lexical default, its perfective counterpart *poljubovat'sja* in (1b) adds a temporal limitation to this process. This limited process can be part of a chain of events; this is indicated by the continuation *i ušla* 'and went away'. Note that *ujti* 'go away' implies a change of state, or otherwise: it is goal-directed (a.k.a. *telic*; see §2.3). However, despite this lexically implied difference between *poljubovat'sja* and *ujti*, both verbs belong to perfective aspect, and in a narrative setting this combination of perfective verbs yields a sequencing effect.

We observe the same difference if we compare (2a, IPFV) and (2b, PFV), although these verbs imply an inherent goal. In turn, a comparison between (3a, PFV) and (3b, IPFV) shows that the perfective verb is used if the eventuality is a single event; if this event is repeated, the imperfective counterpart is the preferred or the only possible option. So far, we may generalize that perfective verbs are used to present an eventuality as limited, the Slavic perfective aspect can therefore be called a limitative aspect (cf., for instance, Breu 2000a: 38). The imperfective aspect, in turn, is used either to defocus limits or to mark the unlimited or regular repetition of an eventuality.

There is a third main function of imperfective verbs, the so-called general-factual meaning (usually restricted to the past), by which the speaker simply states that, or asks whether, an eventuality has occurred or not. It can be compared to the experiential function of perfects, as in 'yes-no' questions, e.g. Russ. *Ty kogda-libo* **platil**[IPFV] *v kafe naličnymi?* 'Have you ever **paid** cash in a café?'. Furthermore, a very salient distinction applies in imperatives. Simplifying somewhat, we can say that, by default, if an imperative is issued with respect to a single action, a perfective verb is used (e.g., Russ. *Zakroj dver'!* 'Close the door!'); if the imperative is denied, an imperfective verb would be used (*Ne zakryvaj dver'!* 'Don't close the door!'), even if it refers to the same single action. Here the aspect opposition already interacts with mood.

Regardless of this slight expansion, it is important to realize that perfective verbs mark limitation even if the eventuality is not lexically preconceived as an event; they recategorize processes, sometimes even states, into events. This happens in (1b) above: *ljubovat'sja* denotes a process, but its perfective counterpart in (1a) makes it into an event. Similarly, the perfective verb *pro-suščestvovat'* 'exist for some time span' denotes a state which existed within some temporal boundaries and which, thereby, is recategorized as an event. Compare a corpus example:

9 *Diachrony and typology of Slavic aspect: What does morphology tell us?*

(4) *Po ocenkam arxeologov, "Strana gorodov"* **pro-suščestvova-l-a**[PFV] *na*
                                                          PFX-exist-PST-SG.F

   Urale okolo trex stoletij.

   'According to archeologists' assessments, the "Country of cities" **existed**
   in the Ural for about three centuries.' (NKRJa; *Znanie-sila*, 2013)

Conversely, imperfective aspect can serve to focus on a process if the verb meaning implies an endpoint, but this endpoint is out of consideration (e.g., Russ. *Celuju nedelju on* **pererabat-yva-l** [IPFV] **stat'ju** 'A whole week he **worked on a re-cast** of his article'), while perfective aspect focuses on the achievement of this goal (*On* **pererabot-a-l** [PFV] **stat'ju** 'He **wrote a re-cast** of his article'). This is the case with goal-directed (i.e. telic) verbs (on which see §2.3).

By a similar token, the aspect distinction is sensitive to eventuality alternations. If a lexeme has two alternative actionality readings these are rigidly distributed over perfective and imperfective counterparts. Consider event—process alternations. Events can alternate with processes if the internal structure of the latter consists of heterogeneous phases, we are then dealing with goal-directed processes denoted by imperfective verbs (e.g., Russ. *otkryvat'*[IPFV] *okno* 'open a window', *stroit'*[IPFV] *dom* 'build a house'); but events can also alternate with processes if the lexemes do not imply any goal. We observe this with the semelfactive—repetitive alternation (compare, e.g., *wave one's hand*, *knock at the door*, *jump*, *kick*). Semelfactives are always perfective, and they are even marked with a specific suffix (*max-**nu**-t'* 'wave', *vil'-**nu**-t'* 'wag one's tail', etc.); their repetitive counterparts are always imperfective (*max-**a**-t'*, *vilj-**a**-t'*, etc.). On this alternation see also Footnote ??; in the literature on Slavic aspect repetitives are usually referred to as 'multiplicatives'. Furthermore, even states can alternate with events. This happens regularly with emotive and perceptual predicates: the imperfective verb denotes the state (*ljubit'* 'love', *volnovat'sja* 'be excited, agitated', *videt'* 'see', *zameč-a-t'* 'notice', *vospri-nim-a-t'* 'perceive'), its perfective counterpart the corresponding inceptive event (*po-ljubit'*, *vz-volnovat'sja*, *u-videt'*, *zamet-i-t'*, *vospri-nja-t'*, respectively).

Already from this cursory glance at core distinctions distributed over stems belonging to perfective or imperfective aspect we can infer two things: (i) On the one hand, both perfective and imperfective aspect are "harmonious" with different eventuality types; namely, the basic function of perfective verbs corresponds to events, while their imperfective counterparts are used if related processes or states are to be named. To some extent, these correspondences motivate the basic

actional functions of aspect. (ii) On the other hand, the PFV:IPFV distinction does not depend on inherent features of verb lexemes, and the concepts underlying these lexemes can be presented as different eventualities, i.e. they can be recategorized in accordance with the opposite aspect (Mende 1999: 289-294; Lehmann 2004: 174-177). That is, aspect must be able to override lexical defaults; it unifies verb stems (or verb forms), regardless of such defaults, for some more abstract functional purpose. Consequently, as parameters by which the degree of grammaticalization can be determined we may regard two things. First, the freedom, or flexibility, with which lexical concepts of processes and states can be recategorized as events by being marked as perfective, and, conversely, with which one can defocus from the boundaries of events, laying stress on a correlated process or on repetition, by using an imperfective verb. Second, the reliability with which more abstract functional oppositions, possibly in combination with other verbal categories, can be marked by perfective or imperfective verbs. In any case, the aspect of a verb is recognized on the basis of derivational patterns. Before we dwell on them, it is expedient to introduce necessary distinctions connected to the notion of *telicity*.

2.3 Aspect and telicity

The preceding discussion should have made it obvious that goal-directedness, or telicity, is not a defining property of (perfective) aspect in Slavic. There is an undisputed association between perfective aspect and telic verbs (forms or stems) inasmuch as perfective verbs are the functionally unmarked choice for telic events (see §3.2.2, Footnote ??), but, as we have already seen, events need not be goal-directed; compare semelfactives and processes or states delimited by mere temporal boundaries. Conversely, processes can be goal-directed. Compare, for instance, verbs with incremental objects or, more broadly, incremental changes, e.g., Pol. *Rodzice już od pięciu lat* **budują**[IPFV] *dom* 'The parents have been [lit. *are*] **building** a house for five years', or *Chłopiec powoli* **zasypiał**[IPFV] 'The boy was slowly **falling asleep**'. (A)telicity is a lexical property of verb stems, or of verb phrases (as the case may be). If, thus, affixes were to mark simply a change of this property (telic → atelic, atelic → telic), this would preclude the rise of lexical equivalents differing only on aspectual core features, such as shifts of focus between some available boundary and an associated process or state. In other words: a system of perfective:imperfective verbs can hardly be established if it is built only on a strict association between telic situations and perfective

verbs vs. atelic situations and imperfective verbs.[4]

The notion of (a)telicity needs some further clarification and refinement, for reasons that will become evident in §3.2 and §5. In the following we draw on Dahl (1981), Łaziński & Wiemer (1996), and Arkadiev [Arkad'ev] (2015: 21-24). We take *telicity$_1$* to mean an inherent feature of a verb lexeme or a predicate that makes the denoted situation imply an inherent endpoint, regardless of whether this endpoint is realized or not.[5] In turn, *telicity$_2$* puts an assertive focus on the realization of this lexically implied endpoint. Telicity$_1$ as an inherent feature of a lexeme, regardless of whether its implied limit has been reached or not, is what traditionally most Russian scholars (following Maslov 1948) have been understanding under this term (Russ. *predel'nost'*, Germ. *Terminativität*). By contrast, scholars working in the tradition going back to Vendler (1957) have been using telicity$_2$ as a property indicating that the inherent endpoint has been reached, often even regardless of whether the predicate implies an inherent endpoint lexically or on clause level. From this perspective, a sentence like *He was writing a letter* would be atelic$_2$ but telic$_1$, while *He wrote (up) a letter* would be telic$_2$ and also telic$_1$. Either of these sentences implies an endpoint, but only in the last one this endpoint is presented as attained. To add to the confusion, telicity$_2$ has also been used as an indication that the situation has been delimited by merely temporal boundaries. This, of course, occurs if some perfective operator applies to a predicate which is atelic$_1$, i.e. activities in Vendlerian terminology. This is a standard function of the aorist with atelic$_1$ predicates, e.g. Italian *cant-ò* (sing.AOR.3SG) 's/he sang', *lavor-ò due ore* (work.AOR.3SG) 's/he worked for two hours' or Ancient Greek. This function is salient also with prefixes like Russ. delimitative *po-*, e.g. *po-guljat' po parku* 'walk (some amount of time) in a park', *po-sporit' s drugom* 'argue (for some time) with a friend', or *po-smotret' televizor* 'watch TV (for some time)'; see the discussion of examples (1a-b) and (4) above. In order to avoid misunderstandings (and clumsy circumscriptions) we supply the term *(a)telic* and all its derivatives with an index whenever we consider it

---

[4] Here we skip over colloquial Upper Sorbian, which is the only exception to this rule in contemporary Slavic (Breu 2000b).

[5] Compare the difference between *walk (around)* and *go (to the shop)*, or between actions without an inherent limit (*cry*, *shout*), or momentary (punctual) verbs like *find*, *notice*, *wince*, on the one hand, and verbs with an inherent endpoint, e.g. *solve (a problem)*, *build (a house)*, *break (a window)*, on the other. Many verb lexemes can have either a telic$_1$ or an atelic$_1$ reading, such as consumption verbs (*eat*, *drink*) or activities like *read*, *write*. Characteristically, in modern Slavic languages perfective counterparts of such verbs tend to have different prefixes depending on the telic$_1$ or atelic$_1$ reading (e.g. Russ. *čitat'*[IPFV] 'read' ⇒ *pro-čitat'*[PFV] 'read through', telic$_1$, vs. *po-čitat'*[PFV] 'read a little bit / for some time', atelic$_1$).

Björn Wiemer & Ilja A. Seržant

appropriate.

To resume, imperfective and perfective verbs can both be telic$_1$, i.e. imply a natural boundary, but only the perfective verb asserts that this natural boundary has been reached (= telic$_2$). In general, perfective verbs only assert that some boundary has been set, whether this boundary is inherent or only a temporal one depends on whether the predicate is telic$_1$ or atelic$_1$. In turn, with the imperfective, other things remaining equal, the focus shifts to other parts of a more complex situation, e.g. the gradual approachment toward an implied goal (= progressive accomplishment reading, or: incremental change)[6] or a state that follows from an event. Perfectivity has to be understood as a grammatical property, since it is the result of an operation by which, regardless of lexical defaults, a situation can be presented (or construed) as bounded and, if necessary, be recategorized as an event. *Bounded* means that the situation is presented with limits, regardless of the telic$_1$ or atelic$_1$ character of the eventuality. Therefore, *bounded* and *perfective* can be treated as practically synonymous notions, although boundedness, at least diachronically, often comes in as a feature which implies the introduction of some inherent endpoint (also called *telicization*); see §3.2. In other words: all events are bounded by definition, and the grammatical function of perfective aspect is to mark a situation as bounded, regardless whether this boundary coincides with some inherent endpoint or not. That is, perfective aspect makes the lexical concept suitable for functions that are associated to boundedness, such as a sequencing effect in narrative discourse; see the discussion of example 1b and 7a.

Thus, telicity$_1$, as a lexical feature, does not entail perfectivity, nor vice versa. Both atelic$_1$ and telic$_1$ predicates can be perfectivized, and the grammatical status of the means which mark perfectivization enhances by the degree of productivity and predictability with which perfectivizers apply not only to telic$_1$ verbs, but also to atelic$_1$ ones. In other words: perfectivization has a broader extension than telicization$_1$, since it does not depend on, or change, the lexical properties of a predicate. This is why, as a rule, telicization$_1$ does not *per se* constitute a perfective:imperfective system (see however Footnote ??); it remains too restricted to certain actional, thus lexically specific, classes of verb lexemes (or predicates).

---

[6] This, of course, works only for verb lexemes which imply such an endpoint. Such telic$_1$ verbs supply the starting point for many aspect systems (see §3.2.2).

2.4 The morphological make-up of classificatory aspect

We turn now to the classificatory character of Slavic aspect. In many other languages for which viewpoint aspect is acknowledged, predictable and reliable form:function correspondences are marked by inflectional desinences, and they are often restricted to the past domain as in Romance (aorist vs. imperfect), or they are marked periphrastically as, e.g., in English (simple vs. progressive forms). Much less known and acknowledged are languages in which such form: function correspondences are based not on inflection on the same verb (stem), but on the classification of different, though morphologically related verb stems. In such a system we encounter regular patterns of stem derivation: the new stem is derived by (i) an additional suffix or (ii) by an additional prefix. To these patterns we can ascribe different sets of functions for each member of a derivational pair; and the more these sets of functions become complementary, i.e. do not intersect with each other, the more reliably the choice of the prefixed vs. suffixed stem marks off contrasting values of stable functional oppositions. What we eventually get is a binary classification of verb stems.

It is important to realize that both fundamental principles cooperate: transparent derivational relations for the absolute majority of verb stems, and a tendency toward complementary distribution of functions for each class, i.e. perfective vs. iperfective stems. Some of the functions were already illustrated above. For productive patterns of stem derivation see examples from modern Polish, for infinitives (5a-b) and inflected forms (6a-b); '*' marks off reconstructed forms:[7]

(5) Perfective/imperfective derivation with infinitives

    a. simplex imperfective $\Rightarrow$ perfective by prefixation
       łowi-ć$^{\text{IPFV}}$ $\Rightarrow$ **z***-łowi-ć*$^{\text{PFV}}$
       catch-INF   **PFX**-catch-INF
       patrze-ć$^{\text{IPFV}}$ $\Rightarrow$ **po***-patrze-ć*$^{\text{PFV}}$,
       observe-INF   **PFX**-observe-INF
       podoba-ć się$^{\text{IPFV}}$ $\Rightarrow$ **s***-podoba-ć się*$^{\text{PFV}}$
       please-INF RFL   **PFX**-please-INF   RFL

---

[7] Here only the most productive and salient patterns are used for illustration. In some cases suffixes are not added, but replaced. However, with one exception, replacement relations have become unproductive. The exception is the nasal suffix. For example, *-ną-* in Polish replaces *-a-*, but only for semelfactive (PFV) vis-à-vis multiplicative (IPFV) verbs; compare *mach-**a**-ć* vs. *mach-**ną**-ć* 'wave', *dźg-**a**-ć* vs. *dźg-**ną**-ć* 'prod, stab', etc. (see §2.2). These suffixes are older than the suffixes used in productive additive patterns of prefixation and suffixation (see §3.2.1).

Björn Wiemer & Ilja A. Seržant

b. perfective stem by prefixation ⇒ secondary imperfective by suffixation

na*-*mówi*-*ć[PFV]   ⇒ ***na*-*mawi*-*a*-*ć***[PFV],
PFX-persuade-INF   PFX-persuade-SFX-INF

prze*-*kona*-*ć[PFV]   ⇒ ***prze*-*kon*-*ywa*-*ć***[IPFV]
PFX-persuade-INF   PFX-persuade-SFX-INF

s*-*po*-*strze*-*ć[PFV]   *(< *s-po-streg-ti)* ⇒
PFX-PFX-take_notice-INF

s*-*po*-*strzeg*-*a*-*ć[IPFV]
PFX-PFX-take_notice-SFX-INF

(6)   Perfective/imperfective derivation with finite forms of past and present

a. simplex imperfective ⇒ perfective by prefixation

pis-a-ł-a[IPFV]   ⇒ ***na*-*pis-a-ł-a***[PFV]
write-THV-PST-SG.F   PFX-write-THV-PST.-SG.F
'she wrote, was writing'   'she wrote (up)'

pisz-ę[IPFV]   ⇒ ***na*-*pisz-ę***[PFV] *(<*(na-)pis-jǫ)*
write.PRS-PRS.1SG   PFX-write.PRS-PRS.1SG
'I write, am writing'   'I will write'

b. perfective stem by prefixation ⇒ secondary imperfective by suffixation

roz*-*wiąz-a-l-i[PFV]   ⇒ ***roz*-*wiąz-ywa-l-i***[IPFV]
PFX-bind-THV-PST-PL.VIR   PFX-bind-SFX-PST -PL.VIR
'they tied/were tying off'

roz*-*wiąz-uj-ą[IPFV]
PFX-bind-SFX-PRS.3PL
'they tie/are tying off'

In Slavic verbal morphology this principle is pervasive, because both prefix-ation and suffixation are not only prolific, but also able to focus on aspectual features alone without restrictions of tense or changes related to argument struc-ture or valency. Most of these prefixes and suffixes are transparent and clearly segmentable from the original stem and desinences marking other categories, despite systematic morphonological alternations between stem and inflectional ending (see (6a) for the present tense stem) or allomorphy of suffixes (see (6b) for past/infinitive vs. present tense stem). This is why we end up with a classi-ficatory system in which the morphological relations between the predominant number of stems remain transparent (for the rise of these relations see §3).

Note that in a persistent classificatory system, the issue whether perfective and imperfective stems always come in pairs becomes less important. On the one hand, even etymologically unrelated stems can be united into aspect pairs, i.e. stems of opposite aspect with an identical lexical meaning. These are suppletive pairs like Russ. *brat'* [IPFV]—*vzjat'* [PFV] 'take', *lovit'* [IPFV]—*pojmat'* [PFV] 'catch', *klast'* [IPFV]—*položit'* [PFV] 'put'. Such pairs distribute among themselves the same sets of functions as do other stem pairs (see §2.5). On the other hand, it proves to be of minor importance that most Slavic prefixes mark not only a shift from the class of imperfective stems into the class of perfective stems, but tend to also modify the lexical meaning of the deriving stem. Compare, for instance, Russ. *pisat'* [IPFV] 'write' ⇒ *pere-pisat'* [PFV] 'write anew', where the change to perfective aspect is accompanied by a lexical modification of the meaning of the imperfective stem. This differs from cases like *pisat'* [IPFV] ⇒ *na-pisat'* [PFV] 'write (up)' / *po-pisat'* [PFV] 'write (for a while / a certain amount of text)', where the prefix only marks a change to perfective aspect. Regardless of this difference, each stem belongs to either perfective or imperfective aspect and the class membership is determined by the restriction to opposed sets of functions.

Prefixed stems whose lexical meaning differs from the lexical meaning of their simplex are often further suffixed, which yields an imperfective equivalent with identical lexical meaning. Compare, for instance, the Polish example in (6b) or Russ. *pere-pisa-t'* [PFV] (PFX-write-INF) ⇒ *pere-pis-**yva**-t'* [IPFV] (PFX-write-SFX-INF); both stems mean 'write anew'. This process is called *secondary suffixation*. In modern Slavic languages, secondary suffixation is highly productive, in the eastern half of Slavic possibly even more than prefixation in that aspect pairs are derived primarily via (secondary) suffixation (see §3.2.3).[8] Moreover, the set of suffixes including allomorphs is much less numerous than that of prefixes. Again, it is essential that the derivational patterns (illustrated in examples 5-6) combine into a systematic classification of verb stems; the distribution of these functions is, by and large, independent from the specific pattern. This insight has been corroborated by Janda & Lyashevskaya (2011) who show that, on average, Russian aspect pairs are characterized by basically identical oppositions of function sets (*aspectual profiles* in their terms), regardless of whether they are based on the pattern simplex stem (A)^{IPFV} ⇒ (prefix+A)^{PFV} or on the pattern (prefix-A)^{PFV} ⇒

---

[8] The eastern half of Slavic comprises East Slavic (with Polish behaving more like East Slavic than like the rest of West Slavic) and Balkan Slavic (Bulgarian, Macedonian). Admittedly, all claims related to suffixation must be verified for languages of the western half of Slavic, in which secondary suffixation appears to be less prominent than in the eastern half (cf. Arkadiev [Arkad'ev] 2015: 124f. and references therein). However, such inner-Slavic differences do not invalidate the principled point which we are making.

Björn Wiemer & Ilja A. Seržant

[(prefix-A)+suffix]$^{\text{IPFV}}$. The relative insensitivity of the two predominant derivational patterns with respect to a more fine-grained functional distribution between stems of aspect pairs is, thus, another strong indicator of the coherence of the classificatory system.

2.5 Aspect pairs and continued increase of grammatical restrictions on aspect choice

Although the complementary inventories of imperfective and perfective aspect are not as such dependent on aspectual pairedness, aspect pairs nonetheless form the backbone of the system, both concerning its diachronic development and their role in the contemporary Slavic languages. Roughly, aspect pairs provide the core of a system in which lexical concepts coded by verb stems are manipulated by aspect to meet various grammatical constraints. Aspect pairs have become a time-honoured subject of never ending disputes in Slavic (mainly, Russian) aspectology. We do not intend to engage into this discussion, but we want to specify the relevance of aspect pairs just for the limited purpose of our investigation. Above, aspect pairs were introduced as pairs of imperfective and perfective verbs whose absolute majority shows a transparent derivational relation, which share the same lexical concept, but which are differently distributed over functions related to actionality and beyond. Since for no (inflected or nonfinite) form of a verb the choice between perfective and imperfective stems can be avoided and other categorial distinctions are expressed by verb forms as well, aspect choice more and more interferes with these distinctions.

Interference can turn into hard constraints. A prominent illustration is provided by narrative passages in modern Russian or Polish. The backbone of any narration is a sequence of events, and these are usually conveyed in past tense by perfective verbs (see example 7a). If however a past tense narration is transferred to the present tense, imperfective verbs must do the job for perfective verbs without altering the lexical meaning (see example 7b), because the morphological present tense of perfective stems has almost entirely been driven out of present tense uses;[9] by default, it has been recategorized as (perfective) future. Compare a made-up textbook example from Polish:

(7) a. past tense: sequence indicated by perfective verbs

---

[9] Remnants exist in the inactual present (irregular habitual events, gnomic present, etc.), and Slavic languages such as Czech or Slovene are less restrictive than Russian in this respect (cf. Dickey 2000: Chapter 4-5, among many others).

> *Wacek siadł*[PFV] *w fotelu, wreszcie doczytał*[PFV] *powieść, odłożył*[PFV]
> *książkę i zasnął*[PFV].
> 'Wacek sat down in an arm-chair, at last finished [10] the novel, put
> away the book and fell asleep.'

b. present tense: sequence indicated by imperfective verbs
Wacek siada[IPFV] *w fotelu, wreszcie doczytuje*[IPFV] *powieść, odkłada*[IPFV]
książkę i zasypia[IPFV].
'Wacek sits down in an arm-chair, at last finishes the novel, puts
away the book and falls asleep.'

Furthermore, as already mentioned in §2.1, more or less tight constraints of aspect choice have encroached into other domains which, on first sight, are rather remote from actionality. These pertain to clause level or clause-combining distinctions, or to illocutionary functions. For instance, the factor [± volition-based], or [± controlled], explains most of the distribution of perfective vs. imperfective stems in the scope of modal auxiliaries or other modalized contexts. Compare, for instance, Russian textbook examples with negated possibility expressed by the auxiliary *nel'zja* 'cannot, must not': the infinitive in its scope tends to be perfective if the action is out of control of the speech-act participants, yielding a circumstantial (a.k.a. dynamic) reading: *Nel'zja rasstegnut'sja* [PFV] 'One cannot unhook' (e.g. because the zipper is broken). If, by contrast, the action can be controlled and the modality tends to be deontic, the imperfective infinitive is the default: *Nel'zja rasstegivat'sja* [IPFV] 'One must not unhook' (e.g. because I, the speaker, disallow it).

Admittedly, the distributional facts discussed in this section are framed primarily on modern standard Russian and Polish. Among Slavic languages we observe variation concerning the distribution of functions over both aspects; the details of this variation are in part considerable. However, we wanted to give an idea of the principles that describe the basic architecture of a classificatory aspect system, and which should be accounted for if the evolution of such a system is to be captured as a type of grammaticalization, though a peculiar one. Thus, at least in the northeastern part of Slavic we are observing, over the past centuries, an increasing tendency toward complementary distribution of perfective vs. imperfective stems into the predicational, clausal and utterance level. This indicates that the very principle of stem classification has been expanding from aspectual core features ([± limitation, boundedness] and [± singular situation]) to features related to modality and discourse pragmatics. The distribution in func-

---

[10] Literally 'read the novel to the end' (or even 'end-read the novel').

tional domains such as narration, directive speech acts, deontic vs. circumstantial modality, etc., was much less clear-cut in former stages and has remained a locus of inner-Slavic differentiation to this day (Wiemer 2008; 2015, with further references).

3 History of aspectual morphology from Proto-Indo-European into Early Slavic

Let us now try to reconstruct that part of the aspect story in Slavic which led to the pervasive employment of prefixes and suffixes and the initial steps towards a classificatory system.

We assume that the emergence of the Slavic aspectual system started and proceeded to an advanced stage at a period when the Slavic dialect continuum was still sufficiently homogeneous for morphological innovations to spread all over this continuum and to be carried by different Slavic varieties to locations of an expanding territory. It counts as more or less accepted knowledge that the expansion of Slavic speakers started from their homeland (somewhere between Oder and central Dnieper) during the 5th-6th century AD; by end of the 7th century AD Slavs occupied most of the Balkan peninsula and had spread further northwest to the Alps, and they reached the Volga-Oka basin in the northeast by the 9th c. AD. The invasion of the Magyars into the Pannonian plain took place in the 9th c. AD, it was one of the reasons for a division of Slavic into North and South (cf. Birnbaum 1979: 5-7, Stieber 1989: 9-11, Holzer 2014: 1123, among many others). The basis of the stem-derivational aspect system must thus have been laid by the middle of the first millennium AD, i.e. in Common Slavic. Otherwise, the perfective/imperfective opposition could hardly have installed itself in the predecessors of modern Slavic varieties based on the same morphological technique and with a shared core domain of functions (sketched in §2).

In this section we present the relative chronology of pertinent developments from Proto-Indo-European into early Slavic. We integrate findings on the development of aspect in later stages as far as these are important for the overall picture. Needless to say, the following division into Proto-Indo-European (more than 6,000 years ago), Early Common Slavic (before 300 AD) and Common Slavic (300-700 AD), early Slavic (Old Church Slavonic and early East Slavic, often also called "Old Russian") and later Slavic up to our days is idealized and glosses over a number of details (cf. Andersen 2003: 46). Note that the reconstructed verbal system of PIE represents an idealized concept. PIE was not homogeneous either in areal or in diachronic terms. We do not, however, regard areal dialectal differ-

ences of PIE as important to our argument here, and so they will not be pursued further.

3.1 The Proto-Indo-European aspectual system

Before we turn to our brief exposition of aspectual distinctions in PIE, two additional remarks are in order. First, note that in the following we will mainly rely on one of the most authoritative compendia of PIE verbal morphology, namely the LIV². We acknowledge that there is considerable disagreement with regard to the exact shape of the morphological schemata involved, issues of historical phonology (related to laryngeals, *a*-vowel, etc.) as well as etymologies of particular verbs and their present vs. aorist stem formations.[11] Having said this, only the general make-up of the verbal system reconstructed for PIE as laid out below is crucial for our purposes and not particular reconstructions, which indeed might be subject of controversy. As concerns development within the long span of PIE, our argument will relate mainly to its later stages.

Second, PIE and the old layer of Slavic verbal derivation is never purely concatenative as there are usually additional phenomena involved such as different types of vowel gradation, alternation of the thematic vowel, etc. In what follows, we refer to these complicated morphological patterns as *(derivational) schemata* (cf. Haspelmath & Sims 2010: 46–53).

Let us now consider the morphological shape of the aspectual system of PIE as reconstructed on the basis of Ancient Greek and Vedic Sanskrit (cf. LIV²). This system was very much conditioned by lexical defaults of roots; it was based on the classification of verbs into two major groups traditionally referred to as (i) *root presents* vs. (ii) *root aorists* based on lexically conditioned aspectual defaults. Morphologically, the latter verbs or, more precisely, verb stems formed the aorist (and related TAM categories) with no additional derivational marker, while the former formed the present (and related TAM categories such as the imperfect) with no additional derivational marker. This is illustrated in Table 1:

The markers used to indicate the change in the aspectual value (schematically X and Y in Table 1) are immediately attached to the root and precede the tense, mood and person/number/voice markers. This placement is an important indication for their originally derivational nature. Moreover, as can be observed from Table 1, the split between root presents and root aorists was itself not dependent

---

[11] A number of particulars are explained differently, e.g. by the Leiden School (beginning from Beekes 1969; 1995). A considerably different model of the PIE verb system is suggested in Jasanoff (2003).

Björn Wiemer & Ilja A. Seržant

Table 1: Morphological patterns for TAM formations in PIE

Time reference		Type (i) (root present)	Type (ii) (root aorist)	Mood	Person+ number
present	no augment	**root**	root + Y		present endings
imperfect	augment	**root**	root + Y	(+ Z)	non-present endings
aorist	augment	root + **X**	**root**		non-present endings

Comments: *X* – perfective schemata, *Y* – imperfective schemata, *Z* – mood schemata (zero in the indicative and imperative), *augment* – the traditional term to refer to the past tense prefix *h_1e-.

on time reference. Time reference was expressed by the presence vs. absence of the past denoting prefix *h_1e- (the so-called augment) and two sets of person-number endings: a set of presential (traditionally called *primary*) endings and a set of non-presential (*secondary*) endings. In other words, both the root presents and root aorists could form present and past tense. Root presents formed their past tense (called *imperfect*) by means of the augment and the set of non-present endings. Root aorists were not restricted to past tense use but could also occur in the present tense (injunctive, gnomic aorist) in specific contexts. The distinction between these two classes was most probably aspectual. It was obeyed in all finite and even non-finite forms (e.g., participles and infinitives based on aorist vs. present stem) as well as in different moods (e.g., aorist vs. present subjunctives). The evidence from the earliest documented Indo-European languages, such as Homeric Greek, suggests that, by and large, root presents behaved like imperfectives and root aorists like perfectives. They resembled the Slavic PFV:IPFV opposition inasmuch as this distinction (i) was inherent to all (finite and non-finite) forms of the verb and (ii) did not depend on tense or mood. Instead, tense was marked independently from the aforementioned aspectual characteristics with different sets of personal endings (for example, non-present *-*t* vs. present *-*ti* for the 3SG.ACT) and, in some varieties, with a prefix, the already mentioned augment *h_1e- (PST-) in PIE.

With this in mind, we can rename root aorists as *perfectives* and root presents as *imperfectives*, but have to emphasize that, in the context of PIE, these terms are not meant as a grammatical opposition in the same way as they are for more recent and modern Slavic (see §2). The PIE perfective:imperfective distinction of roots was probably much closer to actionality features (or, lexical aspect), but the exact semantics of this opposition remains obscure. However, as we will now see, the governing principles of this system were strikingly similar – and partly

etymologically related – to those found in Common and early Slavic.

Once the notions *perfective* vs. *imperfective* are introduced there is no need for the traditional distinction between *aorist* and *imperfect* since the former is just the past form for perfectives while the latter is the past form for imperfectives. In turn, perfectives may be simplexes or derived by means of a derivational schema. Analogically, imperfectives may be simplexes or derived from perfectives by means of some other derivational schema. For example, in order to form the perfective (such as the aorist), an imperfective simplex had to be additionally marked by some morphological marker, e.g. by the suffix *-s-* (concomitantly with the lengthened/full grade of the root vowel)[12] or, more rarely, by the reduplication of the root-initial consonant with the zero grade of the root vowel and the attachment of the thematic vowel (cf. LIV[2] 2001: 21). And, vice versa, a perfective simplex must be marked by an additional marker in order to become imperfective: e.g., by a nasal infix *-n-* (with vowel gradation), more rarely by reduplication of the root-initial consonant or by one of the suffixes, e.g. *-i̯é/ó-*, *-sḱé/ó-* or *-u-*, etc. Note again that the morphological strategies to derive imperfectives or perfectives from roots are very much derivational and not inflectional. The choice of the schema depended on actional defaults of the root, and the different schemata were not etymologically related. Originally, they were not fully synonymous and must have marked different semantic nuances (Meiser 1993; Kölligan 2004; Seržant 2014: 115). The new, derived stem behaved morphologically like a distinct lexeme.

For example, the lexical default of **deh₃-* 'give' was perfective, that is, it formed the aorist and related categories without any additional marker **h₁e-deh₃-t* (PST-ROOT-NPRS.3SG.ACT) 'she/he/it gave', while the present and related categories such as the imperfect were formed by means of reduplication with this verb **de-doh₃-ti* (REDUPL-ROOT-PRS.3SG.ACT) 'she/he/it gives/is giving' (LIV[2] 2001: 105). In turn, the lexical default of the verb **h₁ei̯-* 'go' was an imperfective and marked the present without morphological changes: **h₁ei-ti* (ROOT-PRS.3SG.ACT) 'she/he/it goes/is going'; its perfective counterpart was most probably suppletive (LIV[2] 2001: 232).

Generally, the majority of the PIE underived verbs were perfective, while their imperfective counterpart was morphologically derived by one of the schemata involving an infix, suffix or reduplication, all combined with vowel gradation (ablaut).[13] In total, around 22 productive morphological schemata were available

---

[12] Later, e.g., in Modern Greek to develop into rather an inflectional *s*-suffix.

[13] The vowel gradation patterns were highly diverse with each of these schemata: while some just required *e*-grade or zero grade of the root throughout the paradigm others involved mobile

for imperfectivization (LIV[2] 2001: 14-21). Very little is known about their original meaning distinctions. Crucially, many of them were not purely aspectual but had also bearings on semantic entailments such as Dowty's (1991). For instance, the nasal infix could mark that a participant of the given eventuality was agentive (Meiser 1993). Other schemata such as reduplication or the schemata involving the suffix *-éịe/o- with the *o*-grade of the root vowel, in turn, combined meanings pertaining to both actionality (such as pluractionality) and/or event structure (e.g., agentives, and, rarely, causatives); cf., *inter alia*, Kölligan (2004) and Seržant (2014: 115). These two were also used to derive secondary, marked unbounded verbs (see §3.2.1), most productively the suffix *-e⊠e/o-.[14]

In turn, perfectivization was quite rare in PIE, and except for one schema (suffix *-s- with root ablaut) there were perhaps one or two other options each attested extremely rarely, to the extent that their existence is somewhat hypothetical. To conclude, PIE attests primarily underived perfectives and derived imperfectives, while underived imperfectives and derived perfectives are very rare – a constellation that corresponds to the imperfectivizing-by-suffixes type in the classification which we apply, following Arkadiev & Shluinsky's (2015) typology; see our D-type in Figure 1 in §4.

3.2 Diachrony of the Slavic aspect system

In Early Slavic, aspectually relevant features unfold along two dimensions: (i) the derivational one, i.e. the opposition between two or more distinct verb stems being morphologically in a derivational relation to each other (§3.2.1), and (ii) the inflectional one, i.e. the opposition between Slavic aorist and imperfect, which is restricted to the past domain (§3.2.2). Note that derivational (i) vs. inflectional (ii) types are meant just as approximate labels; diachronically the inflectional type (ii) originated from a derivational one (i), as is argued in §3.2.2 below.

3.2.1 The derivational type: suffixes

While remaining typologically of the same type (namely, the D-type discussed in §4), early Common Slavic undergoes a number of reductive changes. First, the versatile PIE system in which lexical (i.e. actional) defaults of roots determined the application of different kinds of derivational schemata for imperfectives is

---

vowel gradation patterns dependent on the singular (active voice) vs. plural (active) and both singular and plural (middle) forms.

[14] Compare, for instance, Ancient Greek *pét-o-mai* 'I fly' vs. *potá-o-mai* 'I fly around' (LIV[2] 2001: 479)

drastically simplified. Most of the imperfective schemata of PIE are lost in Slavic: reduplication, various imperfective suffixes such as *-s̯é/ó-, *-dʰé/ó-, and many others. Other imperfective PIE schemata survive, but are no longer productive in Common Slavic, such as the nasal infix (see Table 2 below).

In turn, there are only few schemata that remained productive in the early Common Slavic period. It was primarily those that served to derive secondary deverbatives (often iteratives, habituals or duratives, but also causatives) and not primary imperfectives in PIE (such as in the first schema in Table 2). In turn, the second schema in Table 2 is most probably a remodeling of the old PIE schema to derive imperfective stems (see also §3.2.4 below.

The verbs in Table 2 are *marked unbounded verbs*, where *marked* refers to both (i) meaning and (ii) morphology: (i) Their meaning is specified as entailing lack of a boundary as opposed to the default simplex (which can be both bounded and unbounded), and (ii) they are morphologically marked as opposed to the simplex by one of the schemata adduced in Table 2. The term *unbounded* amounts to the same as the notion *atelic₂*, which was introduced and explained in §2.3. Both indicate that an eventuality is conceived of without boundaries, regardless of whether the lexeme implies an endpoint (telic₁) or not (atelic₁).

All three schemata represent remodellings of PIE schemata. While the first schema illustrated in Table 2 is no longer productive already by early Common Slavic, the second and the third schemata are variants that are productive in Common Slavic except for the morphonological ablaut, which was gradually abandoned. It is the second schema that was involved in creating the imperfect in Common Slavic (§3.2.4). In turn, the third schema in Table 2 involving the suffix *-a-je/o-* (past tense: *-a-*) remained productive into Early Slavic and gave rise to a wide range of allomorphic variants which are all, etymologically, morphological extensions thereof (see Table 3 below). Crucially, in the Common Slavic and Early Slavic period, the formation of various modifications of actionality was still highly lexicalized and by no means regular, and a number of simplexes did not have any pluractional or durative correspondent, e.g. *bъra-ti* (take.NPRS-INF) 'take' whose pluractional *-bira-ti* (take.IPFV.NPRS-INF) appeared only later and only in opposition to the respective prefixed verb, e.g. *sъ-bъra-ti* (PFX-collect.NPRS-INF) ⇒ *sъ-bira-ti* (PFV-collect.IPFV.NPRS-INF) 'collect'; cf. Maslov (2004[1959]) and further in §3.2.3.

While most of the schemata based on suffixation coded unbounded situations, there was only one exception to this, namely, the nasal suffix *-nu-* (i-ii) and the nasal infix *-n-* (iii). There are generally three functional types here: (i) gradual accomplishments, as in, e.g. Russ. *vja-nu-t'* 'wither', *sox-nu-t'* 'dry [INTR]'

Björn Wiemer & Ilja A. Seržant

Table 2: Old, non-productive layer of the Common Slavic schemata to yield *marked unbounded verbs*

Present	Non-present	Ablaut	comments and examples
*-e(j)e/o-	*-i-	-o-	Examples of these derivations are rare, cf. (non-present) *vod-i-* vs. (present) *vož-* (< *-j-) derived from the simplex *ved-* 'lead'; (non-present) *laz-i-* vs. (present) *laž-* (< *lōz-j-) derived from the simplex *lĕz-* (<*lēz-) 'climb', etc. This is an inherited PIE derivation to yield marked unbounded verbs, which became unproductive already in Common Slavic.
*-je/o-	*-(j)a-	(-e-)	An old derivation with sometimes imperfectivizing function, e.g. (non-present) *ima-* vs. (present) *jeml'-* (< *em-j-) 'grasp, take' derived from the simplex *ję-ti* (< *im-ti) 'grasp, take'; (non-present) *da-ja-* vs. (present) *da-j-* 'to give' from the simplex *da-* 'give' or (non-present) *kry-ja-* vs. (present) *kry-j-* 'cover' from the simplex *kry-* 'cover' (Silina 1982: 164f.). This derivation was crucial for creating the Slavic imperfect (Ostrowski 2006; see §3.2.4).
*-a-je/o-	*-a-	long grade	Iterative and durative Aktionsarten were formed by means of the lengthened grade of the root vowel and the suffix *-ā-, cf. Russian *po-lož-i-t'* (historically *-log-ej(e)-) 'put' vs. *po-lag-a-t'* (*-lōg-ā-); (non-present) *na-bira-* vs. (present) *na-bir-aj-* with lengthened root vowel *-bīr- from *na-bьra-* (< *-bir-) 'take, collect' (Silina 1982: 167f.). Ivanov (1964: 382) considers this to be an inherited pattern.

(this has yielded imperfective stems), and (ii) semelfactives, cf. Russ. *tolk-nu-t'* 'push once', *stuk-nu-t'* 'knock once' (which now belong to perfective stems); cf. SłPrasłow (Sławski 1974: 45) and Nesset (2013) for the diachronic relations. The original form of the nasal suffix was **-nVn-*. This very suffix – albeit old – is not inherited as such and must be a Slavic (and, perhaps, Germanic) morphological and functional remodeling of the old PIE infix **-n-*. Finally, (iii) there are remnants of the old PIE pattern with this infix **-n-* such as OCS *sęd-ǫ* 'I take a seat', *lęg-ǫ* 'I lie down' from Common Slavic **sē-n-d-* and **le-n-g-*, all with an ingressive meaning denoting the onset of (or transition into) some new state (Ivanov 1964: 383; Rasmussen 1988).

While (i) adheres to the general tendency to relate suffixation with marked unboundedness, (ii) and (iii) clearly denote bounded situations. The nasal suffix *-nu-* (ii) and the archaic infix **-n-* (iii) were the only schemata to derive stems marking limitation via suffixation. Later, these types were integrated into the patterns of aspect pairs and entered into the class of perfective verbs.

Generally, we observe a clear tendency towards *concatenation* that developed from the earlier schemata. The beginning of this process predates Common Slavic, continues into Early Slavic and reaches up to the modern Slavic languages, in which it is still not fully accomplished. Different kinds of modifications of the old suffixes (the second and third schema in Table 2 above) took place, while concomitant morphonological co-effects such as the lengthened grade of the root or the ablaut were abandoned:

Further modifications of these suffixes are found. Old and Middle Russian *-e/o-* (present) / *-a-* (non-present), *-je/o-* / *-ja-*, *-aje/o-* / *-a-*, *-jaje/o-* / *-ja-*, *-vaje/o-* / *-va-*, *-ovaje/o-* / *-ova-* were mostly compatible with contexts associated to unboundedness only (Silina 1982: 162). Crucially, all these suffixes draw on the old second or third schema in Table 2 above.

The schemata in Table 2 and Table 3 played an important role in the rise of the new aspectual system (Maslov 2004[1959]; *Le Slave commun. 2. éd. rev. et*

---

[15] The earliest attestations of *-iva-* in East Slavic are from the 12th c. (Silina 1995: 377; Ševeleva 2010). This suffix became considerably more frequent in the 14th c. and reached a peak of productivity in the 16-17th c. (Andersen 2009: 131). It outranked the other most salient suffix *-ova-* in East Slavic (Vaillant 1966: 492, 499; Mende 1999: 314 referring to Silina 1982: 170-176); see also §3.2.3. The suffix *-ova-* had originally denominalizing function (Vondrák 1924: 718; Vaillant 1966: 488; SłPrasłow (Sławski 1974: 48)). Its capability of deriving imperfective stems in more recent times might have been facilitated by the fact that desubstantival verbs usually have been integrated into imperfective aspect (Miklosich 1926: 486). The suffix *-iva-* (with allomorph *-yva-*) originates on the basis of already established *-Vva-* and verb stems with *-i-* as thematic vowel (Kuznecov 1953b: 262; Vaillant 1966: 490).

Björn Wiemer & Ilja A. Seržant

Table 3: Recent layer of the Slavic marked-unbounded schemata

Present	Non-present	Ablaut	comments and examples
-jaje/o-	*-ja-*	root vowel lengthening (rare)	This is a recent development (productive after 11th c.) based on the extension of the second schema with *-j-* stemming from verbs with the stem in *-i-*, e.g. East Slavic (non-present) *iz-bavl-ja-*, (present) *iz-bavl-jaj-* (<*-bav-jaj-*) derived from *iz-bavi-* 'save from sth.'.
-(V)vaje/o-	*-(V)va-: yva, iva, ova,* etc.)	root vowel lengthening (rare)	These are the most recent suffixes, although they are sporadically attested in early East Slavic (e.g. in the Laurentian Codex from 1377), they became productive in the 13-17 cc. (Ivanov 1964: 387; Andersen 2009: 131) and are formed analogically to the Old Slavic patterns in *-Vva-* (Kiparsky 1967: 212f.). It is obvious that this suffix series historically derives from the second schema in Table 2 by adding *-v-* and a vowel. The latter are originally parts of the root of some verbs which were reanalyzed as belonging to the suffix and then generalized.[15]() The meaning pertains to multiple actions, cf. *kupit'* 'buy' vs. *koup-l-iva-lъ* '(every time) he bought'.

264

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

9 Diachrony and typology of Slavic aspect: What does morphology tell us?

augm. Nouveau tirage avec le concours de A. Vaillant). The number of schemata has considerably diminished from PIE times, and, concomitantly, their morphological make-up changed from schemata causing stem-internal morphonological changes to concatenative suffixation, creating thus morphologically more transparent derivation.

3.2.2 The derivational type: prefixation

Not much can be said about the chronology in which prefixes appeared in Slavic. Certainly, prefixes did not exist in PIE and preverbs developed out of PIE adverbs or nouns (cf. Chantraine 1953: 82; Cuzzolin, Putzu & Ramat 2006; DeLazero 2012). Closely related Baltic attests traces of a looser morphotactic juncture of preverbs; cf. Lith. *per-si-kel-ti* 'through-RFL-raise-INF' meaning 'move (to another place)'. Here, the reflexive marker *-si-* is inserted between the prefix *per-* and the verb root *kel*, hinting at an earlier adverbal nature of *per-* to which *-si-* was cliticized.[16] We are unaware of any attestations of this phenomenon in documented Slavic material. But this observation can at best be interpreted as an indication that coalescence with the stem was finished earlier in Slavic than it was in Baltic; we cannot, however, induce from this at which period preverbs became rigidly tied to the verb stem, thereby turning into prefixes.

If we turn to the semantic side of prefixation, it is commonly assumed that, originally, preverbs (and thence prefixes) were used to modify the lexical meaning of simplex stems. We may call this *verbal orientation* (not only in a spatial sense), as proposed in Plungjan (2000: 176, 291; 2002). That is, the semantic development responsible for the conventionalization of the telic$_1$ meaning of the prefixed stems can to some extent be explained as the effect of local expressions providing an inherent endpoint of a particular situation in space; compare, for instance, OCS *i-ti* 'go', which can have either atelic$_1$ or telic$_1$ reading, and *vъn-i-ti* (inside.PFX-go-INF) 'go in, enter', *prě-i-ti* (across.PFX-go-INF) 'go over, cross', etc., which are only telic$_1$ (Maslov 2004[1959]; Silina 1982: 163; Bermel 1997: 466, among others). We assume that local (and comitative) prefixes/adverbs could have a considerable degree of abstractness early after the split of PIE, so that their abstract meaning should have developed also by, or during, the Common Slavic period. This development subsequently allowed these prefixes to focus on

---

[16] In the literature this has been discussed under the rubric of tmesis phenomena. In older stages such phenomena occurred with other enclitic pronouns as well (and survived in some dialects); e.g. *ap-**mi**-švies-k akis* 'illuminate **my** eyes', lit. 'illuminate the eyes **on/to me**' [Germ. *erleuchte* ***mir*** *die Augen*]' (cit. from Rosinas 1995: 10f. orthography slightly adapted; cf. also Ambrazas 2006: 83-87).

the telicity$_1$ effect and less on local or other meaning components. Thus, we may safely assume that *the first step* (= stage (i) in Table 4 below) in the rise of the aspectual functions of prefixation was its ability to code telicity$_1$ in opposition to functionally unmarked simplexes. At this stage, both simplex and the prefixed verb could be used in both perfective and imperfective core contexts (Maslov 2004[1959]). There emerged thus an asymmetry between the simplex and the respective prefixed verb: the former could be construed as both telic$_1$ and atelic$_1$, the latter could only have the telic$_1$ meaning.

Furthermore, prefixes can serve to emphasize a semantic component that is already inherent to a simplex stem; compare Russ. *nes-ti* $^{\text{IPFV}}$ 'carry' vs. *pri-nes-ti* $^{\text{PFV}}$ 'carry (toward a reference point)'. In particular, they are able to highlight a boundary of the verbal action which otherwise would remain only implied (= telic$_1$) or can even be suppressed (= atelic$_1$). This phenomenon is known as the Vey-Schooneveld effect.[17] Essentially, it says that alleged "empty prefixes" do have a semantic function, namely: their choice is motivated by, or harmonic with, some lexical component of the meaning of the simplex stem, in particular a component introducing some sort of boundary to the denoted action. We assume that the Vey-Schooneveld effect was an important mechanism facilitating the development of the prefixes primarily coding telicity$_1$ and, subsequently, telicity$_2$ in opposition to simplexes.

Telic$_1$ verbs in general show a default focus on the endpoint as having been attained, especially in the past tense, or in narration.[18] With time the focus on

---

[17] The name comes from two pioneering articles (Vey 1952; Schooneveld 1958), whose significance has recently been re-discovered in connection with the description of contemporary Slavic (in particular Russian) aspect by Janda et al. (2013), among others, and been used for an adequate assessment of the role of preverbation in the evolution of aspect systems, above all by Arkadiev [Arkad'ev] (2015).

[18] Converging evidence supporting the existence of such a default has been provided from usage-based, morphological, and typological findings. First, in first language acquisition children start using telic$_1$ verbs predominantly in the past to denote accomplished actions and resultant states. This apparently holds regardless of whether the language has aspect (e.g. Russian, Chinese, French) or not (e.g. German, Swedish); cf. among others, Schlyter (1990); Lehmann (1992), Stoll (1998; 2005: 806), Gagarina (2004), with further references. In Russian and Polish, children acquire imperfective stems of telic$_1$ verbs later than perfective ones (Lehmann 1990; Gagarina 2004). Second, in German, an aspectless language, telic$_1$ verbs in the past are associated primarily with an achieved goal, not with the preceding process (e.g. *Er öffnete ein Fenster* 'He opened a window', *Sie bauten ein Haus* 'They built a house'). This phenomenon has also been dubbed 'default aspect' (cf., for instance, Bohnemeyer & Swift 2004). Third, Russian imperfective stems of telic$_1$ verbs are predominantly derived from perfective stems (via secondary suffixation, type frequency) and they are also less frequent as tokens in the past (Breu 1980; Lehmann 1993; Gagarina 2004).

the realized boundary turned from an implicature into a conventionalized telic$_2$ meaning of the prefixed stem that no longer can be cancelled.[19] This made up *the second step* (= stage (ii) in Table 4) toward a grammatical perfective/imperfective opposition. It favored the strengthening of the functional distribution of prefixed vs. simplex stems over contexts associated with perfective and imperfective aspect.

The strengthening of the telic$_2$-interpretation and, thus, a split between telic events and telic processes shows that, at this stage, non-punctual telic$_1$-verbs (which correspond to Vendlerian Accomplishments) played an important role. These verbs are different from other actional types, such as activities or achievements, in that they consist of two components each of which may be separately highlighted in a particular discourse move: (i) the preparatory activity (e.g. the process of *building*) and (ii) the culminating achievement (as in *A house was built*). To this extent, these verbs are lexically ambiguous. On the basis of this ambiguity the emergent telic$_2$ function of prefixes could gain in significance, helping to morphologically highlight the culminating-achievement component as opposed to the preparatory-activity component of the meaning (cf. *inter alia*, Maslov 2004[1959]; Breu 1992; Bermel 1997). Notably, it is this aspectual class of verbs which appears to be the first in nascent "Slavic-like" systems, for example, in Gothic or Old Irish (see §5.2-5.3) as well as in Latvian (see §5.4).

In fact, this process complements the Vey-Schooneveld effect: the prefix not only emphasizes an already existent lexical component of the verb, it also makes more salient the default focus on the implied endpoint as being reached. The result of this cooperation of prefix functions was the conventionalization of the telic$_2$ implicature. In turn, the simplex stems in the course of time were predominantly (though not exclusively) relegated to imperfective aspect as they remained capable of denoting anything else beyond the telic$_2$ function. Furthermore, this distribution was then transferred to other patterns of aspectual pairings, first of all with secondary imperfectivization (on which see §3.2.3).

In *a further step* (iii), prefixation started being applied to atelic$_1$-stems as well, first of all to activities, i.e. to eventualities which do not entail an inherent endpoint. In this case the perfectivizing function consisted only in setting a temporal boundary as in contemporary Russian, e.g. *po-rabotat'* 'work for a while'. This brought about two consequences. First, the telic$_2$ function of perfective verbs

---

[19] Cf. Breu (1992: 128f.). Strictly speaking, the simplex remains compatible even with a telic$_2$ meaning, but its prefixed counterpart begins replacing it increasingly in this meaning (other conditions, e.g. [± repetition], remaining equal). The prefixed stem is marked in comparison to the simplex stem both morphologically and functionally, since its contexts of usage have become more restricted.

Björn Wiemer & Ilja A. Seržant

was, in a sense, loosened, because prefixes became able to mark delimitation even with stems that did not imply any inherent boundary (atelics$_1$). Eventually, the function of perfective verbs (most of them prefixed) was fixed to focus on boundaries, be they inherent or only temporal delimitations. Second, the prefixed atelic$_1$-stems were reinterpreted as events and started sharing central functions of perfective verbs with telic$_2$-stems. For instance, Russ. *po-leža-l* ^{PFV} 'lay for some time' can be inserted into narrative sequences together with perfective telic$_2$-stems, e.g. *otkry-l* ^{PFV} *xolodil'nik* 'opened the fridge' or *s"e-l* ^{PFV} *salat* 'ate (up) a salad' (cf. Bermel 1997; Lehmann 1999, 2009, among others). See the discussion in §2.2 and §2.5.

Table 4: Functional development of prefixation in early Slavic

0	**verbal orientation**: spatial, etc.	**lexical**	**Early Common Slavic**
(i)	**telicizing$_1$** function: 1. coding telicity$_1$ on the verb (connected to Vey-Schooneveld effect) 2. emphasis on lexically inherent boundary	↓	**Common Slavic**
(ii)	**telicizing$_2$** function: focus on the achievement of the inherent boundary	**actional modification** ↓	**Common and Early Slavic**
(iii)	**limitation** (temporal or natural)	**viewpoint aspect**	**later Slavic, but predominantly in the eastern half**

We may sum up so far. By and large, one can justifiably assume that the role of prefixes proceeded along the stages of functional development (Table 4). The comparative recency of stage (iii) is reflected in the fact that it is less prominent in the western part of Slavic, insofar as merely temporal limitation is concerned (Dickey 2000: Chapter 7; 2011). Support for the assumption that the preceding stages (i-ii) must have advanced considerably earlier comes, for instance, from Eckhoff & Haug (2015: 202-207). In their corpus study on Codex Marianus and

268

Codex Zographensis (10-11th c.), written in Old Church Slavonic, they observed a strong association between prefixed stems (without further suffixation) and contexts of perfective aspect, on the one hand, and between suffixed stems and contexts of imperfective aspect, on the other (although the latter association was less pronounced).[20] Many simplex stems remained underspecified in this respect and were used readily with either aorist or imperfect. A similar situation was observed in some of the oldest texts of the East Slavic recension of OCS (Seržant 2009). There was a stable opposition between nascent perfectives and nascent imperfectives (if judged from modern Russian) in such a way that nascent perfectives were used only with the aorist, perfect, and the passive based on *n/t*-participles, and they denoted future though being realized in the present tense stem, while nascent imperfectives (i.e. both simplex and secondarily suffixed stems) were used with imperfects and in the passive based on *m*-participles, and being coded in the present tense stem they also had present tense reference. By way of example, compare *znaje-**m**-ъ by-st ь* (know-**PASS.PRS**-NOM.SG.M be-AOR.3SG) 'he was known' vs. *po-zna-**n**-a by-stъ* (know-**PASS.PST**-NOM.SG.F be-AOR.3SG) 'she was recognized'.[21] The originally resultative *n/t*-participles allowed only telic$_1$ verbs as their input (Brugmann 1895; Havránek 1937: 101f.; Haig 2008: 41; Seržant 2012: 359–361), while the present passive participles in *-m-* were compatible with progressive meaning with atelic$_1$/atelic$_2$-verbs as their input. In turn, only a few unprefixed and unsuffixed verbs remained indifferent to this distinction in this corpus. These findings may be summarized as in Table 5.

Table 5: Strong preferences of the verbs in tense and voice formations in the oldest Russian Old Church Slavonic (according to Seržant 2009)

	Past tense	Passive	Present tense
nascent perfectives	aorist, perfect	based on -*n/t*-participles	future in the Greek translation
nascent imperfectives	imperfect	based on -*m*-participles	present in the Greek translation

Observations like these make us inclined to assume that a system as in (8) developed into a system as in (9):

---

[20] Remarkably, the parallel Greek texts did not show such a strict correlation, since many Greek prefixed verbs were used in the imperfect (Eckhoff & Haug 2015: 202).

[21] This resembles the situation encountered in modern Lithuanian (Seržant 2009: 321f.).

Björn Wiemer & Ilja A. Seržant

(8) Early Common Slavic
     simplex stems (default) vs. *suffixed stems* (marked unbounded meaning)

(9) Late Common Slavic
     simplex stems (default) vs. *suffixed stems* (marked unbounded meaning)
     vs. *marked prefixed stems* (telic$_2$)

The aspectual behavior of the simplex stems was the least stable one.

3.2.3 Secondary imperfectivization

We are now entering into a period when inner-Slavic differentiation started becoming more pronounced, both in terms of the productivity of patterns of aspectual derivation and in terms of the functional distribution of (nascent) perfective and (nascent) imperfective stems. These differences have since then partially been accompanied by the different fate of the older aorist:imperfect opposition (on which see §3.2.4).

In §3.2.2 we have argued that prefixation was increasingly related to telicity$_1$ and, subsequently, telicity$_2$. They, thus, marked situations as bounded, while suffixation coded unbounded eventualities. Gradually, in the Early Slavic period, the semantic markedness of suffixation bleached and the latter became the unmarked option for expressing unbounded situations of all sorts and, eventually, even progressive and other functions typically associated with imperfective aspect. Simultaneously, prefixed verbs not only gradually became telic$_2$ and, by this property, constituted the class of perfective verbs, but they also started losing the capability of denoting iterative/habitual/generic meanings. Probably, this process started earliest in the northeastern part of Slavic; in the western half of Slavic these functions have remained robustly attested with perfective verbs until today.

Most prefixed stems lexically differ from their simplexes; compare, for example, *sъ-bъra-ti* (together.PFX-take-INF) meaning 'collect, gather' (lit. 'take together') against *bъra-ti* 'take'. Since the meaning range of prefixed stems began to shrink in the domain of iterative and progressive functions, a gap resulted when these functions were to be expressed with lexical concepts that were denoted by prefixed, and therefore perfective, stems. This gap was, as it were, filled by the suffixation patterns as discussed above, i.e. by creating new, secondary imperfectives to the prefixed stems; compare, e.g., Old East Slavic *prě-bi-va-ti* [IPFV] (through.PFX-hit-SFX-INF) from *prě-bi-ti* [PFV] (through.PFX-hit-INF) 'break (through)'; see the third pattern in Table 2. Although secondarily suffixed stems were attested already in OCS, the class of imperfective verbs started filling up steadily

with such stems. Simplexes remained, in turn, ambivalent for a long period of time even in the northeastern part of Slavic. According to Ševeleva (2010: 208–212), secondary imperfectives marked with *-iva-* were already quite productive in 12th c. East Slavic, and according to Andersen (2009: 131), secondary suffixation experienced a steep rise of frequency from the 13th c. onwards (see also Footnote **??**). He describes this increase as an S-curve, whose flat beginning started a long time before: "The bottom part of that curve would correspond to innovations that occurred in prehistory." (2009: 138) Although these findings, again, concern primarily the northeastern part of Slavic, it is indicative of the general line of development for Slavic aspect as a whole.

Morphologically, secondary imperfectivization is based on the same suffixes already discussed and illustrated in §3.2.1 above (Table 2 and Table 3). As has been mentioned, the old schemata (Table 2) became unproductive and were superseded by more transparent ones (Table 3), showing a tendency toward concatenation. Moreover, the number of productive suffixes decreased.

Among the suffixes mentioned in Table 3, the suffix *-iva-* has became the most productive means of secondary imperfectivization in Russian and Polish, whereas traditional Belarusian and Bulgarian have kept *-va-*; the West Slavic languages except Polish prefer *-ova-*. However, the productivity and functional range with which these suffixes are applied in different Slavic languages varies a great deal. By and large, the eastern part of Slavic appears to show more propensity toward secondary imperfectivization (with Bulgarian as the "leader"). Productivity of secondary imperfectivization, in turn, seems to correlate with a specialization of secondary imperfectives in the domain of iteration (cf. Arkadiev [Arkad'ev] 2015: 122-125 for an overview and references). The iterative meaning was facilitated by the opposition to the respective simplex in cases where the simplex stem had been preserved and the prefix acted as a telicizer$_2$, but did not modify the lexical meaning of the deriving simplex. As a consequence, there were two imperfective stems (the simplex and the secondary suffixed one) relating to the same prefixed perfective stem without a change of lexical meaning. This situation holds on up to today;[22] compare, for instance, Russ. *maza-t'* $^{\text{IPFV}}$ $\Rightarrow$ **na**-*maza-t'* $^{\text{PFV}}$ $\Rightarrow$ *na-maz-***yva***-t'* $^{\text{IPFV}}$ 'smear (e.g. butter on bread)'.

Till now, we have been concentrating on the rise of derivational patterns responsible for the perfective:imperfective opposition in Slavic. However, in order to more fully understand the global significance of these patterns over largest possible stretches of time, we have to dwell on the appearance of the imperfect as well. It is a genuinely Slavic innovation, whose roots into earlier Indo-European

---

[22] In Slavic aspectology this phenomenon is known under the name of *trojki*, i.e. triples.

can be found among the suffix schemata discussed in §3.2.1. As we argue below, it acquired inflectional character via analogical expansion from one of these schemata. This technique of stem extension is probably older than the sources of prefixes discussed in §3.2.2.

3.2.4 The inflectional aspectual opposition: the imperfect

While the Slavic imperfect is an innovation etymologically unrelated to the PIE imperfect,[23] the Slavic aorist largely continues the PIE aorist morphology with some few exceptions. However, this is not particularly telling and morphological inheritance does not correlate here with functional inheritance. The loss of the PIE imperfect led to the loss of the PIE aspectual opposition aorist vs. imperfect altogether. It is thus likely to assume that there was a period of Common Slavic when there was only one past tense form, namely, the later aorist (Andersen 2013: 21).

Indeed, the thematic aorist formations historically contain a number of old, PIE imperfects (Leskien 1919: 195; Pohl 1971: 352). After the new Slavic imperfect was created, the aspectual division of labor must have been re-apportioned, making something that was originally simply a general past (Forsyth 1972) into an aorist. Although this path might sound complicated, similar cases of functional development are found elsewhere, for example, in English. Here, the only available form, the simple past (*He drank wine*) narrowed down its meaning to exclude the progressive meaning, which is now served by the respective continuous forms (*He was drinking wine*). Originally, the simple past was able to express also progressive meaning; compare, for instance, Norwegian *Han drakk* or German *Er trank* 'He drank/was drinking'. Of course, the innovated imperfect:aorist distinction in Common and Early Slavic differed from the English simple:progressive opposition in that the latter applies to all tense levels, not only to the past. However, the logic of redistributing parts of the actionality domain when a new aspect gram emerges is basically the same.

Yet, the question is how the imperfect emerged in Common Slavic. The Slavic imperfect *mьn-ě-(j)a-xъ* 'I thought' or *glagol-a-(j)a-xъ* 'I spoke / was speaking' is synchronically formed from the aorist stem (*mьn-ě-* or *glagol-a-*) or from the infinitive stem (both are most frequently homonymous) by means of the marker *-(j)a-* with a set of imperfect endings somewhat different from the ones of the aorist (2sg, 3sg, 2pl, 3pl); cf. Pohl (1971: 359).

---

[23] The reader may be reminded that the PIE imperfect was formed as the past tense of an imperfective (*root present*) stem by means of the past tense prefix and the non-present person-number-voice attached to this stem (Table 1, §3.1 above).

Since Franz Bopp this marker has traditionally been considered to have emerged from a periphrasis with an auxiliary (most accounts assume some form of the verb 'be') and some deverbal noun/participle that subsequently univerbated (Leskien 1919; Stang 1942: 82; Kortlandt 1986: 253ff. Lühr 1999), assuming somewhat unusual word order: participle-auxiliary. Since the initial sound(s) of the auxiliary merged with the final syllable of the alleged deverbal noun/participle, there is space to hypothesize about the exact phonological and morphological shape of the alleged auxiliary. Thus, positing an original auxiliary *$\bar{e}som$ 'I was' explains where the dedicated set of imperfect endings came from, since the latter are neither aorist nor present (nor old perfect) endings. The major difference between the aorist and imperfect ending sets is the absence in the former, but presence in the latter, of a thematic vowel *-e/o- between the aorist/imperfect marker -x- (< *-s-) and the old past-tense endings.[24] This sequence of *-s- and the thematic vowel *-e/o- is indeed found in the alleged *$\bar{e}s$-o-m.

Unfortunately, this traditional explanation faces a number of problems, one of which is that it crucially hinges upon the form of the auxiliary *$\bar{e}som$ 'I was', which as such is not attested anywhere in Slavic (or closely related Baltic) and most probably never existed.[25] Furthermore, it seems that the distinction between thematic and athematic endings is rather one of allomorphs and not of etymologically distinct morphemes. Indeed, athematic imperfect endings are also found, e.g. in the dual -sta (2DU) and -ste (3DU) alongside thematic -*šeta* (2DU), -*šete* (3DU) (Pohl 1971: 349) as well as in all imperfect forms of the verb *by-ti* 'be' with the stem *bě-* for which the traditional periphrasis account sketched above does not offer any explanation. Finally, this model does not account for the morphological shape of the stem of the lexical verb underlying the imperfect (Pohl

---

[24] The only exception is the first person containing the thematic vowel *-o- also in the aorist, cf. both aorist and imperfect: -*xъ* (SG)/-*xomъ* (PL) < *-s-o-m/-s-o-mos. The derivation from a PIE imperfect *e-h₁es-o-m would indeed explain the thematic vowel. However, since the aorist ending set equally attests the thematic vowel in the first person singular and plural, it is more likely to assume that two different sets of endings for the past tense existed in Common Slavic: the ones based on the suffix *-s- with no thematic vowel and the ones with *-s- and the thematic vowel.

[25] This form is allegedly, a morphologically somewhat modified IE imperfect *e-h₁es-o-m (as can be deduced from Leskien 1919: 196) or perfect *h₁e-h₁ós-e (Stang 1942: 82; Kortlandt 1986: 253). The former assumes the past-tense prefix *e- (the augment) which is attested nowhere else in Slavic, nor in the closely related Baltic languages, and is therefore highly hypothetical (Pohl 1971: 349). The latter is equally hypothetical because the IE perfect of the verb 'be' was formed on the basis of the suppletive stem *bueh₂- (Slavic *by-*); moreover, the perfect reduplication is equally unattested in Slavic and Baltic. Finally, the PIE perfect had a different set of endings, which are attested in early Slavic (with the verb *vid-/věd-* 'see, know') and hence would be expected to appear in the imperfect as well if this account were correct.

Björn Wiemer & Ilja A. Seržant

1971: 349f.).

Since the periphrasis-based approaches face some quite unsolvable problems, other scholars have advocated a derivational approach (*inter alia*, Kuryłowicz 1937, 1960, Bech 1971; Arumaa 1985: 285). Maslov 2004[1954]: 142f.) suggested that there must have been some relation of the suffixes of the imperfect *-ě-* and *-a-* with the same suffixes found in aktionsart derivations such as *sěd-ě-ti* 'be seated, sit' vs. *sěd-* (in *sěs-ti*) 'sit down', *bъd-ě-ti* 'be awaken' vs. *budi-ti* 'waken s.o.' and *im-a-mъ* 'I have' vs. *jeml-ju* 'I take', *plav-a-ti* 'swim, drift' vs. *plu-ti* 'idem', etc.

A breakthrough in the discussion between the periphrasis-based and derivational approaches was achieved by Ostrowski (2006), who independently from Slavic data identified the morphological pattern for marked unbounded verbs (derived pluractionals, duratives, iteratives, etc.) in the closely related Baltic languages. In Baltic, the pattern **-j-e/o-*[26] (present tense) / **-jā-* (past tense) is found to mark stems denoting unbounded eventualities. Morphotactically parallel to the Slavic imperfect, this pattern derives unbounded verbs from the past tense stem of the verb. Recall that the Slavic imperfect is mainly derived from the aorist stem of the verb, which was originally the default past stem. Consider the examples from Lithuanian and Latvian in Table 6.

This strategy both morphologically and semantically equals the one found in the Slavic imperfect except, of course, for the person-number desinences. Moreover, there are even one-to-one correspondences between the past form of the Baltic marked-unbounded verbs and the Slavic imperfects (Ostrowski 2006) (Table 7).

Other parallels can be added: Baltic and Slavic **tek-ē-jā-* found in OCS *teč-a-**a**-xomъ* (flow- NPRS-IMPF-1PL) and Lith. *tek-ė-jo-m*, Latv. *tec-ē-**jā**-m* (flow-NPRS-UNBOUND.PST-1PL) 'we flew/were flowing'; Baltic and Slavic **eisk-ā-jā-* found in OCS *isk-a-**a**-xomъ* (search-NPRS-IMPF-1PL) and Lith. *iesk-o-**jo**-m* (search-NPRS-UNBOUND.PST-1PL) 'we searched/were searching'. Although the two suffixes used to form marked unbounded verbs, namely, the suffix **-ē-* or **-ā-* for past/aorist (Stang 1966: 387; Pohl 1971: 356) and the suffix **-jā-* for unboundedness were originally distinct suffixes, they tended to merge into one conglomerate suffix **-ējā-/*-ājā-* in both Baltic and Slavic; compare imperfect forms in *-ě(j)a-xъ* and *-a(j)a-xъ* in Slavic and verbs in *-ējā-/-ājā-* in Baltic.[28] From this derivational his-

---

[26] Baltic, but not Slavic, has generalized the thematic vowels **-e/o-* into **-o-*; in modern Baltic we have *-a-* throughout.

[27] Lithuanian *o* is historically **ā*.

[28] This was additionally facilitated by the change in Common Slavic from **ē* to **ā* after palatal

Table 6: The original derivational pattern of semantically and morphologically *marked-unbounded* verbs (pluractional, durative, etc.) in Baltic

Infinitive	Past tense stem	Marked unbounded (pluractional, durative, etc.)
Lith. *py-ti* 'begin to give milk'	Lith. *pij-o-* (**pij-ā-*) 'began to give milk'	Lith. *pij-o-jo-* (**pij-ā-jā-*) 'was giving milk / gave milk (repeadedly), etc.'
Lith. *gul-ti* 'lie down'	Lith. *gul-ė-* (**gul-ē-*) 'lay down'	Lith. *gul-ė-jo-* (**gul-ē-jā-*) Latv. *gul-ē-jā-* 'was lying / lay repeatedly, etc.'
Lith. *min-ti* 'remember'	Lith. *min-ė-* (**min-ē-*) 'remembered'	Lith. *min-ė-jo-* (**min-ē-jā-*) Latv. *min-ē-jā-* 'was mentioning / mentioned (repeatedly), etc.'

Table 7: Morphological and etymological correspondences between Slavic imperfects and Baltic (Lithuanian) marked-unbounded verbs in the past (Ostrowski 2006)

Past tense	Marked unbounded
znā- OCS. *zna* 'know.AOR.3SG' Lith. *pa-žino*[27] 'PRF-know.PST.3'	***znā-jā-*** OCS. *zna-a-še* 'know-IMPF-3SG' Lith. *žino-jo-me* 'know-UNBOUND.PST-1PL'

torical explanation of the Slavic imperfect, two conclusions immediately follow: (i) It is the stem of the imperfect forms itself which carries the grammatical semantics of the imperfect, not the endings based on *-x-*, whose function is to refer to the past tense. (ii) The Slavic imperfect alongside its Baltic counterparts is of derivational origin and its inflectional status in Early Slavic is secondary.

---

consonants, yielding a merger of **-ējā-/*-ājā-* into **-ājā-*, and, as shown in Ostrowski (2006), some analogical restructuring of Baltic **-ē-* and **-ā-*-based past tenses. Compare, for instance, Old Lithuanian *isch-tirr-**a*** (**iš-tīr-**ā***), which corresponds to modern Lithuanian *iš-tyr-**ė*** (**iš-tīr-**ē***) 'examined'.

Björn Wiemer & Ilja A. Seržant

Moreover, as argued in Seržant (2008: 314), Slavic itself attests this derivational pattern as well. Compare Old Russian *da-ja-ti* 'give' with the imperfect *da-ja-xu* (3PL) and the present *da-j-utъ* (3PL) originally derived from the simplex aorist *da-* (e.g., OCS *da* 'give.AOR.2/3SG'), but also *li-ti* 'pour' vs. *li-ja-ti, dě-ti* 'do' vs. *dě-ja-ti, kry-ti sę* 'hide [INTR]' vs. *kry-ja-ti sę, sta-ti* 'stay' vs. *sta-ja-ti*, etc. Interestingly, while the Old Russian dictionary (Sreznevskij 1893-1912: 635) lists the imperfect *da-ja-xu* (3PL) in the lexical entry of *da-ja-ti*, it is, at the same time, the regular imperfect form of the simplex *da-ti*. In the same way, the imperfect *dě-ja-še* (3SG) may be just the past tense of *dě-ja-ti* or the imperfect of *dě-ti*, etc. In other words, the derivational pattern *-je/o-* (present tense) / *-jā-* (past tense), established by Ostrowski (2006) independently for Baltic, re-occurs here in the following Old East Slavic verbs: the present *da-j-etъ* (3SG) / *da-j-utъ* (3PL) is historically *dā-je-ti* (3SG) / *dā-jo-nti* (3PL), whereas the past of derived atelics, alias imperfect, is historically *dā-jā-*.

The imperfect thus is attested in present tense and infinitive forms for some verbs in Slavic: the imperfect *da-ja-xu* (give-IMPF- IMPF.3PL), the present *da-j-utъ* (give-IMPF- PRS.3PL) and infinitive *da-ja-ti* (give-IMPF-INF). Finally, the semantics of the imperfect confirms this analysis: both the imperfect and verbs in *-je/o-* (present) / *-jā-* (past) signify marked unbounded meanings. Following Ostrowski (2006) and amendments by Seržant (2008), we assume that the Slavic imperfect, e.g. *da-ja-xu,* is just the marked unbounded derivation restricted to past tense for many verbs while some few Old East Slavic verbs just mentioned (*da-j-a-ti* 'give', *kry-j-a-ti sę* 'cover') still attest the whole paradigm. The derivational nature is independently confirmed by the cognate derivation in Baltic. We therefore assume the following system for Common Slavic (and, *mutatis mutandis,* Proto-Baltic Table 8).

Table 8: The Common Slavic system exemplified by the verb *da-* 'give'

	Default	marked unbounded
Present	*dad-* 'give, are giving'	*da-j-* 'are giving, give (repeatedly), etc.'
Past	*da-* 'gave, were giving'	*da-ja-* 'were giving, gave (repeatedly), etc.'

This system is very close to what we find in Baltic, which is generally more conservative than Slavic. The creation of the imperfect as a dedicated category has probably to do with the fact that marked unbounded verbs were used in the present more rarely than in the past tense. The reason for this assumption is that there is generally a strong pragmatic bias of present tense for unbounded

actions, while there is no such bias in the past tense. Thus, the present tense was the place where the distinction between the default simplex verb (e.g. *dad-*) and the marked-unbounded verb (e.g. *da-j-*) was less relevant or unnecessary. Likewise, in Romance languages, inflectional aspectual distinctions were coded only in the past tense but not in the present. Consequently, unbounded actions did not need to be marked as such in the present but must be highlighted in the past. Therefore, we assume that for many verbs the present tense of the marked unbounded equivalent simply was not, or was only rarely used while the simplex covered all the necessary contexts. This asymmetry between past and present uses of marked unbounded verbs formed by the suffix **-j-* (**-je/o-* (present tense) vs. **-jā-* (past tense)) was the reason for the conventionalization of the past tense use into a dedicated category of imperfect and not vice versa, as is sometimes assumed (e.g., in Borodič 1953).

Furthermore, the claim that the Slavic imperfect is historically a derivational pattern based on the suffix **-j-* to derive marked unbounded verbs implies one important aspect for its emergence. As has been briefly illustrated above (§3.1), for the PIE lexical perfectives the aorist was simply the bare root. In turn, their present tense stem had to be additionally marked by some suffix in order to make it imperfective. Crucially, one of these markings was precisely the suffix discussed here, namely, **-je/o-*;[29] compare the paradigm of the perfective **gʷem-* 'come, arrive' in Table 9 (Seržant 2008: 315). Moreover, analogically to the Slavic imperfect, this PIE suffix derived the imperfective from the perfective stem, as illustrated in Table 9.

Table 9: The late PIE pattern for the derivation of the imperfective counterpart from the perfective verb **gʷem-* 'come, arrive'

Perfective	Imperfective
*(h₁e)-gʷem-t	**(h₁e)-gʷem-je-t*
'(PST)-**arrive**-3SG.ACT'	'(PST)-**arrive**-IPFV-3SG.ACT'
attested in	attested in
Vedic Sanskrit aorist *á-gan* ('PST-arrive.3SG'),	Greek *baínō* < **gʷm-jo-* 'I am going',
á-gm-an ('PST-arrive-3PL') and	Latin *uen-iō* < **gʷm-jo-* 'I am going'
Homeric Greek aorist *bá-* (<**gʷm-*)	

To sum up, Slavic inherited from (late) PIE not only the suffix itself but also

[29] This suffix is spelt conventionally as **-⊠é/ó-* in IE linguistics, but we skip this notation for reasons of simplicity of comparison.

Björn Wiemer & Ilja A. Seržant

its morphotactic distribution: it attaches to perfective stems (aorists) to derive imperfectives (Seržant 2008).

3.2.5 Continued functional development

We have discussed the emergence of the imperfect based on morphological evidence. When it comes to the functional load of the imperfect, we observe the following development:

Stage 1: The imperfect develops from marked unbounded verbs that became gradually restricted to the past tense. They were derived with the suffix clusters *-ā-jā-/*-ē-jā-* from the respective simplexes and represented initially distinct lexemes (compare Old Russian *da-ti* 'give' vs. *da-ja-ti* 'give (unbounded)').

Stage 2: The past tense forms of the marked unbounded verbs such as *da-ja-xu* 'they gave' became associated with the simplex (*da-ti* 'give') in terms of an inflectional category (imperfect). This category acquired a more general meaning of imposing an imperfective operator on the meaning of the underlying verb.

Residuals of stage 1 are found in the earliest Old Church Slavonic documents. Here, the verbs that would later be reinterpreted as imperfective via secondary suffixation were predominantly or almost exclusively (depending on the text) used with the imperfect and not with the aorist to code past reference (cf. *inter alia*, Borodič 1953; Maslov 2004[1954]: 141; Køln 1957; Ivanov 1964: 386; Seržant 2009; Eckhoff & Haug 2015).

In later Slavic, at stage 2, the imperfect:aorist distinction – guided by labor division between the old past tense (aorist) and the new marked atelic$_2$ past tense (imperfect) – gradually developed into viewpoint aspect. The latter became orthogonal to actionality distinctions. Thus, Maslov (2004[1954]), drawing on earlier claims by Potebnja and some others, showed that there was a trend towards a division of labor between the imperfect:aorist and the actionality type. This trend appeared incipiently also in texts of a genuinely Early East Slavic character (see below). The imperfect continued to be the default past tense for atelic$_2$ predicates, but with telic$_2$ predicates its function was to superimpose its unbounded meaning on the lexically conditioned telic$_2$ meaning. This yielded what Maslov (2004[1954]: 149) referred to as *multiply-perfective* meaning (Russ. *kratno-perfektivnoe značenie*): the completion of every act is coded by the telic$_2$ stem, while the unboundedness of the chain of these acts is indicated by the imperfect. Consider the well-known example with the imperfect of perfective verbs meaning 'die' and 'carry out':

(10) Early East Slavic ((Codex Laurentius) Kiparsky 1967: 221)

*ašte   kto   **umrj-aše**,   **tvorj-axu**   tryzn-o(y)*
when INDEF.NOM die[PFV]-IMPF.3SG do[PFV]-IMPF.3PL tryzna-ACC.SG

nadъ   nimь
above him.INS

'Each time someone **died** they **carried out** the tryzna [a rite] on him.'

Therefore, the development of the multiply-perfective meaning is old, but probably it was acquired already independently by different Slavic branches. It is well-developed in modern Bulgarian (Breu 1994: 37-39), but, apart from early East Slavic, it is also incipiently attested, for instance, in Old Czech (Maslov 2004[1954]: 172, 175). Additionally, it could also have a conative (11) or a modal (irrealis) reading (12), although the latter is most probably recent (Maslov 2004[1954]); both examples are cited from Maslov (2004[1954]: 142):

(11) Old Church Slavonic

Da-ěxǫ   *emu   pi-ti   ocъtьn-o*
give[PFV]-IMPF.3PL him.DAT drink[IPFV]-INF vinegarish-N.SG

vin-o.   Onъ že ne prijętъ.
wine[N]-ACC.SG

'They **offered** [more lit.: tried giving] him to drink wine with vinegar. But he did not accept/take it.' (Mark 15, 23)

(12) East Slavic

*Ne lěpo li ny   **bj-ašetъ**,   brat-ie,...*
NEG proper Q us.DAT be-IMPF.3SG brother-VOC.PL

'Wouldn't it **be** proper for us, oh brothers, if ...' (Igor's Tale, end of 12th century)

In general, the imperfect is lost earlier than the aorist,[30] but it still existed in early East Slavic and was not a borrowing from Old Church Slavonic, as a number of peculiarities not found in the latter show (Maslov 2004[1954]: 172). While it is well attested in the 11th c. (Silina 1982: 68f.), later, for example, in the Russkaja Pravda (1282), only aorist forms but no imperfect forms are attested

---

[30] This diachronic constant of Slavic is reversed only under conditions of intense contact. Thus, Molisean and Resian Slavic preserve the imperfect, but have lost their aorist in favour of the previous perfect (> past), as have their Italian contact varieties (Breu 2005: 41-43; 2006: 71f.). However, apart from being conditioned by contact, this development belongs to considerably later periods.

Björn Wiemer & Ilja A. Seržant

(Ivanov 1964: 388). Admittedly, the imperfect is attested in genuine East Slavic texts of high registers (such as chronicles) until the 16th c. (Kiparsky 1967: 220).

3.2.6 Summary: the emergence of Slavic aspect

Before we turn to typological background and other IE languages, let us summarize the results assembled so far. First, the aspectual system of PIE marks predominantly imperfectives and leaves the perfectives morphologically unmarked (type D according to the classification used in §4 below). This remains so into the Common Slavic period, at which point this old system disappears and the development towards coding perfectives (type A) begins. Late Common Slavic is already of type A. Second, the emergence of Slavic aspect is partly rooted in some few – considerably remodelled – morphological schemata of PIE and in a new morphological strategy, namely, prefixation that is exclusively associated to the perfective aspect (thence type A). Third, in contrast to its ancestor language, Slavic vehemently strives after concatenation in its aspectual coding inventory, abandoning various fusional co-effects in morphonology by mere deletion or by reanalysis. Finally, by its origin the imperfect is a derivational category restricted to the past tense by conventionalization. In later periods, the imperfect and, consequently, the aorist tend to interact compositionally with the aspectual properties of their input, developing functions that are orthogonal to the $telic_2$:$atelic_2$ distinction of the verb stems.

On this background, the question becomes especially intriguing as to why the newly developed imperfect (together with the aorist) vanished later in the predominant part of the Slavic-speaking territory, whereas the perfective:imperfective opposition gained ground and developed into a very stable system.

4 Classificatory aspect elsewhere in the world

Let us now map the different stages in the development of the Slavic perfective/imperfective opposition onto a typological backdrop, before we come to a comparison with other IE languages in Europe and areal considerations in §5. According to Dahl Velupillai, perfective/imperfective distinctions "seem to be less skewed in their geographical distribution than, for instance, past tenses". While this holds true for the coarse global distribution of the grammatical opposition as such, there seems to be a rather scattered worldwide distribution of how relations between perfective and imperfective values are marked: "There are languages in which the perfective has no marker and the imperfective has

an overt marker, and vice versa, but most often (at least in our sample) no clear marking relations can be identified. (One reason for this is that the distinction is frequently manifested by stem alternations and similar processes.)" (ibid.). The addition in brackets comprises not only morphonological adaptations between stem and suffix, but also combinations with some other morphological devices such as root ablaut in Classical Greek, cf. *leip-ō* 'remain.IPFV-PRS.1SG' vs. *é-lip-on* 'PST-remain.PFV-PST.1SG'.[31] The morphological system of aspect in contemporary Slavic differs from the type found, e.g., in Classical Greek by its higher degree of concatenation: it tends to abandon inherited morpho(no)logical co-effects such as root ablaut and stem alternations and to create derivational transparency. Morphonological fusional co-effects do exist in Slavic as well, but they usually concentrate around present vs. infinitive stem distinctions and not aspect. Compare, for instance, Russ. *pokaza-l* 'show.NPRS-PST.SG.M' vs. *pokaž-u* (< *pokaz-j-u*) 'show.PRS-1SG', which both belong to the same perfective stem, and contrast this with example (6) in §2.4.

We observe a tendency towards concatenation in the history of postclassical Greek, too. However, in contrast to Slavic, Greek considerably abandoned the PIE distinction between tense and aspect, which are coded cumulatively in the finite verbs already in the classical period. From the Koiné period on, we notice a tendency to abandon aspect distinction in the non-finite domain as well. Thus, the concatenative nature of the perfective/imperfective distinction as well as the non-cumulative coding of tense and aspect in Slavic is special cross-linguistically and differs from the archaic IE languages.

Now, although the derivational character of the Slavic perfective/imperfective distinction has in principle been acknowledged by some typologists, classificatory aspect has so far not really been recognized in the typological literature on TAM grams and systems.[32] It therefore does not come as a surprise that a really systematic study on the world-wide distribution of classificatory aspect systems is lacking. The empirical work of Arkadiev & Shluinsky (2015), therefore, is partic-

---

[31] Östen Dahl (e-mail, 9/16/2015): "we had in mind a situation like that in Classical Greek, where the perfective-imperfective distinction is manifested both in endings and in the choice between the present and the aorist stems – and the relationships between these two is highly idiosyncratic, involving ablaut (as in *leip-* vs. *(e)-lip-* 'remain.PRS-' vs. '(PST)-remain[AOR]-'), affixation (as in *kale-* vs. *(e)-kale-s-*, i.e. 'call.PRS-' vs. '(PST)-call-AOR-'), infixation (*la<m>b-an-* vs. *(e)-lab-*, i.e. 'take<PRS>-PRS-' vs. '(PST)-take.AOR-'), suppletion (as in *erkho-* vs. *(e)-elth-*, i.e. 'come.PRS-' vs. '(PST)-come.AOR-') and reduplication (as in *di-do-* vs. *(e)-do-*, i.e. 'REDUPL-give-' vs. '(PST)-give.AOR-')." (Translations and glosses added.)

[32] To our knowledge, Vladimir Plungjan was the first who developed further the idea that Slavic aspect should be conceived of as a classificatory category (cf. Plungjan 2000: 125f.).

Björn Wiemer & Ilja A. Seržant

ularly welcome as an important pilot investigation. For a language to be included into Arkadiev & Shluinsky's convenience sample they required that the perfective member of the opposition present a situation as limited in time. Depending on the more specific type of the perfective/imperfective opposition, the expression of limitation could arise from completion (with $telic_1$ lexemes) or it could simply represent a temporal kind of limitation, i.e. a delimitative or ingressive meaning (with both $telic_1$ and $atelic_1$ lexemes).

Arkadiev & Shluinsky's sample confirms that classificatory aspect systems are by no means a unique property of Slavic; they can be encountered in different parts of the world. Arkadiev & Shluinsky found such systems in the following languages (groups) and areas:

(i) IE: Slavic, Baltic (Latvian, Lithuanian), Yiddish, Istro-Romanian, Ossetic (Indo-Iranian)

(ii) Kartvelian: Georgian[33]

(iii) Uralic: (a) Samoyedic (Enets, Nenets, Nganasan, Selkup), (b) Ugric (Hungarian, Mansi/Vogul), (c) Finnic (Livonian)

(iv) Altaic: Tunguso-Manchu (Even)

(v) Afro-Asiatic: Chadic (Margi)

(vi) Austronesian: Oceanic (Mokilese, Kusaiean)

(vii) Sino-Tibetan: Northern Tibeto-Birman (Qiang, Gyalrong, Tangut)

(viii) Eskimo-Aleutic: West Greenlandic

(ix) Amerindian languages (different genealogical affiliations): (a) Pomo (Kashaya, Eastern Pomo), (b) Araucanian (Mapuche), (c) Quechua (South Conchucos, Imbabura, Huallaga/Huanuco), (d) Aymara.

Note that all IE languages mentioned in (i) are spoken (or developed) in closer vicinity with some varieties of Slavic, predominantly with Russian; for these languages Slavic influence has been assumed as a major factor in the development of aspect (cf. *inter alia* Breu (1992); Tomelleri 2009; 2010; Arkadiev 2014: 384 also with references). As for Ossetic, however, Arkadiev (2014: 399) has raised doubts

---

[33] The monograph Arkadiev [Arkad'ev] (2015) contains a more comprehensive account of Kartvelian as a whole.

that similarities with the Slavic-style system can be explained from language contact, because contact between Iranian and (Balto-)Slavic populations "must have significantly predated the time when the modern grammatical systems and especially their functional make-up started emerging. Rather, the Balto-Slavic and Ossetic systems of prefixal perfectivization are independent developments based on a common genetic inheritance." Moreover, groups (ii-iv) include non-IE languages spoken in northern Eurasia. These should be taken into account when considering macro-areal patterns in the affixation of verb stems and their possible relation to developing systems of classificatory aspect, although only part of them can be considered as possible substrata that might have strengthened suffixation patterns of Slavic aspect (see §6).[34]

Of course, this synchronic assembly hardly says anything about chronology, nor about the reasons why areal clusters evolve. Moreover, it does not tell much about peculiarities of Slavic aspect even in the context of the broader area surrounding Slavic-speaking territory. Additional parameters applied by Arkadiev & Shluinsky to subclassify the aspect systems of the languages mentioned in (i-ix) help recognize that "Slavic-style aspect" nonetheless deserves particular attention, both from a systematic and an areal perspective. Here, we take up two of their parameters.

First, one should examine the direction of derivation, determined on the basis of the predominating pattern[35] for underived stems (simplexes) along Arkadiev & Shluinsky's classification. On the basis of their approach, four theoretically possible patterns can be established: simplexes may be either perfective or imperfective, and each of them may either be prefixed or suffixed to derive an equivalent of the opposite aspect. These patterns can be derived from Figure 1, which we draw after Arkadiev & Shluinsky's investigation:

Slavic illustrates type (A): underived stems are predominantly imperfective and derive a perfective counterpart via prefixation. Here belong the Baltic languages, Georgian, Hungarian, Yiddish, Ossetic and Sino-Tibetan languages, too. By contrast, pattern (B), which includes simplexes that are predominantly imper-

---

[34] A closer investigation of non-IE languages of northern Eurasia may reveal perfective/imperfective distinctions which have been unnoticed so far in typological research dealing with the more global distribution of grammatical distinctions. For instance, Nenets is claimed to lack grammatical marking of perfective/imperfective aspect in WALS (Chapter 65), whereas it figures in the sample of Arkadiev & Shluinsky (2015). Thus, while the assertion "Northern Europe outside the Slavic area has very little perfectivity/imperfectivity marking" (WALS, ibid.) may more or less remain tenable, the picture might change if a broader range of languages from entire northern Eurasia is investigated with more scrutiny.

[35] Dominance is here understood in terms of type-frequency.

Björn Wiemer & Ilja A. Seržant

(A)   via suffixes   (C)

perfectivizing ⟨   ⟩ imperfectivizing

(B)   via prefixes   (D)
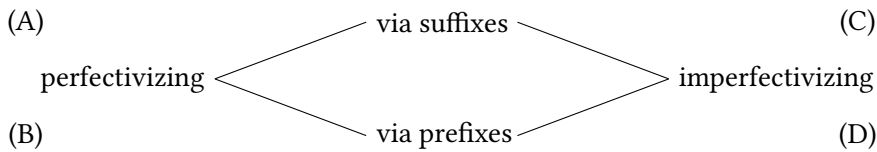
Figure 1: Patterns of derivation from simplex stems

fective, but which derive perfective counterparts via suffixation – is encountered
in Margi (Chadic) and the Micronesian languages. More interesting is pattern (D)
– simplexes are predominantly perfective and derive imperfective equivalents
via suffixes – since it occurs in Samoyedic languages and in Even, which are
spoken in northern Eurasia. Moreover, this pattern corresponds to the prevail-
ing strategy of Proto-Indo-European to derive "imperfectives" from "perfective"
simplexes by means of various morphological schemata most of which involve
suffixation. Pattern (C) – the same as for (D), but with prefixes – has so far re-
mained unattested.

Second, following Arkadiev & Shluinsky (2015), we may ask whether the lan-
guage shows secondary imperfectivization or perfectivization, i.e. whether it
allows already prefixed or suffixed stems to be additionally suffixed or prefixed
in order to cause a change to the opposite aspect.[36] On this basis we can further
distinguish whether secondary (im)perfectivization is achieved via a pair of pre-
fixes or suffixes, or whether the secondary affix attaches to the stem from the
other side of the already attached prefix or suffix, respectively. Thus, this param-
eter classifies according to a combination of direction of function (perfective →
imperfective or imperfective → perfective) and the position of the affixes to each
other (one after another or on opposite sides of the initial stem). The predominant
Slavic pattern of secondary imperfectivization is suffixation of already prefixed
stems. Another example of this pattern is Lithuanian (but not Latvian; see below).

---

[36] We disregard the existence of double prefixation (or 'preverb stacking') that does not change
the aspect (e.g. Russ. *po-ras-stavit*'[PFV] 'put each other on their places' (distributive) ⇐ *ras-
stavit*'[PFV] 'put on their places'). We also ignore prefixes added „on top" of already secondarily
suffixed stems (e.g. Russ. *po-ot-kry-va-t*'[PFV] 'open one after another' ⇐ *ot-kry-va-t*'[IPFV] ⇐
ot-kry-t'[PFV] 'open'). All these are cases of so-called *external prefixes* among which quanti-
fying (accumulative and distributive) functions prevail. Semantically they are of a different
type, and for the system they have a different status, than "simple" prefixation and secondary
suffixation. We also neglect isolated cases and perfectivation of simplexes via suffixes. The lat-
ter is semantically restricted to semelfactives from atelic₁ simplexes denoting repetitive (often
cyclical, mostly motoric) action (e.g. Pol. *mach-a-ć* 'wave' ⇒ *mach-nǫ-ć* 'wave once', *kiw-a-ć*
'nod' ⇒ *kiw-nǫ-ć* 'nod once'), though productive in these confines. See the comments in §2.

9 Diachrony and typology of Slavic aspect: What does morphology tell us?

Arkadiev & Shluinsky do not adduce any other language with this pattern. Other languages considered by them show secondary imperfectivization via suffixes added to other suffixes (used as perfectivizers), e.g. Kashaya and Mansi. Chaining of suffixes is encountered for secondary perfectivization among Samoyedic languages (like Nenets), too. In turn, chaining of prefixes (with change of aspect) is attested in Mingrelian.[37]

In general, however, the number of languages with any kind of secondary perfectivization or imperfectivization appears to be rather limited in contrast to the investigated sample. In particular, Arkadiev & Shluinsky (2015) argue that Latvian, Yiddish, Hungarian, Livonian, Georgian, Margi, Mapuche, Aymara and the Austronesian group do not attest such patterns. One gets the impression that many languages with a classificatory aspect system do not have a possibility to derive another stem (belonging to the opposite aspect) from an already derived stem. However, again, the reasons (and chronology) may differ: either such a possibility was never acquired (as probably in Yiddish or Latvian), or it might have been lost.

If both aforementioned parameters are considered jointly, we see that Slavic stands out against almost all areally contiguous languages and even against a larger northern Eurasian backdrop. Apart from Lithuanian, only Istro-Romanian is known as a non-Slavic language in which contacts with speakers of Slavic have led to the appearance of, and increase in, secondary suffixation (cf. Arkadiev forthcoming). In other words: Slavic (plus Lithuanian and Istro-Romanian) appear to be the only languages on a broader areal background which show productive patterns of prefixation and (secondary) suffixation used for the purpose of perfectivization and imperfectivization, respectively. Leaving aside now Istro-Romanian, the consistency with which this happens in Slavic and Lithuanian differs; Lithuanian in many respects shows a less grammaticalized stage than the Slavic languages which surround it. But the morphological patterns are fully parallel and some of them – such as *-āje/o- (present)/*-ājā- (past) and *-ēje/o- (present)/*-ējā- (past), discussed at length in §3.2.4 – are most probably inherited from a common Baltic and Slavic dialect continuum. Lithuanian has kept this "heritage" and revitalized it at a later stage with the new suffix -inė-, while Latvian has not developed any new productive aspectual suffixation which would go beyond strong lexical restrictions (see §5.4).[38] For example, the common suffix

---

[37] One has to admit that the imperfectivizing prefix comes between the stem and the perfectivizing prefix (which comes first also in the "derivational history"). In the following Mingrelian example the perfectivizing prefix is in **bold**, the imperfectivizing one is underlined: **ge**-<u>tmi</u>-a-ʒic-en-d-u; this has to be translated as 'was laughing at him/her' (Arkadiev 2014: 391).

[38] It did, however, create a somewhat productive suffix -inā- to derive morphological causatives

(*-āje/o-…*) is retained in just a few verbs such as *brauk-t* 'drive-INF' vs. *brauk-ā-t* 'drive-HAB-INF'.

To conclude, crucial for the rise of the Slavic aspect system based on stem derivation was the fact that one productive set of affixes (prefixation) at some point in history started being combined with another productive set of affixes (suffixes). It follows from the areal overview given above that these morphological preconditions are met only rarely in languages of the world. It is our conviction that this constellation is the key to understanding the rise of the Slavic aspect system. Above we have traced back the development of prefixation and suffixation of verb stems and argued that they developed from separate sources and diachronically at different periods of time: imperfectivization schemata represent old – albeit highly remodelled – patterns while the exploitation of prefixes for perfectivization is a much more recent development. In contrast, other IE languages in Europe that have exploited prefixation to code actionality (which is a pre-stage to aspect) have lost the old imperfectivization strategies altogether. This topic will be addressed in the next section.

5 Verb stem derivation in ancient Indo-European languages of Europe

In §3, we supplied a diachronic account of verbal prefixation and suffixation in Slavic. In turn, the preceding discussion in §4 served the purpose of recognizing the typological peculiarities of Slavic aspect and of relativizing claims concerning its alleged rarity. In this section, we want to critically assess some facts and findings that help cast light on the role verbal affixation might have played in shaping the aspectual character of verb stems in other IE languages outside of Slavic. Our survey is selective: we do not pretend to give an exhaustive account of preverbation and prefixation (or of suffixation) in these languages, but we focus on languages (or language groups) with some closer areal affinity to at least some Slavic-speaking territory during the first millenium AD.

Many scholars have mentioned the widespread existence of preverbs (often also included into inventories of particles) attached to verbal stems in different old IE languages of Europe. The morphological status of these preverbs varies, as they are sometimes characterized as proclitics, on other occasions as already tightly agglutinated parts of verbal stems, i.e. as prefixes. Possibly, this variation reflects different stages on a clitic—affix cline of morphologization. Admittedly,

---

and other deverbatives.

there is no straightforward correlation between this assumed cline and the development of a preverb into a prefix. One of the problems is that neither preverbs nor prefixes need be unstressed, so that we cannot be sure that it is cliticization as such which triggers the processes.[39] In the first place, however, tightness of coalescence with lexical stems does not *per se* give reliable information concerning the function of morphemes on a preverb > prefix-cline and their role in forming systematic oppositions pertaining to actionality and/or grammatical aspect. Note that investigations into preverbation in ancient IE languages have concentrated largely on issues of morphologization (cliticization > agglutination) of preverbs originating from adverbs or so-called *verb particles* and on the question of what processes of coalescence tell us about constituent and argument structure in early IE.[40] Preverbs as mere aspectual bounders are mentioned rather occasionally, so that it is hardly possible to draw any conclusions as to whether the bounder function should be characterized as modification or as telicization$_2$, cf. for instance, Cuzzolin, Putzu & Ramat (2006: 10). Among others, this applies to Ancient Greek, too, and for this reason we will not deal specifically with Greek anymore in this article.

5.1 Romance

This general picture obviously holds true also for Latin. In Classical Latin, many prefixes still functioned as markers of telicity$_1$, but this function deteriorated by Late Latin (after 300 AD; cf. Haverling 2003), thus more or less at the time of the Great Migrations. Therefore, we feel justified to say that, by and large, neither Romance nor its ancestor Latin pushed the use of preverbs further than the modificational stage (see the upper part of Table 4) and maybe some incipient stage (ii).

In the Romance successor languages of Latin, prefixes usually became lexicalized and opaque when they could no longer be separated from the stem; compare It. *in/segnare* 'teach', Fr. *s'en/dormir* 'fall asleep' (Haverling 2003: 125; Cuzzolin, Putzu & Ramat 2006: 12), or the prefixes used did not carry any aspectual function, being restricted, as a rule, to spatial and related functions (e.g. It. *ag-giungere—dis-giungere* 'add, attach—separate') or to comitative meaning or redo-

---

[39] We want to thank Christian Lehmann for pointing this out to us.

[40] Cf. for instance, Vincent (1999) on Latin and Romance, Boley (2004); Cuzzolin, Putzu & Ramat (2006) more generally on preverbation in diverse stages of early IE languages. Among others, Luraghi (2003) and Viti (2008) investigated the role of spatial prefixes, their relation to cognate prepositions and adverbs as well as their role in the syntax of Ancient Greek. As for Latin cf. Leumann, Hofmann & Szantyr (1977[1926^5]: 557-566) and Haverling (2003).

Björn Wiemer & Ilja A. Seržant

ing (compare Romance *re-*, *con-* and their translational equivalents). Obviously, in older stages of Romance, e.g. Old French, preverbs were used widely, but according to the examples adduced in relevant publications (e.g., Dufresne, Dupuis & Tremblay 2003) the function of these preverbs was restricted to modifications of the verbal action more or less like in modern German or Dutch.

Suffixes, in turn, proved unproductive, or they were incorporated into inflectional paradigms. The latter happened to Latin *-sc-*, which occurs in some forms of the present tense conjugations of Romance successor languages (e.g., It. *capi-sc-e* 's/he understands' from *cap-ire*.INF 'understand'). Cf. Allen (1995) on this process whereby a former derivational morpheme turns into a merely formal marker incorporated into inflectional paradigms. In Greenberg (1991) this process was called *regrammaticalization*. As we saw in the preceding sections, this is clearly not what happens when we distinguish perfective and imperfective stems. Only the development of the Slavic imperfect shows a change from derivation into inflection (see §3.2.4).

In sum, neither (late) Latin nor its Romance successor languages relied on productive prefixation strategies to code telicity$_1$. The same applies, *mutatis mutandis*, to suffixation strategies to mark actionality functions associated to imperfectivity. We are unaware of any reliable findings concerning possible contact relations of Vulgar Latin or its successor varieties in early Romance with Slavic. We thus refrain here from any comments on this issue.

5.2 Celtic

It is not entirely clear whether there were considerable contacts between Celtic and Slavic populations, in particular during the Great Migrations (cf. the critical remarks in Polomé 1972: 64–69 and Andersen 2003: 48). Although toponyms of Celtic origin have been attested as far east as in the Danubian delta and the upper Dniester basin (Blažek 2015), these traits of Celtic influence could have been due to settlements from the last centuries BC, when Celtic tribes had spread over vast territories of Europe and into Asia Minor. In fact, "Celtic speech, apart from possible enclaves, appears to have died out on the European continent by AD 500" (MacAulay 1992: 2), and the earliest form of Celtic that could be reconstructed more or less completely from extant sources, Old Irish, reflects a stage just after this time (approx. 6th-9th century AD; Thurneysen 1975[1946]: 1–11). Moreover, Old Irish was spoken in the northwestern periphery of an earlier Celtic dialect continuum, while contacts with Slavs could have occurred only on its opposite end, and we do not know to which extent other Celtic dialects were comparable to Old Irish in terms of preverbation (or suffixation). Gvozdanović (2009; 2015)

9 Diachrony and typology of Slavic aspect: What does morphology tell us?

wonders whether certain important typological changes in word prosody such as syllable structure and the direction of palatalization from regressive to progressive assimilation of the velars /*k, g, x*/ could not have been due to some Celtic influence. She links her argument to the Venetian region to which Celtic is supposed to have once spread. However, apart from Gvozdanović's observations on phonology (mainly word prosody) there are no really "hard core" arguments able to substantiate Celtic influence on Slavic. After all, "we do not have sufficient evidence to identify the individual contacting language, which may well have been the eastern European Venetic (…) of which we have no direct linguistic evidence" (Gvozdanović 2015: 97). The relation to Celtic, thus, remains unclear.

Therefore, the following brief remarks on preverbation in Old Irish have to be taken with caution, at least insofar as we cannot say whether Old Irish did not differ, with respect to verbal stem derivation, from Celtic varieties which previously had been spoken on the European continent, some of them possibly in some proximity to speakers of Common Slavic.

Old Irish had some dozen preverbs (prefixes), most of them obviously in a transitional stage between clitics and affixes; the most widespread and prominent was *ro-*. Gvozdanović (2015: 104), summarizing Thurneysen (1975[1946]: 339-348), concludes that Old Irish *ro-* "perfectivizes the verb on the level of grammatical aspect, not only lexical aspect". She even goes further saying that the functional properties of this preverb, "as part of the verb phrase, are fully paralleled by the perfective aspect in Slavic". These parallels concern the combination with the imperfect, which yields repetition in the past (compare modern Bulgarian, see §3.2.5), and, first of all, prefixation of present tense stems which occurred only in gnomic or other inactual functions of the present (including dispositional modality, e.g. *as·ro-b(a)ir* 'can [= is able to] say' vs. *as·beir* 'says'). However, the term *perfective* probably entered the English translation of Thurneysen's authoritative grammar (written originally in German) as an inadequate rendition of Germ. *perfektisch* or *Perfekt* (Lambert 1995: 251, following McCone 1987),[41] where it seems to mean accomplished action (Germ. *vollendete Handlung*), i.e. telic$_2$ predicates. Moreover, *ro-* (leaving aside other preverbs) was optional, verbs with an inherently telic meaning (= telic$_1$) could convey perfective values without *ro-* as well (cf. Lewis & Pedersen 1937: 141f., 245-248; Lambert 1995: 231-239). For this reason, Schumacher (2004: 81) proposed to consider *ro-* (and other preverbs) just as an augment of the stem that does not constitute any new category (differently

---

[41] Cf. Thurneysen (1909: 319). According to West (1981/1982: 252), the facts allow for an interpretation as mere anteriority marker as well, so that stems prefixed with *ro-* should probably be considered 'perfect forms'.

Lambert 1995: 251f.).

These observations, as fragmentary as they are, seem to be indicative rather of a stage in which preverbs (prefixes) frequently, but, optionally, were used to mark inherent boundaries, i.e. to create telic$_2$-predicates under favorable conditions and independently from tense. This corresponds to stage (ii) in Table 4. This resembles the situation we encounter in Gothic, to which we now turn.

5.3 Gothic

Germanic has been regarded as being much closer to Slavic and Baltic than any other of the IE groups in Europe. It is very probable that the speakers of Gothic, as the best-documented old Germanic language, were in rather close contact with Baltic and Slavic tribes, before they fell victim to the Great Migration (in which they intensely participated), so that by the 6th c. AD they disappeared from history in eastern Europe (Kotin 2012: 13-15), while the Visigoths on the Iberian Peninsula eventually abandoned their language at the beginning of the 7th c. AD.

The Gothic verbal prefix *ga-* was the most salient representative of a series of prefixes, and its behaviour was very similar to Old Irish *ro-*.[42] The known documents (primarily Wulfila's Bible) reflect the state of the language from approx. the 4th c. AD (i.e. slightly earlier than Old Irish). These doculects were, of course, influenced especially by Greek, and also by Latin (Kotin 2012: 21). In particular verb stems prefixed with *ga-* have, since Streitberg (1891), been evoking divergent claims about their status as "perfectivizers". As with the Old Irish preverbs, most researchers (except Maslov 1959a) have remained rather vague as for what they understand by *aspect*, in particular which role is played by prefixes, and whether the designation *perfective* characterizes a lexical or a grammatical feature. In Gothic, *ga-* and some more prefixes[43] functioned not only as lexical modifiers, but they often fulfilled functions that are reminiscent of mere bounders of the action denoted by the simplex stem (see below). Thus, Kotin (2012: 287) writes that Gothic demonstrated "a relatively stable opposition of simplexes and so-called *ga-*composites (…), that can largely be interpreted as aspectual" (our translation). However, *aspectual* here does not have the value of a grammatical opposition, but, rather, of a complex of actional and voice-related features. Remarkably, Kotin also mentions that in quite a few cases, prefixes did not so

---

[42] The same applies to Old High German (*gi-*) and other prefixes in the most ancient stages of documented Germanic languages (on which cf. Kotin 2012: 297, 397).

[43] Kotin (2012: 393f.) names *dis-*, *fra-*, *faír-*, *ga-*, his examples also show *us-* (2012: 397), cf. also Guxman (1998: 205-209). Braune & Ebbinghaus (1961: 124) name some more, and they mention tmesis and other signs of a looser juncture between prefix and stem (ibid: 124f.).

much modify the lexical semantics of the simplex, but rather made it more pronounced; conversely, some simplex stems "selected" a prefix depending on their own inherent semantics (2012: 394f.).[44] This observation brings to mind the Vey-Schooneveld effect of prefixes discussed in §3.2.2.

However, even if we regard *ga-* as a perfectivizer proper, its application remained restricted both in terms of the range of verb stems with which it could be combined (type-frequency) and the reliability with which it was encountered in cases when it should be expected from the meaning in discourse (token-frequency). The application of *ga-* (or any other prefix) was by no means very consistent. Moreover, the extant texts do not allow for too far-reaching conclusions about which pairs of simplex/prefixed stems were distributed over different aspectual functions, in particular as concerns finite forms (cf. also West 1981/1982: 250f. and the review of the literature until the mid-1950s in Maslov 1959a). It is symptomatic that even Kotin's thorough examination of Gothic texts brought to light such pairs only for inherently telic$_1$ verbs (cf. *ga-swiltan* vs. *swiltan* 'die – be dying [Germ. *im Sterben liegen*]', *fullnan* 'fill [INTR]' vs. *ga-fullnan* 'fill [INTR], become full (to its limits)') and for verbs of passive perception (e.g. *saíhvan* 'see' vs. *ga-saíhvan* 'catch sight of [Germ. *erblicken*]'). With these verbs, *ga-* served to mark off the initial boundary of the perceptual state (= atelic$_1$), whereas with telic$_1$ verbs, namely those denoting more punctual changes *ga-* modified the lexical meaning (e.g. *niman* 'take' vs. *ga-niman* 'take with o.s., take along', *qiman* 'come, arrive' vs. *ga-qiman* 'gather, assemble [INTR]'); cf. Kotin (2012: 294-300, 395-397)).

In sum: it is not entirely clear whether Gothic *ga-* should be analyzed as a marker of telicity$_1$ or telicity$_2$, not least due to some terminological confusion in the literature. Examples such as Kotin's *bindan* ⇒ *ga-bindan* 'bind, tie (up)', *swiltan* ⇒ *ga-wiltan* 'die' (see above) suggest that there was, at least, a considerable progress from telicity$_1$ towards telicity$_2$ in Gothic (our stage (ii) in Table 4 above). By contrast, Maslov's (1959a) analysis leads rather to a characterisation of *ga-* as a marker of telicity$_1$ (stage (i) in Table 4).

---

[44] With the exception of *ga-*, Kotin (2012: 394) ascribed a prototypical semantic function to each particular prefix: „in connection with various verb stems this function could either have remained practically unaltered, or it was modified to different degrees, depending on the modifications allowed or even required by the semantics of the verb stem. This property of the derivational basis exerts an impact not only on modifications of the basic semantic function of the prefix, but it also restricts the selection of the latter." (our translation) For instance, *taíran* 'tear' ⇒ *dis-taíran* (≈ *ga-taíran*) 'ditto', *qistjan* 'destroy' ⇒ *fra-qistjan* 'ditto': *dis-* was lexically associated with separation, *fra-* with destruction and loss (2012: 394f.).

Björn Wiemer & Ilja A. Seržant

5.4 Baltic

The morphological prerequisites necessary for the innovations common to all later Slavic languages are present in Baltic as well. This allows to infer that the premises of these innovations must have developed in a larger dialectal region of early IE, of which Slavic and Baltic formed part. Concomitantly, the old layer of Baltic suffixation is etymologically related to the respective old layer of Slavic (see §3.2.1 and §3.2.4). However, old suffixes have ceased to be productive. While varieties of Lithuanian created new productive suffixes such as *-inė-* (iterative, durative, etc.) or *-dav-* (habitual past), Latvian did not introduce new verbal suffixes with aspectual functions.

Latvian shows a certain opposition of atelic$_2$/telic$_2$ predicates comprising non-punctual telic$_1$-verbs. This opposition is lexically restricted and builds on prefixed stems contrasted with the respective simplex stems that take verb particles part of which are cognates of the prefixes; for instance, **ie**-*nāca istabā* 'in-come.PST.3 room.LOC.SG' 'entered into the room' (usually telic$_2$) vs. *nāca* **iekšā** *istabā* 'come.PST.3 inside.PRT room.LOC.SG' 'was entering the room' (telic$_1$); cf. Holvoet (2001: 132-141), Arkadiev [Arkad'ev] (2015: 132-134). Verb particles are a relatively recent phenomenon, which most probably arose from contact with Germanic (Low and High German, Swedish) and Finnic (Wälchli 2001).[45]

Lithuanian is different, since in its Aukštaitian dialects and the standard language it has introduced two new suffixes relevant for differentiation in actionality: semelfactive *-ėre-/-ėle-* and *-(d)inė-*; the latter takes on functions associated to unboundedness.[46] Remarkably, especially the latter suffix has been attested as particularly frequent (on type and token level) in southeastern Lithuania, i.e. in close vicinity with (East) Slavic. The suffix *-(d)inė-* has been extraordinarily frequent in (now extinct or moribund) insular dialects in Belarus. It is thus apparent that this new suffix gained frequency from contact with Slavic speakers (and Lithuanian-East Slavic bilingualism), but only in recent times (Wiemer 2009: 359-363; Arkadiev [Arkad'ev] 2015: 125-131). The same may hold true for double prefixation, which is otherwise unusual in Lithuanian and Baltic in general, but quite widespread in East Slavic. It would be risky to try to extrapolate into a remote past these facts about the distribution and frequency of these younger

---

[45] There is much of mutual influence between Latvian and Finnic (Estonian, Livonian) contact in here, and Estonian, in turn, is probable to have introduced this technique under contact with Germanic (Hasselblatt 1990; Metslang 2001). Anyway, this recent innovation is in stark contrast to the otherwise strong suffixing strategy of Finno-Ugric.

[46] Both suffixes certainly arose from some morphological reanalysis (as did most of the Slavic suffixes mentioned in §3.2.3).

verbal affixes that are relevant for aspectual distinctions.

6 Conclusion and an outlook

In this paper it has been our main concern to give a comprehensive assessment of the internal preconditions which made the rise of the contemporary aspect system based on stem derivation (perfective/imperfective verbs) of Slavic languages possible. We have restricted ourselves to the core of the system and inquired into morphological changes that affected the formation of verb stems in the prehistory and early history of Slavic. The analysis concerned both particular morphemes and patterns of affixation in relation to each other and to initial verb stems; we tried to trace back these patterns and morphemes from PIE into Slavic, pointed out genuinely Slavic innovations based on an IE heritage and discussed the further expansion or loss of early inner-Slavic developments. We have focused on changes that affected Slavic as a whole and stopped short at the point when, after the consolidation of the core system, inner-Slavic differences both in formal expression and in the range and hierarchy of functions became more pronounced.

Favorable inherited conditions are visible in the internal changes of Slavic since times prior to documentation (see §3). In asking whether or not Slavic aspect continues aspectually relevant oppositions found in PIE (cf. *inter alia*, van Wijk 1929; Stang 1942) we have to be careful not to mix up morphological schemata with functions of grammatical aspect or aspectual functions in general. Once this is taken into account, we can claim that the morphological devices used in Common Slavic to mark unboundedness *grosso modo* represent – albeit highly restructured and modified – heritage from late PIE. An important feature of Slavic morphemes to mark unboundedness is that there is a strong tendency towards morphological concatenation, away from non-concatenative PIE schemata. This trend can be reconstructed for Common Slavic and it continues to this day.

Preverbation (particles, prefixes and intermediate stages) developed at the time of ancient IE languages. Especially the comparison with Old Irish (see §5.2) demonstrates that preverbs used as bounders of verbal action, evolved in very different regions of Europe by the middle of the first millenium AD. Whether this testifies to spontaneous independent parallelism triggered by some propensity on the basis of inherited adverbs or should rather be explained by mutual contacts between subgroups in Europe (e.g. between early Slavic and Gothic), cannot be ascertained. However, preverbation has been prominent especially for

changes of valency or argument structure whereas, apparently, apart from Slavic, in none of these IE languages did prefixes (or other preverbs) start to productively function as mere bounders of verbal action, without additional functions on the syntax-semantics interface. These patterns were then, in Slavic, strengthened by the combination with suffixes, which decayed in other IE languages of Europe.

Thus, for late Common Slavic we can also assume a tendency of extending the distinction between telic and atelic stems from a purely lexical opposition (i.e. from $telic_1$ vs. $atelic_1$) into an opposition in which realized telicity (= $telic_2$ meanings) gets marked via prefixation. At a later stage, prefixes start serving also the differentiation of other aspectual meanings such as ingressivity or mere temporal delimitation (see Table 4 in §3.2.2). Simplexes, unmarked also with respect to $telicity_2$ (as there were both $telic_1$ or $atelic_1$ simplexes to begin with), underwent different, lexeme-specific developments still into recent centuries to stabilize aspect assignment of stems. Unprefixed, but suffixed stems – representing the oldest layer – played a subsidiary role in the emergence of the opposition. From the point of view of morphological patterns, the last step was taken when suffixes started being productively attached to already prefixed stems (so-called *secondary imperfectivization*). This pattern has remained less productive in the western half of Slavic, while it is very prominent in the eastern half (East Slavic, Bulgarian, Polish).

This being settled, a further aim of this paper consisted in demonstrating that, although Slavic is by no means unique in having developed a classificatory aspect system, it nevertheless stands out on a larger areal, namely Eurasian, background and in comparison to other IE groups in Europe. We, thus, compared diachronic and synchronic data of Slavic with somewhat fragmentary data against an areal and typological backdrop (§4) as well as with likewise fragmentary data from earlier stages of IE languages (§5).

Now, on the basis of this comparison, there arises a more intricate question, which, for the moment, we only want to state. Namely: one wonders to what extent the rise and consolidation of the Slavic aspect system can be explained as only a spontaneous evolution that just continued already existing preconditions of stem derivation. To what extent might contact with non-IE-speaking populations have helped trigger, or support, the consolidation of such continued development, which we do not find in areally close IE languages? To put it differently: there is no doubt that the morphological prerequisites necessary for the evolution of stem-derivational aspect in Slavic continued earlier patterns that were partly rooted in PIE. But why has only Slavic developed these prerequisites in such a consistent manner during the last, say, two millenia, whereas in other

IE groups suffixation and/or preverbation have gone other ways? In the latter, such prerequisites disappeared or were renewed (for instance, by separable verb particles), but nowhere else have prefixes and suffixes come together to jointly build a grammatical system as in the Slavic languages (except for much more recent developments as in some varieties of Lithuanian, obviously under East Slavic influence; see §5.4).

Slavic expanded over a large territory all over eastern Europe since about 600 AD; contacts with groups of speakers of Uralic or Turkic were, thus, very likely. In general, the existence of Finno-Ugric and Turkic adstrata and even substrata in the eastern part of Slavic can hardly be doubted.[47] However, whether contact with Finno-Ugric or Turkic-speaking populations might have been sufficient to strengthen suffixing strategies, must be inquired in a well-considered manner, taking into account various kinds of (often indirect) evidence and equilibrating findings from different approaches. Among other things, it should be asked what morphological techniques of stem extension in possible contact languages looked like, which types were productive and resembled, in some way or other, stem derivation in early Slavic. For instance, according to Serebrennikov, many Finno-Ugric languages show suffixal extensions of verb stems with various functions from the domains of iterativity (repetitive or habitual action) or of semelfactivity. These aspectual meanings can be interpreted as remnants of an earlier stage, when dialectal differentiation was less advanced (Serebrennikov 1960: 31-34, 188).

As concerns Turkic, we may assume that its oldest reconstructable layer "operated entirely by adding suffixes at the end of the word and had a fully developed system of suffixes" (Clauson 1962: 27). Throughout, Turkic languages have experienced several renewals of suffixation, among others of suffixes modifying the aspectual character of the verb. Such suffixes developed via morphologization (enclisis > agglutination) from converb or auxiliary constructions (Johanson 1998a: 41-43, 1998b: 113-115; cf. also Erdal 2004: 262-272). In general, Turkic languages can be regarded as having remained astonishingly homogeneous in this respect (Menges 1968: 181; Johanson 1998b: 111). One wonders whether such findings cannot be more substantiated with respect to aspectual functions of suffixes (or postverbs, which preceded them in morphologization) at the dawn of written documentation of Turkic, i.e. from the early eight century AD, and in subsequent centuries. One should seek contacts of Uralic and Turkic speaking populations

---

[47] Cf. Veenker (1967) and Haarmann (2014) on Finno-Ugric, Stachowski (2014) on Turkic. Consider also the history of the Bulgars from the middle of the first millenium AD, a Turkic (Oghur) tribal union which was later ethnically and linguistically assimilated by eastern South Slavic people and henceforth gave its name to this mixed population and the later state.

especially for the Common Slavic period (400-900 AD), when the Slavic dialect continuum must have been still sufficiently homogeneous and compact for innovations to spread across Slavic from East to West, including the strengthening of already developed patterns.

Now, if we want to explain the early stages in the rise of the Slavic perfective/imperfective opposition, we should take into account the following considerations: (i) In contrast to many other IE languages, Slavic has partly preserved stem extensional patterns (suffixing strategies, though less concatenative), (ii) frequent patterns of preverbation (prefixation) in later IE languages in Europe outside Slavic did not further participate in the formation of viewpoint aspect despite some incipient developments, and (iii) the predominant suffixing strategies of non-IE languages with which speakers of prehistoric and later Slavic must have come into considerable contact, in particular since the Great Migrations. Considering all these pieces of a puzzle, one is tempted to formulate the following hypothesis:

(13)   While, during the first millenium AD, prefixation of verb stems was
       shared with other IE language groups as a new development in Europe,
       suffixation patterns were sustained by similar patterns in Finno-Ugric
       and Turkic speaking populations.

In some sense, Common Slavic came to be sandwiched between an area with predominant preverbalizing strategies in the IE speaking West and an area with a clear preference for suffixing strategies in the East where speakers of Finno-Ugric and Turkic dominated. The morphological prerequisites for a system of viewpoint aspect based on the combination of prefixes and suffixes in verbal stem derivation had developed by Common Slavic times, but *only* in Slavic *both* suffixes and prefixes eventually turned out as being capable of marking aspectual distinctions without voice or valency-related changes. Support for this assumption comes from the observation that secondary suffixation has been much more productive in the eastern half of Slavic than in the western one (see §3.2.3).

This hypothesis and the issues related to it wait for an investigation, if one wishes to complement an internal reconstruction with contact-induced considerations.

Acknowledgments

9 Diachrony and typology of Slavic aspect: What does morphology tell us?

ing Indo-European and other issues. We also owe special gratitude to Peter Arkadiev and Christian Lehmann, whose comments on an earlier version of this article made us aware not only of some unfortunate formulations but, first of all, of pitfalls and fallacies which we had overlooked. Furthermore, we have to thank (alphabetically) Henning Andersen, Walter Breu, Pino Marco Pizzo and Florian Sommer, who drew our attention to other possible pitfalls. Beyond them, we want to thank Silvia Luraghi and Chiara Zanchi for discussion of Ancient Greek (which remained in the background), Östen Dahl and Ljuba Veselinova for discussion of typological matters, Bohumil Vykypěl for some advice concerning sources of prehistoric Slavic morphology, Jadranka Gvozdanović and Ranko Matasović concerning Celtic and Lázsló Károly concerning Turkic. Needless to say, none of these colleagues should be made responsible for any possible drawbacks of this article, as the sole responsibility for any remaining misinterpretations or mistakes is ours.

Abbreviations

ACC	accusative	PFX	prefix
ACT	active	PL	plural
DAT	dative	PRS	present
F	feminine	PST	past
HAB	habitual	Q	question particle
IMPF	imperfect	REDUPL	reduplication
INDEF	indefinite	RFL	reflexive
INF	infinitive	SFX	suffix
INTR	intransitive	SG	singular
N	neuter	THV	thematic vowel
NEG	negation	UNBOUND.PST	marked unbounded past
NOM	nominative	VIR	virile
NPRS	non-present stem	VOC	vocative

Acknowledgements

References

Allen, Andrew A. 1995. Regrammaticalization and degrammaticalization of the inchoative suffix. In Henning Andersen (ed.), *Historical linguistics 1993 (selected*

papers from the 11th International Conference on Historical Linguistics, Los Angeles, 16-20 August 1993), 1–8. Amsterdam, Philadelphia: Benjamins.

Ambrazas, Vytautas. 2006. *Lietuvių kalbos istorinė sintaksė*. Vilnius: Lietuvių kalbos institutas.

Andersen, Henning. 1985. Protoslavic and Common Slavic – questions of periodization and terminology. *International Journal of Slavic Linguistics and Poetics* 31-32. 67–82.

Andersen, Henning. 2003. Slavic and the Indo-European migrations. In Henning Andersen (ed.), *Language contacts in prehistory: Studies in stratigraphy*, 45–76. Amsterdam, Philadelphia: Benjamins.

Andersen, Henning. 2009. On the origin of the Slavic aspects: Questions of chronology. In Vik Bubenik, John Hewson & Sarah Rose (eds.), *Grammatical change in Indo-European languages (Papers presented at the workshop on Indo-European linguistics at the XVIIIth International Conference on Historical Linguistics, Montreal, 2007)*, 123–140. Amsterdam, Philadelphia: Benjamins.

Andersen, Henning. 2013. On the origin of the Slavic aspects: Aorist and imperfect. *Special issue of Journal of Slavic Linguistics* 21(1). Laura Janda (ed.). 17–44.

Arkadiev, Peter. 2014. Towards an areal typology of prefixal perfectivation. *Scando-Slavica* 60(2). 384–405.

Arkadiev, Peter. Forthcoming. Borrowed preverbs and the limits of contact-induced change in aspectual systems. In Rosanna Benacchio, Alessio Muro & Svetlana Slavkova (eds.), *Grammatikalizacija i leksikalizacija: Rol' prefiksov v aspektual'nyx processax*. Firenze: Firenze University Press.

Arkadiev, Petr M. & Andrej B. Shluinsky. 2015. Slovoklassificirujuščie aspektual'nye sistemy: Opyt tipologii. *Vestnik SPbGU* 2015(3). 4–24.

Arkadiev [Arkad'ev], Petr M. 2015. *Areal'naja tipologija prefiksal'nogo perfektiva (na materiale jazykov Evropy i Kavkaza)*. Moskva: Jazyki slavjanskoj kul'tury.

Arumaa, Peeter. 1985. *Urslavische Grammatik: Einführung in das vergleichende Studium der slavischen Sprachen. Band 3: Formenlehre*. Heidelberg: Winter.

Bech, Gunnar. 1971. *Beiträge zur genetischen idg. Verbalmorphologie*. Vol. 44 (Historisk-filosofiske meddelelser 5). København: Munksgaard.

Beekes, Robert Stephen Paul. 1969. *The development of Proto-Indo-European laryngeals in Greek*. The Hague: Mouton.

Beekes, Robert Stephen Paul. 1995. *Comparative Indo-European linguistics*. Amsterdam, Philadelphia: Benjamins.

Bermel, Neil. 1997. *Context and the lexicon in the development of Russian aspect*. Berkeley: University of California Press.

Birnbaum, Henrik. 1979. *Common Slavic: Progress and problems in its reconstruction.* Columbus: Slavica.

Blažek, Václav. 2015. The eastern border of Celtic settlement: A toponymic perspective. In Dafydd Johnston, Elena Parina & Maxim Fomin (eds.), *'yn llawen iauwn, yn llawn iaith': Proceedings of the 6th international colloquium of societas Celto-Slavica*, 23–29. Aberystwyth: University of Wales Centre for Advanced Welsh & Celtic Studies.

Bohnemeyer, Jürgen & Mary Swift. 2004. Event realization and default aspect. *Linguistics and Philosophy* 27(3). 263–296.

Boley, Jacqueline. 2004. *Tmesis and Proto-Indo-European syntax.* Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.

Borodič, Vera V. 1953. K voprosu o formirovanii soveršennogo i nesoveršennogo vida v slavjanskix jazykax. *Voprosy jazykoznanija* 1953(2). 68–86.

Boye, Kasper & Peter Harder. 2012. A usage-based theory of grammatical status and grammaticalization. *Language* 88. 1–44.

Braune, Wilhelm & Ernst A. Ebbinghaus. 1961. *Gotische Grammatik (mit Lesestücken und Wörterverzeichnis). 16. Auflage neu bearbeitet von Ernst A. Ebbinghaus.* Tübingen: Niemeyer.

Breu, Walter. 1980. *Semantische Untersuchungen zum Verbalaspekt im Russischen.* München: Sagner.

Breu, Walter. 1992. Zur Rolle der Präfigierung bei der Entstehung von Aspektsystemen. In Marguerite Guiraud-Weber & Charles Zaremba (eds.), *Linguistique et slavistique: Melanges offerts à Paul Garde*, 119–135. Aix-en-Provence: Université de Provence / Paris: Institut d'études slaves.

Breu, Walter. 1994. Interactions between lexical, temporal and aspectual meanings. *Studies in Language* 18. 23–44.

Breu, Walter. 2000a. Zur Position des Slavischen in einer Typologie des Verbalaspekts (Form, Funktion, Ebenenhierarchie und lexikalische Interaktion). In Walter Breu (ed.), *Probleme der Interaktion von Lexik und Aspekt (ILA)*, 21–54. Tübingen: Niemeyer.

Breu, Walter. 2000b. Der Verbalaspekt in der obersorbischen Umgangssprache im Rahmen des ILA-Modells. In Walter Breu (ed.), *Slavistische Linguistik 1999*, 37–76. München: Sagner.

Breu, Walter. 2005. Verbalaspekt und Sprachkontakt: Ein Vergleich der Systeme zweier slavischer Minderheitensprachen (SWR/MSL). In Sebastian Kempgen (ed.), *Slavistische Linguistik 2003*, 37–95. München: Sagner.

Breu, Walter. 2006. Flektivnyj i derivacionnyj glagol'nyj vid v molizsko-slavjanskom jazyke. *Voprosy jazykoznanija* 2006(3). 70–87.

Björn Wiemer & Ilja A. Seržant

Brugmann, Karl. 1895. Die mit dem Suffix *-to-* gebildeten Partizipia im Verbalsystem des Lateinischen und des Umbrisch-Oskischen: Eine syntaktische Untersuchung. *Indogermanische Forschungen* 5. 89–152.

Chantraine, Pierre. 1953. *Grammaire homérique: II. Syntaxe.* Paris: Klincksieck.

Clauson, Gerard. 1962. *Studies in Turkic and Mongolic linguistics.* London, New York: Routledge.

Cuzzolin, Pierluigi, Ignazio Putzu & Paolo Ramat. 2006. The Indo-European adverb in diachronic and typological perspective. *Indogermanische Forschungen* 111. 1–38.

Dahl, Östen. 1981. On the definition of the telic-atelic (bounded-unbounded) distinction. In Philip J. Tedeschi & Annie Zaenen (eds.), *Syntax and semantics, vol. 14: Tense and aspect*, 79–90. New York: Academic Press.

DeLazero, Octav Eugen. 2012. *Aspect in syntax.* Cornell University PhD thesis.

Dickey, Stephen M. 2000. *Parameters of Slavic aspect: A cognitive approach.* Stanford: CSLI Publications.

Dickey, Stephen M. 2011. The varying role of *po-* in the grammaticalization of Slavic aspectual systems: Sequences of events, delimitatives, and German language contact. *Journal of Slavic Linguistics* 19(2). 175–230.

Dowty, D. R. 1991. Thematic Proto-Roles and argument selection. *Language* 67. 547–619.

Dufresne, Monique, Fernande Dupuis & Mireille Tremblay. 2003. Preverbs and particles in Old French. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology 2003,* 33–60. Dordrecht: Kluwer.

Eckhoff, Hanne & Dag Haug. 2015. Aspect and prefixation in Old Church Slavonic. *Diachronica* 32(2). 186–230.

Erdal, Marcel. 2004. *A grammar of Old Turkic.* Leiden, Boston: Brill.

Forsyth, James. 1972. The nature and development of the aspectual opposition in the Russian verb. *The Slavonic and East European Review* 50(121). 493–506.

Gagarina, Natalia. 2004. Does the acquisition of aspect have anything to do with aspectual pairs? *ZAS Papers in Linguistics* 33. 39–61.

Greenberg, Joseph H. 1991. The last stages of grammatical elements: Contractive and expansive desemanticization. In Elizabeth C. Traugott & Bernd Heine (eds.), *Approaches to grammaticalization, vol. 1: Focus on theoretical and methodological issues*, 301–314. Amsterdam, Philadelphia: Benjamins.

Guxman, M. M. 1998. *Gotskij jazyk.* Moskva: Izd-vo MGU.

Gvozdanović, Jadranka. 2009. *Celtic and Slavic and the Great Migrations: Reconstructing linguistic prehistory.* Heidelberg: Winter.

Gvozdanović, Jadranka. 2015. Evaluating prehistoric and early historic linguistic contact. In Dag T. T. Haug (ed.), *Historical linguistics 2013 (Selected papers from the 21st International Conference on Historical Linguistics, Oslo, 5-9 August 2013)*, 89–108. Amsterdam, Philadelphia: Benjamins.

Haarmann, Harald. 2014. Finnougrisch-slavische Sprachkontakte. In Karl Gutschmidt, Sebastian Kempgen, Tilman Berger & Peter Kosta (eds.), *Die slavischen sprachen: Ein internationales handbuch zu ihrer struktur, ihrer geschichte und ihrer Erforschung, vol. 2*, 1181–1198. Berlin, Boston: De Gruyter Mouton.

Haig, Geoffrey L. J. 2008. *Alignment change in Iranian languages: A construction grammar approach.* Berlin, New York: De Gruyter Mouton.

Haspelmath, Martin & Andrea D. Sims. 2010. *Understanding morphology. 2nd edition.* London: Hodder Education.

Hasselblatt, Cornelius. 1990. *Das estnische Partikelverb als Lehnübersetzung aus dem Deutschen.* Wiesbaden: Harrassowitz.

Haverling, Gerd. 2003. On prefixes and actionality in Classical and Late Latin. *Acta Linguistica Hungarica* 50(1-2). 113–135.

Havránek, Bohuslav. 1937. *Genera verbi v slovanských jazycích II.* Praha.

Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization.* Cambridge: Cambridge University Press.

Holvoet, Axel. 2001. *Studies in the Latvian verb.* Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.

Holzer, Georg. 2014. Vorhistorische Periode. In Karl Gutschmidt, Sebastian Kempgen, Tilman Berger & Peter Kosta (eds.), *Die slavischen Sprachen: Ein internationales Handbuch zu ihrer Struktur, ihrer Geschichte und ihrer Erforschung*, vol. 2, 1117–1131. Berlin, Boston: De Gruyter Mouton.

Ivanov, Vjačeslav V. 1964. *Istoričeskaja grammatika russkogo jazyka.* Moskva: Prosveščenie.

Janda, Laura & Olga Lyashevskaya. 2011. Grammatical profiles and the interaction of the lexicon with aspect, tense, and mood in Russian. *Cognitive Linguistics* 22(4). 719–763.

Janda, Laura, Anna Endresen, Julia Kuznetsova, Olga Lyashevskaya, Anastasia Makarova, Tore Nesset & Svetlana Sokolova. 2013. *Why Russian aspectual prefixes aren't empty: Prefixes as verb classifiers.* Bloomington: Slavica Publ.

Jasanoff, Jay H. 2003. *Hittite and the Indo-European verb.* Oxford: Oxford University Press.

Johanson, Lars. 1998a. The structure of Turkic. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic languages*, 30–66. London, New York: Routledge.

Björn Wiemer & Ilja A. Seržant

Johanson, Lars. 1998b. The history of Turkic. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic languages*, 81–125. London, New York: Routledge.

Johanson, Lars. 2000. Viewpoint operators in European languages. In Östen Dahl (ed.), *Tense and aspect in the languages of Europe*, 27–187. Berlin, New York: De Gruyter Mouton.

Kiparsky, Valentin. 1967. *Russische historische Grammatik, Bd. II: Die Entwicklung des Formensystems*. Heidelberg: Winter.

Kölligan, Daniel. 2004. Wenn zwei dasselbe tun: Iterativa und Kausativa. In Maria Kozianka, Rosemarie Lühr & Susanne Zeilfelder (eds.), *Indogermanistik – Germanistik – Linguistik: Akten der Arbeitstagung der Indogermanischen Gesellschaft, Jena 18.-20.09.2002*, 193–247. Hamburg: Kovač.

Kølln, Hermann. 1957. Vidové problémy v staroslověnštině. *Philologica* 3(1). 67–100.

Kortlandt, Frederik H. H. 1986. The origin of the Slavic imperfect. In R. Olesch & H. Rothe (eds.), *Festschrift für Herbert Bräuer zum 65. Geburtstag*, 253–258. Köln/Wien: Böhlau.

Kotin, Michael L. 2012. *Gotisch. Im (diachronen und typologischen) Vergleich*. Heidelberg: Winter.

Kuryłowicz, Jerzy. 1937. La structure de l'imparfait slave. Acta Jutlandica 9(9(1)). 85–392.

Kuryłowicz, Jerzy. 1960. *Esquisses linguistiques*. Wroclaw-Kraków: Polska Akademia Nauk.

Kuznecov, Petr S. 1953b. *Istoričeskaja grammatika russkogo jazyka*. Moskva: Nauka.

Lambert, Pierre-Yves. 1995. Les préverbes perfectifs du vieil-irlandais. In André Rousseau (ed.), *Les préverbes dans les langues d'Europe: Introduction à l'étude de la préverbation,* 227–254. Villeneuve d'Ascq (Nord): Septentrion.

Łaziński, Marek & Björn Wiemer. 1996. Terminatywność jako kategoria stopniowalna. *Prace Filologiczne* XL. 99–126.

Lehmann, Christian. 1995[1982]. *Thoughts on grammaticalization*. Munich: LINCOM EUROPA.

Lehmann, Volkmar. 1990. Zur ontogenetischen und diachronen Entwicklung grammatischer Kategorien des Polnischen. In Walter Breu (ed.), *Slavistische linguistik 1989*, 167–188. München: Sagner.

Lehmann, Volkmar. 1992. Grammatische Zeitkonzepte und ihre Erklärung. *Kognitionswissenschaft* 2. 156–170.

Lehmann, Volkmar. 1993. Die russischen Aspekte als gestufte Kategorien: Ein Beispiel für die Bedeutung der kognitiven Linguistik in der slavistischen Sprachwissenschaft. *Die Welt der Slaven* 38(2). 265–297.

Lehmann, Volkmar. 1999. Sprachliche Entwicklung als Expansion und Reduktion. In Tanja Anstatt (ed.), *Entwicklungen in slavischen Sprachen*, 169–254. München: Sagner.

Lehmann, Volkmar. 2004. Grammatication via extending derivation. In Walter Bisang, Nikolaus P. Himmelmann & Björn Wiemer (eds.), *What makes grammaticalization? (A look from its fringes and its components)*, 169–186. Berlin, New York: De Gruyter Mouton.

Lehmann, Volkmar. 2009. Aspekt und Tempus im Slavischen. In Sebastian Kempgen, Peter Kosta, Tilman Berger & Karl Gutschmidt (eds.), *Die slavischen Sprachen: Ein internationales Handbuch zu ihrer Struktur, ihrer Geschichte und ihrer Erforschung*, vol. 1, 526–556. Berlin, New York: De Gruyter Mouton.

Leskien, August. 1919. *Grammatik der altbulgarischen (altkirchenslavischen) Sprache.* Heidelberg: Carl Winter.

Leumann, Manu, Johann Baptist Hofmann & Anton Szantyr. 1977[1926[5]]. *Lateinische Grammatik: Lateinische Laut- und Formenlehre.* Vol. 1. München: C.H. Beck'sche Verlagsbuchhandlung.

Lewis, Henry & Holger Pedersen. 1937. *A concise comparative Celtic grammar.* Göttingen: Vandenhoeck & Ruprecht.

LIV[2]. 2001. *Lexikon der Indogermanischen Verben. Die Wurzeln und ihre Primärstammbildungen.* Unter Leitung von Rix, Helmut und der Mitarbeit vieler anderer bearb. von Kümmel, M. and Zehnder, Th. and Lipp, R. and Schirmer, B. Zweite, erweiterte und verbesserte Auflage, bearb. von Kümmel, Martin and Rix, Helmut. Wiesbaden: Reichert Verlag.

Lühr, Rosemarie. 1999. Das slavische Imperfekt. Chronologie einer Periphrase mit dem Instrumental. In Jürgen Habisreitinger, Robert Plath & Sabine Ziegler (eds.), *Gering und doch von herzen, 25 indogermanistische beiträge: Bernhard forssman zum 65. Geburtstag*, 167–182. Wiesbaden: Reichert.

Luraghi, Silvia. 2003. *On the meaning of prepositions and cases: The expression of semantic roles in ancient Greek.* Amsterdam, Philadelphia: Benjamins.

MacAulay, Donald. 1992. *The Celtic languages.* Cambridge: Cambridge University Press.

Maslov, Jurij S. 1948. Vid i leksičeskoe značenie glagola v sovremennom russkom literaturnom jazyke. *Izvestija AN SSSR, Otdelenie literatury i jazyka, t. 7, vyp* 4. 303–316.

Maslov, Jurij S. 1959a. Kategorija predel'nosti/nepredel'nosti glagol'nogo dejstvija v gotskom jazyke. *Voprosy jazykoznanija* 1959(5). 69–80.

Maslov, Jurij S. 2004[1954]. Perfektivnyj imperfekt v drevnerusskom literaturnom jazyke. In Aleksandr V. Bondarko, Timur A. Majsak & Vladimir A. Plungjan (eds.), *Ju. S. Maslov. Izbrannye trudy. Aspektologija i obščee jazykoznanie*, 141–175. Reprint of Maslov, Jurij S. 1954. Imperfekt glagolov soveršennogo vida v slavjanskix jazykax. *Voprosy slavjanskogo jazykoznanija* 1. 68-138. Moskva: Jazyki slavjanskoj kul'tury.

Maslov, Jurij S. 2004[1959]. Vozniknovenie kategorii soveršennogo/nesoveršennogo vida. In Aleksandr V. Bondarko, Timur A. Majsak & Vladimir A. Plungjan (eds.), *Ju. S. Maslov. Izbrannye trudy. Aspektologija i obščee jazykoznanie*, 131–140. Russian translation of the paper: Maslov, Jurij S. 1959. Zur Entstehungsgeschichte des slavischen Verbalaspektes. *Zeitschrift für Slawistik* 4. 560-568. Moskva: Jazyki slavjanskoj kul'tury.

McCone, Kim. 1987. *The early Irish verb*. Maynooth: An Sagaart.

Meillet, Antoine. *Le slave commun. 2. éd. rev. et augm. Nouveau tirage avec le concours de a. Vaillant*. Paris: Champion.

Meiser, Gerhard. 1993. Zur Funktion des Nasalpräsens im Urindogermanischen. In Gerhard Meiser (ed.), *Indogermanica et Italica, Festschrift für Helmut Rix zum 65. Geburtstag, unter Mitarbeit von Jadwiga Bendahman, Jón Axel Harðarson und Christiane Schaefer*, 280–330. Innsbruck: Institut für Sprachwissenschaft.

Mende, Julia. 1999. Derivation und Reinterpretation: Die Grammatikalisierung des russischen Aspekts. In Tanja Anstatt (ed.), *Entwicklungen in slavischen Sprachen*, 285–325. München: Sagner.

Menges, Karl H. 1968. *The Turkic languages and peoples: An introduction to Turkic studies*. Wiesbaden: Harrassowitz.

Metslang, Helle. 2001. On the development of the estonian aspect: The verbal particle *ära*. In Östen Dahl & Maria Koptjevskaja-Tamm (eds.), *Circum-Baltic languages, vol. 2: Grammar and typology*, 443–479. Amsterdam, Philadelphia: Benjamins.

Miklosich, Franz. 1926. *Vergleichende Grammatik der slavischen Sprachen, II. Bd.: Stammbildungslehre*. Reprint of first edition 1875. Heidelberg: Winter.

Nesset, Tore. 2013. The history of the Russian semelfactive: The development of a radial category. *Special issue of Journal of Slavic Linguistics* 21(1). Laura Janda (ed.). 123–170.

Ostrowski, Norbert. 2006. *Studia z historii czasownika litewskiego. Iteratiwa. Denominatiwa*. Poznań: Wydawnictwo Naukowe UAM.

Plungjan, Vladimir A. 2000. *Obščaja morfologija. Vvedenie v problematiku.* Moskva: URSS.

Plungjan, Vladimir A. 2002. O specifike vyraženija imennyx prostranstvennyx xarakteristik v glagole: Kategorija glagol'noj orientacii. In Vladmir A. Plungjan (ed.), *Issledovanija po teorii grammatiki, vyp. 2: Grammatikalizacija prostranstvennyx značenij*, 57–98. Moskva: Russkie slovari.

Plungjan, Vladimir A. 2011. *Vvedenie v grammatičeskuju semantiku: Grammatičeskie značenija i grammatičeskie sistemy jazykov mira.* Moskva: RGGU.

Pohl, Heinz Dieter. 1971. Das slavische Imperfekt auf -*ěax*- und -*aax*-. *Zeitschrift für slavische Philologie* 38(2). 349–360.

Polomé, Edgar C. 1972. Germanic and the other Indo-European languages. In Frans van Coetsem & Herbert L. Kufner (eds.), *Toward a grammar of Proto-Germanic*, 43–69. Tübingen: Niemeyer.

Rasmussen, Jens Elmegård. 1988. The IE background of the Slavic nasal-infix presents. *Arbejdspapirer [af] Institut for Lingvistik Københavns Universitet* 7. 195–201.

Rosinas, Albertas. 1995. *Baltų kalbų įvardžiai: Morfologijos raida.* Vilnius.

Sasse, Hans-Jürgen. 2002. Recent activity in the theory of aspect: Accomplishments, achievements, or just non-progressive state? *Linguistic Typology* 6-2. 199–271.

Schlyter, Suzanne. 1990. The acquisition of tense and aspect. In Jürgen M. Meisel (ed.), *Two first languages: Early grammatical development in bilingual children*, 87–121. Dordrecht: Foris.

Schooneveld, Cornelis H. 1958. The so called 'préverbes vides' and neutralization. In *Dutch contributions to the Fourth International Congress of Slavists. Moscow, September 1958*, 159–161. The Hague: Mouton.

Schumacher, Stefan. 2004. *Die keltischen präverbien. Ein vergleichendes, etymologisches und morphologisches lexikon* (Innsbrucker Beiträge zur Sprachwissenschaft). Unter Mitarbeit von Britta Schulze-Thulin und Caroline aan de Wiel. Innsbruck.

Serebrennikov, Boris A. 1960. *Kategorii vremeni i vida v finno-ugorskix jazykax permskoj i volžskoj grupp.* Moskva: Izdatel'stvo AN SSSR.

Seržant, Ilja. 2008. Review: N. Ostrowski, Norbert Ostrowski. Studia z historii czasownika litewskiego. Iteratiwa. Denominatiwa. 2006. *Historische Sprachforschung* 121. 308–320.

Seržant, Ilja. 2009. Tempus und Aspekt im ältesten Russisch-Kirchenslavischen, untersucht an den Texten des Gottesdienstmenäums für Dezember. In Dagmar Christians, Dieter Stern & Vittorio S. Tomelleri (eds.), *Bibel, Liturgie und Fröm-*

migkeit in der Slavia Byzantina: Festgabe für Hans Rothe zum 80. Geburtstag, 309–328. München, Berlin: Sagner.

Seržant, Ilja A. 2012. The so-called possessive perfect in North Russian and the Circum-Baltic area. A diachronic and areal approach. *Lingua* 122. 356–385.

Seržant, Ilja A. 2014. *Das Kausativ im Tocharischen*. München: LINCOM.

Ševeleva, Maria N. 2010. Vtoričnye imperfektivy s suffiksom -*yva-*/-*iva-* v letopis- jax XII-XVI vv. *Russkij jazyk v naučnom osveščenii* 20(2). 200–243.

Silina, Vera B. 1982. Istorija kategorii glagol'nogo vida. In Ruben I. Avanesov & Vjačeslav V. Ivanov (eds.), *Istoričeskaja grammatika russkogo jazyka*, 158–279. Moskva: Nauka.

Silina, Vera B. 1995. Glagol. In Vjačeslav V. Ivanov (ed.), *Drevnerusskaja gram- matika XII-XIII vv*, 374–506. Moskva: Nauka.

SłPrasłow, Franciszek. Sławski 1974. *Słownik prasłowiański, t. I: A—B*. Wrocław etc.: Ossolineum.

Smith, Carlotta. 1991. *The parameter of aspect*. Dordrecht: Kluwer.

Sreznevskij, Izmail Ivanovič. 1893-1912. *Materialy dlja slovarja drevne-russkago jazyka po pismennym pamjatnikam*. Sankt-Peterburg: Tipografija impera- torskoj akademii naukъ.

Stachowski, Stanisław. 2014. Türkischer Einfluß auf den slavischen Wortschatz. In Karl Gutschmidt, Sebastian Kempgen, Tilman Berger & Peter Kosta (eds.), *Die slavischen Sprachen: Ein internationales Handbuch zu ihrer Struktur, ihrer Geschichte und ihrer Erforschung*, vol. 2, 1198–1210. Berlin: De Gruyter Mouton.

Stang, Christian S. 1942. *Das slavische und baltische Verbum*. Oslo: Universitets- forlaget.

Stang, Christian S. 1966. *Vergleichende Grammatik der baltischen Sprachen*. Oslo, Bergen, Tromsø: Universitetsforlaget.

Stieber, Zdzisław. 1989. *Zarys gramatyki porównawczej języków słowiańskich*. Warszawa: PWN.

Stoll, Sabine. 1998. The role of aktionsart in the acquisition of Russian aspect. *First Language* 18. 351–377.

Stoll, Sabine. 2005. Beginning and end in the acquisition of the perfective aspect in Russian. *Journal of Child Language* 32. 805–825.

Streitberg, Wilhelm. 1891. Perfektive und imperfektive Aktionsart im Germani- schen. Einleitung und I. Teil: Gotisch. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 15. 70–177.

Thurneysen, Rudolf. 1909. *Handbuch des Alt-Irischen: Grammatik, Texte und Wörterbuch, Bd. I: Grammatik*. Heidelberg: Winter.

Thurneysen, Rudolf. 1975[1946]. *A grammar of Old Irish*. Dublin: The Dublin Institute for Advanced Studies.

Tomelleri, Vittorio. 2009. The category of aspect in Georgian, Ossetic and Russian: Some areal and typological observations. *Faits des langues* 1. 245–272.

Tomelleri, Vittorio. 2010. Slavic-style aspect in the Caucasus. *Suvremena lingvistika* 69. 65–97.

Vaillant, André. 1966. *Grammaire comparée des langues slaves: Le verbe*. Vol. III. Paris: Klincksieck.

van Wijk, Nicolaas. 1929. Sur l'origine des aspects du verbe slave. *Revue des études slaves* 9(3-4). 237–252.

Veenker, Wolfgang. 1967. *Die Frage des finnougrischen Substrats in der russischen Sprache*. Indiana: Bloomington.

Vendler, Zenon. 1957. Verbs and times. *Philosophical Review* 46. 143–160.

Vey, Marc. 1952. Les préverbes "vides" en tchèque moderne. *Revue des études slaves* 29(1-4). 82–107.

Vincent, Nigel. 1999. The evolution of c-structure: Prepositions and PPs from Indo-European to Romance. *Linguistics* 37. 1111–1154.

Viti, Carlotta. 2008. Coding spatial relations in Homeric Greek: Preverbs vs. Prepositions. *Historische Sprachforschung* 121. 114–161.

Vondrák, Wenzel. 1924. *Vergleichende Slavische Grammatik, I. Bd.: Lautlehre und Stammbildungslehre.* (2., stark vermehrte und verbesserte Auflage). Göttingen: Vandenhoeck & Ruprecht.

Wälchli, Bernhard. 2001. Lexical evidence for the parallel development of the Latvian and Livonian verb particles. In Östen Dahl & Maria Koptjevskaja-Tamm (eds.), *Circum-Baltic languages, vol. 2: Grammar and typology*, 413–441. Amsterdam, Philadelphia: Benjamins.

West, Jonathan. 1981/1982. Preverbs in Gothic and Old Irish: A typological parallel. *Studia Celtica* XVI/XVII. 248–258.

Wiemer, Björn. 2008. Zur innerslavischen Variation bei der Aspektwahl und der Gewichtung ihrer Faktoren. In Karl Gutschmidt, Ulrike Jekutsch, Sebastian Kempgen & Ludger Udolph (eds.), *Deutsche Beiträge zum 14. Internationalen Slavistenkongreß, Ohrid 2008*, 383–409. München: Sagner.

Wiemer, Björn. 2009. Zu entlehnten Präfixen und anderen morphosyntaktischen Slavismen in litauischen Insel- und Grenzmundarten. In Lenka Scholze & Björn Wiemer (eds.), *Von zuständen, dynamik und veränderung bei pygmäen und giganten (festschrift für Walter breu zu seinem 60. Geburtstag)*, 347–390. Bochum: Brockmeyer.

Björn Wiemer & Ilja A. Seržant

Wiemer, Björn. 2015. O roli vida v oblasti kratnosti i pragmatičeskix funkcij (ėskiz s točki zrenija xronotopii). In Rosanna Benacchio (ed.), *Glagol'nyj vid: Grammatičeskoe značenie i kontekst / verbal aspect: Grammatical meaning and contex*, 585–609. München: Sagner.

Name index

Name index

Name index

Name index

Language index

Did you like this book?

This book was brought to you for free

Please help us in providing free access to linguistic research worldwide. Visit http://www.langsci-press.org/donate to provide financial support or register as a community proofreader or typesetter at http://www.langsci-press.org/register.

language
science
press

Unity and diversity in grammaticalization scenarios

Empirical research is carried out in a cyclic way: approaching a research area bottom-up, data lead to interpretations and ideally to the abstraction of laws, on the basis of which a theory can be derived. Deductive research is based on a theory, on the basis of which hypotheses can be formulated and tested against the background of empirical data. Looking at the state-of-the-art in translation studies, either theories as well as models are designed or empirical data are collected and interpreted. However, the final step is still lacking: so far, empirical data has not lead to the formulation of theories or models, whereas existing theories and models have not yet been comprehensively tested with empirical methods.

This publication addresses these issues from several perspectives: multi-method product- as well as process-based research may gain insights into translation as well as interpreting phenomena. These phenomena may include cognitive and organizational processes, procedures and strategies, competence and performance, translation properties and universals, etc. Empirical findings about the deeper structures of translation and interpreting will reduce the gap between translation and interpreting practice and model and theory building. Furthermore, the availability of more large-scale empirical testing triggers the development of models and theories concerning translation and interpreting phenomena and behavior based on quantifiable, replicable and transparent data.