

Origins of Grammatical Forms and Evidence from Latin¹

Brigitte L. M. Bauer

The University of Texas at Austin, U.S.A.

blm.bauer@austin.utexas.edu

Analyses of grammaticalization in Indo-European languages tend to focus on Germanic and Romance forms, for which we generally are able to identify the original autonomous composing elements. By contrast, in early Indo-European languages we find many inflected formations or other grammatical elements that are opaque. The question then is: are these forms opaque (1) because of the length of their existence; or (2) because the composing elements in origin were not autonomous lexical forms, but rather of a different type? We may never be able to fully answer that question. Yet comparison of phenomena of grammaticalization in different stages of Latin will help clarify certain aspects of it.

This article compares manifestations of grammaticalization in the transition from Latin to Romance with earlier formation processes in Latin and Indo-European. I will submit that the instances of grammaticalization that we find in the later stages of Latin and that resulted in a number of new grammatical forms in Romance, reflect a major linguistic innovation. While the new grammatical forms are created out of lexical or mildly grammatical autonomous elements, earlier processes seem to primarily involve particles with a certain semantic value, and freezing. This fundamental difference explains why the attempts of early Indo-Europeanists such as Franz Bopp at tracing the lexical origins of Indo-European inflected forms were unsuccessful and strongly criticized by the Neo-Grammarians.

1. Introduction

Over the last thirty or so years the study of grammaticalization has been one of the hotbeds of historical linguistics. The term grammaticalization refers to

¹Acknowledgements. I am grateful to the anonymous referees for their useful suggestions and comments on an earlier version of this paper.

a type of language change whereby a lexical autonomous element gradually acquires grammatical functions and/or eventually becomes a fully grammatical element, cf. for example the Old English noun *lic(e)* ‘body’ turning into an adverbial suffix (-ly, e.g. *lightly* < OE *leohtlice* lit. ‘light.body-Dat.’).

The process also includes phenomena whereby a grammatical element acquires stronger grammatical features, e.g. a demonstrative developing into a definite article, losing its deictic value in the process (e.g. the Latin demonstrative *ille* ‘that’ becoming a definite article in many Romance languages). Because of its inherently innovative nature (as opposed to analogy for example; cf. Meillet 1982 [1912]), grammaticalization may result in the creation of new forms (e.g. the English indefinite article *a* [< numeral ‘one’] or periphrastic prepositions tracing back to a noun [e.g. *in front of* < N *front*]), new paradigms (e.g. the Romance future tenses that originated in an infinitive and finite *habere* ‘have’),² or even new grammatical categories, such as definiteness, which came to be explicitly marked with the development of definite articles in many Indo-European languages (e.g. Bauer 2007).

Grammaticalization is attested in numerous Indo-European and non-Indo-European languages. Handbooks and case studies provide ample data and analyses of these phenomena, such as Heine and Kuteva’s *World Lexicon of Grammaticalization* (2002) registering numerous instances from African, Indo-European, Semitic, Caucasian, Oceanic, Amerindian, and various other language groups. See also e.g. Heine and Reh (1984; African languages), Hopper and Traugott (1993; a variety of languages), and more recent collections of articles, such as the volumes edited by Verhoeven *et al.* (2008) and Van Linden *et al.* (2010).

This paper aims to demonstrate that the phenomena of grammaticalization in the transition from Latin to Romance do not merely result in new forms, but that the

| | |
|---------------------------------------|--|
| ² La. <i>cantare habeo</i> | > Fr. (<i>je</i>) <i>chanterai</i> /Sp. <i>cantaré</i> |
| La. <i>cantare habes</i> | > Fr. (<i>tu</i>) <i>chanteras</i> /Sp. <i>cantarás</i> |
| La. <i>cantare habet</i> | > Fr. (<i>il</i>) <i>chantera</i> /Sp. <i>cantará</i> |
| La. <i>cantare habemus</i> | > Fr. (<i>nous</i>) <i>chanterons</i> /Sp. <i>cantaremos</i> |
| La. <i>cantare habete</i> | > Fr. (<i>vous</i>) <i>chanterez</i> /Sp. <i>cantaréis</i> |
| La. <i>cantare habent</i> | > Fr. (<i>ils</i>) <i>chanteront</i> /Sp. <i>cantarán</i> |

processes themselves were innovative: grammatical forms were created out of lexical or mildly grammatical autonomous elements, e.g. the possessive *habeo* in combination with an infinitive or a perfective participle eventually becoming an auxiliary of tense (e.g. La. *cantare* + *habeo* ‘sing-Inf. have-1sg.’ > Fr. [je] *chanterai* ‘(I) will sing’) or the noun *mente* in combination with an adjective turning into an adverbial suffix (e.g. La. *clara* + *mente* ‘with a clear mind’ > It. *chiaramente* ‘clearly’). Comparative analysis of these formations and those from earlier stages of Latin and Indo-European will reveal that the earlier processes typically involved elements of a different type. This conclusion sheds new light on the contributions of early Indo-Europeanists — a topic on which the handbooks on grammaticalization are remarkably silent. Their historical overviews indeed tend to ignore or misrepresent attempts by (early) Indo-Europeanists at reconstructing the origins of early grammatical forms. Yet these contributions and debates are crucial to the analysis of emerging grammatical forms in Indo-European, both in early times and later as will become clear in this paper. Our formal analysis will therefore start with an evaluation of these attempts.

In this article we will first briefly assess the historical overviews of three often-quoted textbooks on grammaticalization (Section 2). Subsequently we will review contributions of Indo-Europeanists to the study of the origins of grammatical forms and the debates these entailed (Section 3). With this theoretical embedding we will then discuss phenomena of grammaticalization in the early period of Latin (Section 4.1), comparing their patterns to those of the later stages (Section 4.2). Finally we will evaluate our findings in Section 5, Conclusions.

2. Historical overviews of the study of grammaticalization

In this section we will assess the historical overviews provided by three textbooks on grammaticalization: Hopper and Traugott (1993), Heine, Claudi, and Hünemeyer (1991), and C. Lehmann (1995). Their historical sketches discuss early scholars who have referred to processes or aspects of grammaticalization without however using that terminology, such as Etienne Bonnot

de Condillac (1715-1780), according to whom verbal inflection — e.g. tense suffixes — traces back to autonomous words (Heine *et al.* 1991:5). Or John Horne Took (1736-1812), the “father of grammaticalization studies”, who posited that inflectional and derivational elements were “fragments of earlier independent words agglutinated to the root word” (Heine *et al.* 1991:5).

Reference is also made to less known scholars who assumed processes of grammaticalization — again without using that term — such as Riis (1854) and Christaller (1875), German missionaries in Ghana who specialized in Twi (Akan) and presented “new frameworks for discussing developments” whereby lexical elements became grammatical (Heine *et al.* 1991:8). Wegener (1885) on the other hand argued that discourse elements could develop into “morphosyntactic constructions” (Heine *et al.* 1991:8).

Moreover, the three textbooks discuss the work by Wilhelm von Humboldt (1767-1835) and Georg von der Gabelentz (1840-1893).³ Humboldt’s (1822 [1963]; 1836) contribution to the study of grammaticalization concerns his assumption that his different types of language based on grammatical marking (isolating, agglutinating, inflectional; Humboldt 1963 [1822]:42-43) parallel the stages of language development: Stage I > Stage II-Isolation > Stage III-Agglutination > Stage IV-Inflection (Humboldt 1963:54-55). Moreover, mention is made of the correlation put forward by Humboldt between the stages of higher complexity and superior levels of intellectual and cultural achievement (cf. Humboldt 1963:56-63). Interestingly, Humboldt’s further specification about the origins of grammatical endings is not provided in the handbooks, but is revealing: “alle haben vermutlich, nach Horne Took’s richtigerer Theorie, ihren Ursprung in wirklichen, Gegenstände bezeichnenden Wörtern. Die grammatisch-formale Wirkung der Sprache beruht daher auf dem Grade, in welchem diese Partikeln noch ihrem Ursprunge näher,

³Humboldt: Heine *et al.* (1991:6-9; 22); Traugott and Hopper (1993:18-21; 25); C. Lehmann (1995:3-4).

Gabelentz: Heine *et al.* (1991:8-9; 22; 214; 245); Traugott and Hopper (1993:19-21); C. Lehmann (1995:3-4).

oder entfernter stehen" (Humboldt 1963:51).

While Humboldt took an evolutionary hence linear perspective, Gabelentz emphasized the recurring nature of the creation of grammatical forms, putting forth the metaphor of the spiral with two driving forces, *Bequemlichkeitstrieb* 'drive for ease' and *Deutlichkeitstrieb* 'drive for clarity' (Gabelentz 1901:356; 355-358). In Gabelentz's view endings of the finite verb originated in personal pronouns or finite verbs (e.g. La. "*videbo* = *vide-fuo*"; Gabelentz 1901:256): "was heute Affixe sind, das waren einst selbstständige Wörter" (Gabelentz 1901:256).

It is important to note that Humboldt and Gabelentz, like Condillac before them, hypothesized about the origins of forms in general terms, but in fact typically did not trace nor analyze the assumed developments on the basis of data, even if there is occasional reference to well-known later examples such as the development of Latin *habeo* into an auxiliary in compound and future tense forms (e.g. Gabelentz 1901:348). The contributions of early Indo-Europeanists were a marked exception to this trend.

It is remarkable, however, that the historical overviews tend to ignore research on the origins of grammatical forms by Indo-Europeanists in the 19th century. Or if there is reference, it is rather selective. Yet, the origins of grammatical forms was a prominent topic on which many have worked in the early days. In fact only one Indo-Europeanist is mentioned often, Franz Bopp (Heine *et al.* 1991:6; C. Lehmann 1995:2-3) and his work is then saluted as an important contribution to grammaticalization studies: "Bopp presented numerous examples of the development from lexical material to auxiliaries, affixes, and, finally, inflections. Grammaticalization, as conceived of by Bopp, forms an important parameter in understanding diachronic Indo-European linguistics" (Heine *et al.* 1991:6). Yet there is no allusion to the fact that the results of analyses by Bopp and others were rather controversial among close colleagues, as we will see in greater detail below. Despite these disappointing results, the efforts may be significant but in a sense different from what is generally assumed in the literature on grammaticalization: the early — 19th century — attempts at tracing the origins of grammatical forms and the debates these entailed, put the grammatical

forms in perspective that emerged in the later periods of Latin and the role of autonomous lexical elements in them.

In his 1816 analysis and more so in his 1833 handbook Bopp traced the origins of grammatical forms in detail, indeed establishing a link between lexical elements and the later grammatical endings. According to C. Lehmann, Brugmann was “favorably inclined to hypotheses of this kind” (C. Lehmann 1995:3). Yet, the situation was more complex than that and in fact Brugmann did not approve of this type of analysis, as we will see in greater detail below (Section 3).

Another Indo-Europeanist who is mentioned in some studies is the Sanskritist William Whitney (1827-1894), who argued that lexical items could lose their concrete meaning and become an element of “formal grammatical expression” (1875:90-91; Heine *et al.* 1991:7). Finally, reference is made to Wilhelm von Schlegel’s analysis of elements that lose their original meaning to become grammatical elements, such as demonstratives and the numeral ‘one’ (1818:27-28; Heine *et al.* 1991:6-7; C. Lehmann 1995:1). Yet his observations were made within the context of Provençal grammar, rather than Indo-European in general.

As the names provided above indicate, “grammaticalization” in the early days was a topic of investigation within typology as well as Indo-European linguistics. Yet the approach of the two fields was different: while typologists made an inventory of the structures attested cross-linguistically and attempted — if at all — to relate them chronologically (e.g. Humboldt), Indo-Europeanists in the 19th century primarily focused on data-oriented research and on the origins of grammatical forms as found in the early daughter languages and as reconstructed for the protolanguage. Yet despite the effort put into it, the results were at best speculative, mainly because of the lack of proof. This observation is a vital aspect of Meillet’s 1912 article. Meillet generally is referred to as the first scholar to use (and “coin”) the term ‘grammaticalization’ (Heine *et al.* 1991:8-9; Hopper and Traugott 1993:18; C. Lehmann 1995:1). He will be discussed in greater detail below.

In sum it is remarkable that the textbooks on grammaticalization refer to the work by typologists, but only superficially glance — if at all — at the work done by comparativists, who worked with actual data. It is even more remarkable that those that do refer to Bopp, ignore the debate about the results. This selective approach may give the mistaken impression that studies in grammaticalization solely are success stories.⁴

3. Indo-Europeanists and the origins of grammatical forms

The best known exponent of research on the origins of Indo-European grammatical forms was indeed Franz Bopp — the father of Comparative Linguistics. In his *Conjugationssystem* (1816) and more so in his *Vergleichende Grammatik* (1833 [1999]), Bopp compared Sanskrit, Greek, Latin, Zend, Gothic, German, and Lithuanian, establishing on the basis of formal parallels a genetic relationship between the languages included, as for finite verb forms: cf. Skt *dádihāti* - Gk *δίδωσι* ('he gives') - La. *amat* (with loss of *-i*, 'he loves') - Go. *habaiþ* ('he has', La. *habet*). His aim was also to analyze the "Ursprung der die grammatischen Verhältnisse bezeichnenden Formen" (1999 [1833]:iii). Consequently, in an attempt to trace the origins of the forms above, Bopp also put forth a hypothesis by which the ending *-ti*, for example, is identified as a third singular demonstrative personal pronoun, parallel to Gk *-σι*, in a way similar to Skt *tvām* paralleling Gk *σύ* (La. *tu*), and so forth (Bopp 1999:659; for the pronoun, see [1999:489]). On the whole Bopp submitted strong parallels between endings and the isolated pronouns (e.g. Bopp 1999:109) and emphasized the main principle of Sanskrit and the other early daughter languages: that of combining verbal and pronominal "Wurzeln" (1999:112-113). Bopp's approach had many followers (e.g. Pott and Schleicher [cf. e.g. Delbrück 1884 (1880)]) and the origins of

⁴This inaccurate impression may partially be accounted for by Gabelentz's sweeping reference to Indo-Europeanists in the relevant pages: "im Wesentlichen dürfte die sogenannte Agglutinationstheorie, wie sie heute wohl von allen Indogermanisten angenommen ist, unumstößlich und gemeingültig sein; alle Affirmativen waren ursprünglich selbständige Wörter" (Gabelentz 1901:256).

grammatical forms was a recurring topic in 19th century research.

3.1 Bopp and the Neo-Grammarians

The year 1878 is a landmark in the history of Indo-European linguistics. In the “Vorwort” to their new journal *Morphologische Untersuchungen auf dem Gebiete der indogermanischen Sprachen*, Brugmann and Osthoff defined the aims and methods of the Neo-Grammarians. The importance of this text may be reflected in its common qualification as being the “*Manifesto*” of the Neo-Grammarians. The article embraces the idea put forth by Scherer (1868) that so-called “recent” languages are not merely degenerated forms of their ancestors, but languages in their own right that provide the comparative linguist with crucial information. Their documented history allows us to evaluate and understand how language works and changes. In line with this observation, Brugmann and Osthoff rejected the exclusive focus on the protolanguage and advocated the inclusion of recent languages in linguistic analysis because of their continued documentation in terms of change over time and the existence of real-living systems as found in the dialects (Brugmann and Osthoff 1878:vii-ix). While we need the earliest attestations for the reconstruction of the protolanguage, we need other — later — data for our analysis of the evolution of language and the renewal of grammatical forms. Neither the earliest attested forms nor the reconstructed forms would suffice. The more so since the reconstructed forms are “alle rein hypothetische gebilde” (1878:vi). Consequently Neo-Grammarians should not participate in the ““idealistischen flug” in die ursprachlichen und vorursprachlichen zeiträume, wie es jetzt schon so vielfach gewagt wird” (Brugmann and Osthoff 1878:xviii). With this statement Brugmann and Osthoff also rejected the study of the origins of (reconstructed) grammatical forms as an area of investigation.

Shortly after the *Manifesto*, Delbrück as well criticized Bopp’s reconstruction of the autonomous origins of inflected forms (1884; especially Chapter 5), and more openly so: Delbrück concluded his discussion of Bopp’s

analysis of noun and verb inflection (mood, tense, person marking): “im besten Fall hat sich uns für die Einzelanalysen eine gewisse Wahrscheinlichkeit, nicht selten das kahle “non liquet” ergeben” (Delbrück 1884:100). Below, we will give further details about Delbrück’s position.

In a similar vein Hirt took a position against Bopp, who “hat auch Vermutungen über die Herkunft der Personalendungen ..., die heute in der eigentlichen Sprachwissenschaft ... wenig geschätzt werden” (Hirt 1904:36-37). One of the reasons for Hirt to reject Bopp’s interpretation was the lack of formal parallels between the personal endings (Hirt 1904:37).

Brugmann expanded his criticism in his *Grundriß* (1886-1900),⁵ summarizing the hypotheses of Bopp and his followers: “nach der Ansicht Bopp’s und der meisten seiner Nachfolger entsprang diese Formkategorie in der uridg. Zeit dadurch, daß eine ein Sein oder ein Tun bezeichnende Wortform und ein persönliches Pronomen, ..., sich zu einer Worteinheit verbanden” (Brugmann 1913:5-7). Stating that “grundsätzlich ... gegen die Anschauung nichts einzuwenden [ist]”, Brugmann questioned the scenario whereby “Suffigierung solcher Pronomina” led to the entire system of finite forms. He also noticed that many other formations thus “explained” were distinctly different. Finally, observing a difference between 1st and 2nd person and 3rd person verb forms, Brugmann also had his doubts about the actual choice of the individual pronominal elements (Brugmann 1913:6-8). Consequently Brugmann both in the *Manifesto* and his *Grundriß* dissociated himself from early attempts at tracing the origins of Proto-Indo-European grammatical forms, emphasizing the lack of data and methodological rigor.

Subsequently Meillet, in his well-known 1912 (1982) article assessed Bopp’s research: “Bopp croyait que l’examen des plus anciens types de chaque idiome lui donnait le moyen de remonter à des formes en quelque

⁵The first edition of the *Grundriß* was published between 1886 and 1900 (five volumes). A second edition (1897-1916) included an enlarged version of the parts on phonology and morphology. The volumes on syntax, by Delbrück, had no second edition (cf. W. Lehmann 1993:50).

sorte primitives, susceptibles d'être analysées en leurs éléments composants ... Mais, pour une analyse comme celle-ci, qui est plausible — quoique naturellement indémontrable — Bopp était amené à en proposer cent autres qui étaient ou peu vraisemblables ou tout à fait inadmissibles” (Meillet 1982:131). Referring to the work by Bopp's followers, Meillet continues: “après une cinquantaine d'années d'**essais infructueux de ce genre**, on a compris que l'origine première des formes grammaticales est hors de nos prises” (1982:131-132; emphasis added).

In fact, Meillet's (1912) article underscores Brugmann's call for research on extant and documented languages. Moreover the article was strongly innovative for several reasons. Denouncing attempts to identify the original composing elements of the **earliest** grammatical forms, Meillet strongly advocated analysis of “les transformations des systèmes grammaticaux”, including their origins (1982:132). Moreover, Meillet explicitly distinguished between grammaticalization and analogy, which at the time was the main focus of Indo-Europeanists, as is clear e.g. in Brugmann and Osthoff (1878:xiii; xv). Meillet's emphasis on “grammaticalisation” parallels its primordial role in language innovation: processes of “grammaticalisation” effectively create new forms, whereas analogy will never point to the “origine première” of forms, because this type of change inherently will follow an already existing (predominant) pattern in the linguistic system and the system therefore remains the same. In grammaticalization by contrast we find innovation, both in terms of form and grammatical function or category: the system may change. Meillet explained the historical emphasis on analogy referring to the dearth of data in the early Indo-European languages (1982:133). Consequently while the texts by Brugmann and Osthoff (1878) and Meillet (1912) both underscore the importance of documented languages for the analysis of language change, the focus with Brugmann is on analogy, but with Meillet it is on grammaticalization because of its inherently innovative character.

In conclusion, the Neo-Grammarians did not embrace the attempts by Bopp and his followers although they acknowledged formal parallels between certain endings

and autonomous elements (e.g. person indication on verbs; e.g. Delbrück 1884:61). These were, however, not systematic enough to be conclusive or to account for the emergence of the entire system.

All Indo-Europeanists did not judge Bopp in a similar way nor follow Brugmann's "ban" on origins of grammatical forms. A case in point is Max Müller, who was rather influential and well-known with the public at large. Well aware that not all forms are equally easy to account for (1862:219), Müller argued that "however complicated the declension's regular and irregular forms may be in Greek and Latin, we may be certain that originally they were formed by this simple method of composition" (1862:218), referring to the combining of a root and an independent word. And "though in the present state of our science it would be too much to say that all grammatical terminations have been traced back to original independent words, so many of them have, even in cases where only a single letter was left, that we may well lay it down as a rule that all formal elements of language were originally substantial" (Müller 1899:346). Consequently Müller did not refrain from tracing the origins of early Indo-European grammatical forms and he aligned these with the more recent ones, as found in well-attested languages, such as the Romance languages.

3.2 Freezing, particles, and the origins of grammatical forms in Indo-European

Instead of stating that Proto-Indo-European inflectional and derivational endings trace back to independent elements such as pronouns, nouns, or other autonomous items, the handbooks by Brugmann, Delbrück, and Hirt provide ample evidence showing that particles and freezing play a major role in the origins of Indo-European morphology.

In Indo-European freezing is most commonly attested in the context of adverbs and conjunctions. The majority of adverbs, for example, trace back to frozen case forms, which in addition often are isolated because the rest of the paradigm has disappeared (Brugmann 1911:115). Brugmann provides an extensive overview of adverbs in the various early daughter languages, including frozen

nominatives through instrumentals (1911:671-720). Another process of adverbial freezing pertains to adpositions in combination with inflected nouns e.g. La. *denuo* ‘again’ < *de* + *nouo* ‘new’, *obviam* ‘against’ (< *ob* + *viam* ‘way’), Gk *ἔνδον* ‘within’ (< *ἐν* + *δομ-* ‘house’), Gm *zuhanden*, and so forth (Brugmann 1911:721-727). A smaller group is formed by adverbs that are isolated elements, either with particles (e.g. *dh*, *t*, *s*) or without (e.g. **ne* ‘not’). Freezing in relation to conjunctions will be discussed in Section 4.

Particles are class-less elements that convey a certain semantic value. Because their grammatical category is elusive, particles combine with any part of speech: (pro)nouns, verbs, adverbs, and so forth. Moreover they do not convey distinct grammatical categories nor have a clear formal status: they may occur as independent items or attached to autonomous elements. Instead of having a distinct meaning, they convey a broad notion that we may be able to grasp only by comparing their different uses — if necessary across the various Indo-European languages. The PIE particle **-i*, for example, has deictic value, but rendering its meaning as ‘here’ would be too restrictive for the uses it covers. Instead Hirt proposed ‘hic et nunc’, conveying the notion of ‘here’ both in space and time (1904:46; see more below).

In the early daughter languages such as Hittite, Sanskrit and Greek, we find many particles, reflexes of which may be attested e.g. in endings in the other daughter languages. These particles occur both as enclitics and as independent elements (cf. Friedrich 1960:147-154 on Hittite; Denniston 1954 on Greek). In Greek, particles reflecting a later stage mainly are independent elements that have certain values, such as emphatic ($\delta\epsilon/\gamma\epsilon$), responsive ($\mu\epsilon\nu\ \delta\epsilon$), apodotic ($\alpha\lambda\lambda\acute{\alpha}$), and so forth (cf. Denniston 1954). Denniston defines Greek particles as “a word expressing a mode of thought, considered either in isolation or in relation to another thought, or a mood of emotion: (1954:xxix), but “few Greek particles possess one meaning” (1954:lvi). Cognates often take the form of various parts of speech, such as conjunctions, adverbs, adpositions, pronouns, cf. e.g.: PIE **anti/*anta* is reflected

in Hittite *hantezzi* ‘the first...’ and the adverb *hantezzi* ‘first’ (Friedrich 1960:61; 1967:85), the Greek preposition *αὐτί* ‘opposite’, La. *ante* + Acc., and OHG *andi*, as in *andalanui* (Gk. ἀντιμισθία ‘requital’).

Moreover, numerous particles are integrated in Indo-European morphology, both in inflection and derivation, with the same element combining with a variety of word classes. The particle **-i*, mentioned earlier, for example, recurs in pronouns such as Greek dative *μοί/σοί* ‘for me/you’, adverbs as in Gk *ἐνί* ‘therein’ or La. *ita* ‘so’ or *ibi* ‘here’, and in nouns featuring locative meaning, not only in Greek (*οἴκοι* ‘at home’), but also in Latin, as in *Romae* ‘in Rome’, *domi* ‘at home’ (Meillet and Vendryes 1924:409; Lehmann 2003:93-96). The suffix marked location in time as well, cf.: e.g. Gk *νυκτί* ‘at night’ (Hom., *Od.* 15.34), La. *vesperi* (Meillet and Vendryes 1924:516). Moreover because of its deictic value, **-i* is also related to Latin anaphoric *is* (Brugmann 1916:979-980).⁶

Similarly the particle PIE **-bʰ-* features both in case endings and adpositions. In Mycenaean the element *-pi* occurs most commonly with plural nouns and conveys the value of instrumental, ablative, or locative (Ventris and Chadwick 1956:86-87). In Homeric Greek we find *-φι* in singular and plural nouns, conveying genitive, dative, ablative, instrumental, and locative function (e.g. Hirt 1904:51-52; Smyth 1956:71; Lehmann 2003:56). In Greek the element therefore is number-independent and conveys a variety of functions, which reminds us of an adposition rather than case suffix. As suffix, PIE **-bʰ-* is attested in plural case endings in Indo-Iranian (**-bhyas* <

⁶The question “why should all these different elements *-i* be just one particle **-i* rather than, say, four homonymous particles?”, which was brought up by one of the anonymous referees of this article, is interesting. The current state-of-the-art does not allow us to tell with certainty which interpretation prevails. Yet there are several considerations that support the hypothesis that we are dealing with a single element: (a) the different uses are semantically close; (b) location in time and space is often conveyed by the same case ending in the early daughter languages (e.g., the locative; Meillet and Vendryes 1924:515-516); (c) the dative and locative are historically connected; and (d) other particles as well combine with a variety of forms of speech.

*-*bhyos*; e.g. Av. *urvarābyas* ‘to plants’) and Italic (*-*bhos*, e.g. La. *legibus* ‘with laws’), which function as dative-ablatives. In Germanic — which features a dative plural in *-*mis* — the particle is attested in the Gothic preposition *bi* ‘at, around’ (Gm. *bei*, OE *bi*, Engl. *by*, and so forth; Hirt 1904:51; Meillet 1964 [1903]:298-299; Meillet and Vendryes 1924:424; 466; 476; Szemerényi 1990:196-197; Gamkrelidze and Ivanov 1995:334; Baldi 1999:316-317). In the history of PIE *-*b^h*- we therefore find a two-pronged development whereby the particle *-*b^h*- functions as an adposition in certain daughter languages and as a case suffix in others. These and similar instances do not reveal a scenario whereby a postposition turned into a suffix. In fact, although it may be rather tempting to assume that endings of so-called concrete cases in Indo-European originally were postpositions, there are no data to identify them — with a few exceptions: e.g., the “secondary cases” in Lithuanian (Senn 1966:92-93): the allative, adessive, and illative, which were “recently created cases” (Endzelins’ 1971:166; 166-167).⁷ Instead we find particles.

So far, we have exclusively examined the formal origins of inflected forms, such as cases. The semantic aspect of the story has not been addressed. Yet a case system is not something that spontaneously emerges out of the blue. A fundamental question to address therefore is: how did the case system come about in the absence of case endings?

In 1958 and then again in 1993 and 2003 Winfred Lehmann put forth a scenario for the emerging case system in Indo-European. He posited that grammatical cases originated in a system in which the noun came to combine with a number of particles each with a given semantic value. The particles in question were -*s* conveying an individual or agent, -*m* indicating a target or product,

⁷In am grateful to Martin West (Oxford) for pointing out this example (PC, March 2013). In Old Lithuanian the allative was “formed by adding the postposition -*pi* to the genitive case, whereas the adessive was formed by adding this postposition to the locative case” (Endzelins’ 1971:167). The illative “attaches postposition =*na* to the accusative” (Hewson and Bubenik 2006:206). In today’s Lithuanian these secondary forms survive in certain dialects only, often as lexicalized elements or expressions (cf. Hewson and Bubenik 2006:206-207).

and *-h* expressing collective value. This set of particles reflects “a system in the process of development, with an earlier set of nouns marked for semantic categories and subsequently a shift into a case system” (2003:155). Residues of the system are still identified — according to Lehmann — in isolated lexical items, as in Sanskrit, where on the basis of *hima-* three different words are composed, each with one of the particles, cf.:

| | | |
|------------------------------------|--|-----------------------------------|
| Skt <i>himás</i> (M) ‘cold, frost’ | (< <i>himá</i> + <i>s</i>) | ‘that which is cold, frost’ |
| <i>himam</i> (N) ‘snow’ | (< <i>hima</i> + <i>m</i>) | ‘the product of cold’ |
| <i>hímā</i> (F) ‘winter’ | (< <i>himah</i> , with loss of laryngeal and compensatory lengthening) | ‘the collection of cold entities’ |

(Example from Lehmann 1958:185)

Similarly, Skt *citra-s* ‘name of a king’ vs. *citra-m* ‘splendor’, *mitrá-s* (M) ‘friend’ vs. *mitrá-(m)* (N) ‘friendship’, and so forth (Lehmann 1958:190; Elizarenkova 1995:105). In other languages as well we find residues of this early situation, cf. La. *jugum* (N) ‘yoke’ (product) vs. *conju(n)x* (M) ‘spouse’ (agent). Moreover the early situation may be reflected in divergent developments in the different languages, resulting in cross-linguistic doublets: cf. La. *pēs* (M) ‘foot’ vs. Skt *padám* ‘footprint’ or Skt *āyus* ‘living being, man’ but La. *aevum* ‘life’ (Lehmann 1958:190).

It is important to underscore that in the original system these forms are assumed to represent “four different paradigms with characteristically different meanings” (Lehmann 1958:191). The combinations of noun + particle resulted — therefore — in my view in derivational rather than grammatical entities. It is when these individual elements came to represent different forms of the same paradigm with distinct *grammatical* meanings that the case system came into being, with the form in *-s* becoming nominative, and *-m* becoming the marker of the accusative singular of animate nouns and the nominative/accusative singular of neuter nouns (Lehmann 2003:185-186).

Lehmann’s reconstruction of the emergence of case — rooted in the analysis of the early handbooks — is interesting in that it posits a transition that is plausible: a

shift from a derivational system in which particles express certain semantic non-grammatical values to a system in which these elements increasingly become grammatical. The new grammatical functions are in line with the original lexical value of the elements in question: the particle marking an individual agent (-s) becoming the nominative marker and the particle expressing patient or target (-m) becoming the accusative ending, conveying the direct object. The various types of residue in the early daughter languages support the scenario. Once the grammatical cases were established, the system expanded and came to include concrete cases as well, which accounts for the consistency of nominative and accusative marking in the daughter languages and the lack of consistency in the other cases. Further systematic analysis is required, however, to assess the accuracy of this scenario. But if accurate, the development would involve a shift whereby derivational items become grammatical elements. We may consider this a major process of grammaticalization, but different from what we “normally” find.

While details require further investigation as said, it is legitimate to conclude that particles played a primordial role in the formation of endings in Proto-Indo-European. Inflectional paradigms therefore in likelihood were not the result of grammaticalization processes by which a **lexical** autonomous element developed grammatical functions (e.g., pronoun becoming a suffix marking person on the verb) or a grammatical element increased its grammatical value (e.g., a postposition becoming an inflectional suffix). Consequently, we cannot posit that grammaticalization processes as we know them from the later Indo-European languages account for Proto-Indo-European inflection.

Instead freezing, particles, and the potential shift from derivation to grammatical forms seem to be at the core of Indo-European grammatical formations. In the next section we will assess the grammatical forms in the various stages and varieties of Latin

4. Grammatical forms in Latin

In this section we discuss grammatical forms in Latin: the underlying grammatical processes and composing

elements. We will compare the early forms to what we found in Indo-European (Section 4.1). Subsequently we examine new formations that gradually emerged in the various stages of Latin, and that survived as grammatical items in Romance (Section 4.2).

4.1 Early stages

The starting point for our analysis is Fruyt's (2011) overview of "grammaticalization in Latin" with numerous examples, even if perhaps not all qualify as such. Nouns such as *agricultura* are identified as instances of "agglutination" and "transcategorization" ("*agri* + *cultura*", Fruyt 2011:676; 'land' + 'culture') — i.e. a nominal "syntagm" becoming a noun — and therefore as instances of grammaticalization: "one could say that every case of agglutination or lexicalization is a case of grammaticalization" because of the "downgrading" of the elements involved (Fruyt 2011:673-675). In my view, nouns such as *agricultura* are plain nominal compounds reflecting an underlying Noun Phrase⁸ and the question whether a noun is "more grammatical" than a Noun Phrase is open to discussion. Similarly, verbs like *manumittere* 'release from one's power' (< *manu-* 'hand' + *mittere* 'send') primarily are verbal compounds, hence lexical rather than grammatical: neither *manumittere* nor *agricultura* convey grammatical categories. Along the same lines, we may wonder according to what criteria adverbs that are frozen case forms (e.g., *modo* 'only, but' [< *modus* 'way'] or *partim* 'partly', an old frozen accusative of *pars* 'part')⁹ are more "grammatical" than the original inflected nouns or adjectives (see also Section 3.2). In terms of "parts of

⁸The status of Latin *-i* in formations such as *agricultura* or *agricola*, is a topic of discussion. In Latin as opposed to Romance compounds the first components typically are thematic stems (cf. Bauer 2011, with references, Forthcoming).

⁹The old accusative suffix *-tim/-sim* is also attested in *paullatim* 'gradually' (< *paullus* 'very little') and in other formations: *cursim* 'rapidly' (< perf. stem *curs-us* 'run'), *passim* 'scattered about' (< perf. stem *pass-us* 'spread'), *viritim* 'individually' (< *vir* 'man'), or *tuatim* 'after your manner' (< *tu* 'you'). The variety of bases suggests that the frozen case ending at some point was productive in adverbial formations, even if the precise morphological scenario is not quite clear (for more forms, see, e.g., Meillet and Vendryes 1924:477).

speech” an attributive adjective should be similar to an adverb: both are qualifying elements. One could argue that loss of formal variation — freezing — is a criterion. Yet while the process of grammaticalization may indeed include freezing, not all freezing reflects grammaticalization. The “freezing” of *manu-* in *manumittere*, for example, does not provide *manu-* nor the compound with grammatical function. Yet the acquisition of grammatical function is decisive in grammaticalization. *Manu-* therefore is fundamentally different e.g. from the Romance adverbial suffix *-mente*, originally a frozen case form as well. The questions raised here need further discussion, but space does not allow us to do so in this publication. In the following pages we will assess the most important types of grammatical form in Latin in order to pinpoint certain fundamental characteristics.

Inflection. Many of the declensional and conjugational formations were inherited and discussion in Section 3.2 has shown that the role of particles — rather than autonomous elements — was prominent in their emergence. Yet among Latin verbal forms there are two paradigms that seem to be relatively transparent and need further discussion. In the future forms of 1st and 2nd conjugation verbs (e.g. *lauda-bo* ‘I will praise’), the first part is recognized as an imperfective stem, whereas the precise nature of *-bo-* is not quite clear: it is cognate to La. perfective *fui* (‘I have been’) and Engl. *be*, for example, and identified by some as an aorist subjunctive (Meillet and Vendryes 1924:274), by others as “reflecting some form of the Proto-Indo-European subjunctive (details are obscure)” (Baldi 1999:398), or as an “element *-b-* with the ending of *ero* ...” (Sihler 1995:558). The formation “is partly transparent and partly opaque ... [but] obviously the root is **bhū-*, as in *fui*” (Sihler 1995:558). Similarly, the Latin imperfective forms in *-ba* (e.g. *lauda-bam* ‘I praised’) combine an element *-ba-*, which often is associated with the future element *-bo-*. The imperfective stem with which *-ba-* combines (*laud-* not *laudav-*; but see Sihler’s hesitations for *capiebam* [1999:555]), “has been derived alternatively from an infinitive (**amasī*), a root form (**ama*), and a participle (*amantis*)” (Baldi 1999:397). While Meillet and Vendryes did not take a clear stand and

only suggested a possible link with the morphological base of the infinitive (1924:575), for Baldi the “most reasonable interpretation is participle + inherited optative ‘be’” (Baldi 1999:398). Yet without motivation or attested forms, this interpretation remains hypothetical. Sihler as well suggests that “among the known verbals of Latin, the likeliest candidate for the original stem is the pres[ent] p[articiple]. But phrases which coalesce into single phonological words undergo changes for which there are no testable hypotheses” (1995:555).

Consequently, the interpretation of verbal forms in *-bo-* and *-ba-*, which were not inherited from Proto-Indo-European and therefore were later developments, remains problematic: while the consonantal part (*-b-*) is related to the verb ‘be’, the interpretation of the distinctive vowel and the base remain rather elusive.

Adpositions. Adpositions in origin were autonomous mobile particles that came to be added as prefix to verbs (preverbs) in a process of univerbation or came to combine with a noun. Even if both processes are similar, the linguistic outcomes differ: adpositions are head of adpositional phrases, whereas preverbs do not change the grammatical status of the verb. The developments took place in the early stages of Proto-Indo-European and residues of the earlier situation are commonly attested in the early daughter languages as in Homeric Greek, where these particles function as adposition, adverb, or preverb (Chantraine 1953:84; Hewson and Bubenik 2006:3-9). Moreover, early word order patterns in Greek (Bauer 1995:87-88;129-131) and residues in the early stages of Latin reveal patterns that precede univerbation (e.g. *ob vos sacro* ‘I beseech you’ [Paul., *Festus* 218])¹⁰ for later *obsecro vos*; Meillet and Vendryes 1924:521;527).

Latin prepositions originally may also be frozen participles, both present (e.g. *trans* ‘through’ from a Latin verb **tra-re* [cf. *intrare*; Fruyt 2011:700])¹¹ and perfective (e.g. *adversus* ‘against’ (< *adverto* ‘turn’; Fruyt

¹⁰Paul. *Festus*: epitomized version by Paulus Diaconus of Sextus Pompeius Festus’s version of *De verborum significatu* by the grammarian Verrius Flaccus (Lindsay 1913). The text is also referred to as *Paulus ex Festo*.

¹¹Ernout and Meillet rejected this interpretation (1959:699-700).

2011:693;700; Ernout and Meillet 1959:725-726). Frozen case forms also provide an important source of adpositions as examples like *causa* ‘reason’ > ‘on account of’, *gratia* ‘favor’ > ‘on account of’, *tenus* ‘cord’ > ‘as far as’, and several others clearly illustrate (Fruyt 2011:700; Ernout and Meillet 1959:685, who identify the etymology of *tenus* as uncertain). Several of these instances of grammaticalization were early and the original noun did not always survive in Latin; others emerged during the Latin period, e.g. *causa* or *gratia*. These last adpositions typically occur in postposition (Bauer 1995:134-135). Fruyt also includes “quasi-prepositions” in this context (e.g. *a fronte* ‘in front’, *ante frontem* + Gen. ‘in front of’, 2011:700-701), which were rather late.

Conjunctions. Several conjunctions trace back to an earlier finite verb form, such as *vel* (< *velli* < *vel-si*, 2nd sg. of *volo* ‘want’; Fruyt 2011:678). Yet for Ernout and Meillet “la forme fait quelque difficulté ... **weli* ne fournit pas d’explication sûre” (1959:718). According to Fruyt, *igitur* is an earlier finite verb form as well (Fruyt 2011:694), but Ernout and Meillet rejected this interpretation and posited that the form is “sans étymologie” (1959:307).

For subordinating conjunctions the origins are more certain: several originally were correlatives, such as *cum* (*quom*) ‘when’, which traces back to correlative *quom/cum* ... *tum* (Ernout and Meillet 1959:561; Fruyt 2011:678). Subordination is an acquired feature in Indo-European and in early times we find paratactic constructions instead, as well as participles and correlative constructions (Haudry 1973; Bauer 1995:259-265). A question to address is to what extent a subordinating element is more “grammatical” than a correlative one. The construction *quom/cum* ... *tum* ... originally connected two co-ordinating terms (‘not only... but also’)¹² or two verbs conveying simultaneous actions. At a later stage when *cum* came to prevail as a subordinating element, it expressed temporal meaning with the indicative but concessive-causal value with the

¹²E.g.: *consilia cum patriae tum sibi inimica capiebat* (Nep. 4.3.8) ‘he conceived plans that hurt both his country and himself’. *Quom/cum* ... *tum* ... has parallel structures in *quam* ... *tam* ..., *quod* ... *id*, and others which also gradually developed into subordinate clauses (cf. Bauer 1995).

subjunctive. We therefore observe in this instance a shift from connecting to subordinating function, which profoundly affects the grammatical organization of clauses, and the emergence of semantic values closely linked to mood variation. These changes therefore qualify as “grammaticalization”.

Negation. Latin’s negating particles present clear instances of grammaticalization. *Non*, tracing back to the combining of the particle *ne* with the numeral **oinom* ‘one’ > *noeuum* ‘not one’ (Meillet 1982:140; Fruyt 2011:709-712), is an excellent example of grammaticalization with its etymology still attested in Early Latin *noenum/noenu* (Ernout and Meillet 1959:444; Meillet 1982:140): *noenum mecator quid ego ... dicam ... evenisse* ‘by Castor, I cannot say what... has happened’ (Pl., *Aul.* 67). A similar pattern underlies *nihil*, which in origin was a combination of *ne* and the noun *hilum* ‘trifle, eye of a seed’, referring to a small object (Ernout and Meillet 1959:294-295). The occurrence of both *oinom* and *hilum* can be accounted for by emphatic use: ‘not even one/a trifle’.

The particle *ne* also functions in combination with other elements, as in *n-umquam* ‘not-ever’, *n-ullus* ‘not-any’, *ne-fas* ‘not-lawful’, *n-olo* ‘not-want’, *ne-scio* ‘not-know’, and so forth (e.g. Fruyt 2011:708-709;714-715). Yet here again the question to address is whether all instances reflect grammaticalization or simply a common form of word formation.

Demonstratives. Based on a three-way distinction (‘this here’ [*hic*], ‘that there with you’ [*iste*], and ‘that there with a third person’ [*ille*]), Latin demonstratives include a pronominal element that combines with particles indicating distance. In the singular *hic* includes the enclitic particle *-c(e)*, cognate to Latin *ecce* ‘see, here’ and marking proximity. In the plural we find it in nominative/accusative neuter *haec*, but also in early genitives *harunc/horunc*, in the accusative masculine *hosce*, and dative/ablative *hisce*. The base form is assumed to be identical to Latin anaphoric *is*, and cognate to Sanskrit *ayām* (e.g., Sihler 1995:393; Ernout and Meillet 1957:293). *Iste* includes an element *is*, which some qualify as particle (Ernout and Meillet 1957:324), others as pronoun (Sihler 1995:394), in combination with a demonstrative element (Ernout and

Meillet 1957:324) or deictic pronoun (Sihler 1995:394). The origins of *is*, which does not take inflection, remains vague. The demonstrative part of the formation (*-te*) is assumed to be related to Skt *ta*, Gk *το*, Go. *þa* and is inflected. Finally, *ille* as well includes two elements, which remain rather obscure in terms of origins. The deictic value seems to be present in the first element, which reflects **ol-* ‘that, yonder’ also found in La. *uls* ‘beyond’, *ultra* ‘on the other side, beyond’, *olim* ‘at that time’ (Ernout and Meillet 1957:309; Sihler 1995:393-394).

Finally *ipse*, another source for definite articles in Romance,¹³ combines the anaphor *is* with a particle *-pse*, which conveys the emphatic notion of ‘self’. In the early days the inflection was exclusively on the anaphor, cf. *ipse*, *eapse* (e.g. *ex eapse* [Pl., *Ep.* 254] ‘from the girl herself’), *eampse* (e.g. Pl., *Men.* 772), *eumpse* (Pl., *Most.* 346), etc. Only later did case come to be marked on the second syllable.

Adverbs. Adverbs reflect an important group of formations, many of which trace back to a frozen case form, predominantly of nouns but also of adjectives: “The freezing of a noun in the ablative singular into an adverb is also well documented and must have occurred productively throughout all periods of Latin” (Fruyt 2011:690). We find predominantly temporal and spatial adverbs in this category:

temperi Abl.sg., *tempus* ‘time’ > ‘at the right time’
vesperi Abl.sg., *vesper* ‘evening’ > ‘in the evening, late’

Both *temperi* (later *tempori*) and *vesperi* were early ablative sg. forms,¹⁴ cf.

tam vesperi domum revertor (Ter., *Heaut.* 67)
 ‘however late I return home’

¹³ *Ipse* survives as a definite articles in certain Southern Romance varieties: in Sardinia, parts of Sicily, and in certain dialects in Southern and Northern Italy. It is also attested — to various degrees — in Western varieties: in Provence, Gascony, and Catalonia. Today *ipse* survives most prominently in Sardinian.

¹⁴ The element *-i* occurs in certain early third declension nouns as a locative-ablative marker (cf. Ernout and Thomas 1964:96-97; see also Section 3.2 above).

Formations of this type include frozen accusative forms as well, as in:

foras Acc.pl., **fora* ‘door’ > ‘outside’

Frozen case forms with an adjectival base are found as well:

continuo Abl.sg., *continuus* ‘continuous’ > ‘immediately’

Since many of these formations include old case forms (see also the formations in *-sim/tim* above), it is clear that they are early.

The overview above covers the main types of grammatical element in Latin that often are qualified as results of grammaticalization. We observe the following trends:

- ∞ with the exception of declensions and conjugations, the large majority of grammatical forms in Latin are isolated instances, e.g. *vesperi*, *non*, *vel*, etc., each with a specific meaning;
- ∞ the origins of most forms resides in freezing: a case form or finite verb freezes and comes to be used as a grammatical element;
- ∞ adverbs often are the result of freezing and transcategorization. The grammaticalness of adverbs requires further discussion;
- ∞ adpositions trace back to either particles or frozen participles or nouns;
- ∞ several subordinating conjunctions originated in correlative elements. Their development affecting the grammatical organization of clauses qualifies as “grammaticalization”;
- ∞ declensions and conjugations are among the earliest — inherited — formations and their origins reside in the combining of lexical items with particles. We discussed the nature and possible evolution of particles in Section 3.
- ∞ the deictic elements in demonstratives are particles that are found in adverbs, adpositions, and declension as well, conveying the same deictic value;
- ∞ the development of *non/nihil* is exceptional in the total of the Latin phenomena discussed here: it combines a

grammatical particle with an originally autonomous lexical element conveying a meaning that is independent of its ultimate function. The difference between *non* and *nihil* resides in the inclusion of a numeral (**oinom*) as opposed to a noun (*hilum*).

In sum, we submit that with the exception of negating elements, grammatical forms in Latin continue the inherited practice of Indo-European: particles and freezing play a crucial role in their formation as reflected both in Latin's inherited inflectional morphology, demonstratives, adpositions, and the other isolated grammatical elements discussed here. Negators, including autonomous lexical elements, are an exception to that pattern: they are instances of grammaticalization as we know it from later Indo-European languages. In the next section we will discuss similar forms that emerged in the history of Latin and survived as new grammatical forms in Romance.

4.2 Innovation in Latin: new grammatical forms in the later stages

In post-classical Latin we find grammatical elements that continue to be the result of freezing, such as finite verbs turning into subordinators (e.g. *licet*) or adverbs becoming quantifiers or intensifiers (e.g. *valde*, *bene* [cf. Fruyt 2011:841-843; *passim*]). Moreover, in the transition from Latin to Romance a series of new forms emerges, similar to *non/nihil*, whose development involves elements that originally are lexical and autonomous: the noun *mente* in combination with qualifying adjectives turned into an adverbial suffix, possessive *habeo* became an auxiliary of tense in combination with infinitives and perfective particles, the noun *homo* evolved into a third person indefinite pronoun, nouns conveying small quantity became negators, the numeral *unus* became an indefinite article, demonstratives turned into definite articles and third person pronouns, possessive *habere* became an impersonal verb conveying existence, *facere* turned into an impersonal verb conveying weather conditions and time reference, adverbs *magis/plus* became degree markers, and the conjunction *quam* became a marker of comparative

reference, cf.:

| | |
|---|--|
| NOUN <i>MENTE</i> 'mind' La. <i>clara mente</i> 'with a clear mind' | > ADVERBIAL SUFFIX <i>-MENT(E)</i> > e.g. It. <i>chiaramente</i> 'clearly' |
| <i>HABEO</i> 'have' VERB OF POSSESSION La. <i>amare + habeo</i> | > <i>HABEO</i> AUXILIARY > e.g. Sp. <i>amaré</i> |
| <i>HABEO</i> 'have' VERB OF POSSESSION La. <i>epistulas scriptas + habeo</i> | > <i>HABEO</i> AUXILIARY > e.g. It. <i>ho scritto</i> |
| NOUN <i>HOMO</i> 'man, human being' La. <i>homo</i> | > INDEFINITE THIRD PERSON PRONOUN > e.g. Fr./Occ. <i>on</i> |
| NOUN [SMALL QUANTITY] La. <i>passum</i> 'step' / <i>micam</i> 'crumb' | > NEGATOR > e.g. [O]Fr./It. <i>mie/mica</i> , Fr. <i>pas</i> |
| NUMERAL <i>UNUS</i> 'one' La. <i>unus</i> | > INDEFINITE ARTICLE 'a' > e.g. Sp./Fr. <i>un</i> |
| DEMONSTRATIVE (<i>ille</i> 'that') La. <i>illum/illam/illos</i> | > DEFINITE ARTICLE > e.g. Fr. <i>le/la/les</i> |
| DEMONSTRATIVE (<i>ille</i> 'that') La. <i>ille/illa</i> | > THIRD PERSON SINGULAR PRONOUN > e.g. Fr. <i>il/elle</i> |
| <i>HABEO</i> 'have' VERB OF POSSESSION La. <i>libros habet</i> 'there are books' | > IMPERSONAL <i>HABET</i> + ACCUSATIVE ('THERE IS/ARE') > e.g. Sp. <i>hay</i> ; Fr. <i>il y a</i> |
| VERB <i>FACERE</i> 'make/do' | > IMPERSONAL <i>FACIT</i> + ACCUSATIVE [WEATHER/TIME] > e.g. It. <i>fa</i> |

| | |
|--|---|
| ADVERBS <i>PLUS, MAGIS</i> La. <i>magis/plus</i> | > COMPARATIVE/ SUPERLATIVE ELEMENT > e.g. Sp. <i>más</i> , Fr. <i>plus</i> |
| CONJUNCTION <i>QUAM</i> /PREP. <i>DE</i> La. <i>de/(tam) ... quam</i> | > COMPARATIVE REFERENCE > <i>quam/de</i> > e.g. Fr. <i>que/de</i> , It. <i>di</i> |

These forms emerged *grosso modo* in the same period, but it is not our contention that they were Late Latin developments. In Early and Classical Latin we often find early occurrences anticipating the development, but the grammaticalized forms materialize only much later. In fact, several of these developments came to a conclusion only in the individual Romance languages. The early attestations in Latin are found in a variety of documents and it is important to emphasize that the composing lexical elements then have their full-fledged lexical value, cf. the following instance featuring the noun *mente* and a qualifying adjective:

sed mente simplicissima et vera fide ...
but mind-Abl. simple-Superl.-Abl.sg.-F. and true-Abl.sg.-F. faith-Abl.
comites induxisse (Petr. 101.3)
companions-Acc. take-Pf.-Inf.

'he had taken (us) as his friends with a sincere mind and in good faith'

Similarly the example *epistulas scriptas habeo* did not mean 'I have written letters', but rather 'I have letters that are in the state of being written', cf.:

mares a feminis secretos habeant
males-Acc. from females-Abl. separate-PP-Acc. pl.-M. have-Subju.-3pl.

'that they keep the males away from the females' (Varro, *R.R.* 2.1.18)

In the linguistic literature fore-runners easily are interpreted as grammaticalized forms and instances are treated as if they all are identical (e.g. Karlsson 1981 [*mente*]). The processes of grammaticalization, however,

were long-term developments expanding over several centuries (see also Fruyt 2011:845-846).

Close analysis further shows that these developments typically do not reflect a smooth evolution. In earlier publications, I have examined the emergence of several of the grammaticalization phenomena in Latin/Romance mentioned above: *mente* as an adverbial suffix (Bauer 2001; 2003; 2010), the definite article (Bauer 2007; 2009), the auxiliary *habeo* in synthetic future and compound tense forms (Bauer 1995; 2006), impersonal *habet* constructions (Bauer 1999), and fore-runners of indefinite *homo* (Bauer 2014). The outcome of these analyses revealed the following patterns:

- ∞ the spread of evidence may be erratic, found e.g. in certain texts only of any given author;
- ∞ the new structure may first occur in unexpected (con)texts, rather than so-called “popular” texts (e.g. adjective + *mente* originating in poetry);
- ∞ the spread is not as such time-related: all structures do not emerge in Early Latin and then spread with time in exclusively popular varieties. Instances of *habeo* + perfective participle are attested in Classical texts (e.g. Cic., *Verr.* 5.74), but rarely so later, as in the *Peregrinatio* (one instance);¹⁵
- ∞ the new form competes not only with its predecessor, but — more so — with a host of alternative forms, e.g. *mente* vs. *animo* vs. *voce* vs. *pede*, ... ; or the nouns *passum*, *micam*, *guttam* occurring in reinforced negation and surviving in various (early) Romance dialects;
- ∞ the surviving structure may not be widespread. *Mente*, for example, was much less common in Adjective-Noun combinations than *animo*;

¹⁵This observation troubles the traditional picture of many Romance forms exclusively tracing back to “Vulgar Latin” rather than other varieties of Latin; it eventually rekindles the discussion of the origins of the Romance languages. In this light it is important to note that Adams’ (2013) study of thirty topics in Latin (including perfective participle + *habeo*) has revealed that “some Proto-Romance changes in Latin might have come from above” (2013:842). In fact, Adams has succeeded in “refining the view that it was Vulgar Latin that was the source of Romance” (2013:842).

- ∞ the spread of a structure in certain areas may not be indicative for its eventual survival. *Ipsa*, for example, was frequent in areas where *ille* eventually survived. For that reason, we have to identify factors other than frequency as driving force in grammaticalization processes.

Despite variation in data, occurrence, and development, the new grammatical forms share one important characteristic: they include a shift from an autonomous lexical or slightly grammatical into a fully grammatical element — a suffix, adposition, or an independent item. The type of structuration and branching may determine the precise outcome (e.g. Fr. *ai chanté* vs. *chanterai*; Bauer 1995; 2006). The shared autonomous origin distinguishes the forms from earlier grammatical elements, which were based on the inclusion of particles and freezing.

5. Conclusions

If we were to believe the historical overviews in textbooks on grammaticalization, attempts by early Indo-Europeanists such as Bopp at tracing the lexical origins of grammatical forms in Indo-European were very successful. Yet the Neo-Grammarians not only criticized these efforts, but in their handbooks also presented inventories of Indo-European composing elements that today contribute to our understanding of the early morphological system. They reflect the prominence of particles in early Indo-European, which eventually survived as independent elements or as part of inflected forms. Consequently, rather than lexical autonomous elements, particles seem to play a prominent role in the formation of early grammatical forms.

Another process underlying grammatical entities in Indo-European is freezing: an inflected form or another element “freezes” and comes to be used as a grammatical element (e.g. adpositions). Many adverbs originated in frozen case forms as well, but the question of why adverbs are more “grammatical” than the original noun or adjective needs further defining.

The overview of grammatical formations in (Early) Latin presented here has revealed the continued practice of strategies found in Indo-European: most forms are

either inherited (declensions/conjugations) — including particles as composing elements rather than lexical items — or the result of freezing. A few point to grammaticalization. While various conjunctions and negators — as we saw — indeed are the result of grammaticalization, it is important to note that these were isolated instances and not part of a paradigm.

Observing the prominence of particles in grammatical forms in Indo-European and (early) Latin, can we speak of grammaticalization? Particles were of uncertain grammatical status and had — often rather broad — semantic value. Integrated in endings these elements eventually acquired grammatical function. Since they were not in origin autonomous lexical or mildly grammatical items, they do not qualify as instances of grammaticalization. Yet if we accept the hypothesis by which case forms originated in derivation, then we can indeed argue that the development of case reflects grammaticalization, because lexical derivatives acquired grammatical function. For reasons of space we limited our discussion here to case forms.

While (early) Latin continues Indo-European formation practices, we find in Romance a series of grammatical forms that trace back to autonomous lexical or slightly grammatical elements in Latin. We saw that this series started in early Latin, with negating elements as the earliest fully-developed attestations. For the other forms we find early manifestations in the different varieties of Latin. Consequently the grammaticalized forms that emerged in the transition from Latin to Romance, were not merely “new”, they were fundamentally innovative as well. I do not argue that the merger of (quasi) autonomous elements did not occur before in Latin. It did, even if the first component of Latin compounds predominantly is a thematic stem rather than an independent element. More importantly, these formations were lexical. By contrast, the new formations in Latin/Romance that we discussed here, eventually resulted in grammatical forms.

That the new grammatical forms in the transition from Latin to Romance were fundamentally innovative not only sheds a different light on the processes of grammaticalization as we know it from the later attested

Indo-European languages, but also on the contributions of early Indo-Europeanists to the study of the origins of grammatical forms.

References

- Adams, James N.
 2013 *Social Variation and the Latin Language*. Cambridge: Cambridge University Press.
- Baldi, Philip
 1999 *The Foundations of Latin*. Berlin: Mouton de Gruyter.
- Bauer, Brigitte L. M.
 1995 *The Emergence and Development of SVO Patterning in Latin and French*. Oxford: Oxford University Press.
 1999 Impersonal *habet* constructions: At the cross-roads of Indo-European innovation. In: Edgar Polomé and Carol Justus (eds.) *Language Change and Typological Variation: In Honor of Winfred P. Lehmann on the Occasion of his 83rd Birthday*. Vol II. Grammatical Universals and Typology. Journal of Indo-European Studies. Monograph 31, 590-612.
 2001 Syntactic innovation in Latin poetry? The origins of the Romance adverbial Fformation in *-ment(e)*. In: M.G.M. van der Poel and A.P. Orbán (eds.) *Ad Litteras. Studies in Honour of J. H. Brouwers*, 29-43. Nijmegen (NL): Nijmegen University Press.
 2003 The adverbial formation in *mente* in Vulgar and Late Latin. A problem in grammaticalization. In: Heikki Solin, Martti Leiwo, and Hilla Halla-aho (eds.) *Latin tardif et latin vulgaire VI*, 439-457. Hildesheim: Olms.
 2006 “Synthetic” vs. “analytic” in Romance: The importance of varieties. In: Randall Gess and Debbie Arteaga (eds.) *Historical Romance Linguistics. Retrospective and Perspectives*, 287-304. Amsterdam: Benjamins.
 2007 The definite article in Indo-European: Emergence of a new grammatical category? In: Elisabeth Stark, Elisabeth Leiss, and Werner Abraham (eds.) *Nominal Determination. Typology, Context Constraints, and Historical Emergence*, 103-139. Amsterdam: Benjamins.
 2009 Strategies of definiteness in Latin: Implications for early Indo-European. In: Vit Bubenik, John Hewson, and Sarah Rose (eds.) *Grammatical Change in Indo-European Languages*, 71-87. Amsterdam: Benjamins.
 2010 Fore-runners of Romance *-mente* adverbs in Latin prose and poetry. In: Eleanor Dickey and Anna Chahoud (eds.) *Colloquial and Literary Latin*, 339-353. Cambridge: Cambridge University Press.

- 2011 Word formation. In: Martin Maiden, John Charles Smith, and Adam Ledgeway, eds. *The Cambridge History of the Romance Languages*. I. *Structures*, 532-563. Cambridge: Cambridge University Press.
- 2014 Indefinite *homo* in the Gospels of the *Vulgata*. In: Piera Molinelli, Pierluigi Cuzzolin, and Chiara Feciani, eds. *Latin vulgaire – latin tardif X*. 415-435. Bergamo: Bergamo University Press.
- Forthcoming. *Nominal Apposition in Indo-European. Forms – Functions & its Development in Latin/Romance*.
- Bopp, Franz
- 1816 *Über das Conjugationssystem der Sanskrit-sprache in Vergleichung mit jenem der griechischen, lateinischen, persischen und germanischen Sprachen*. K.J. Windischmann (ed.). Frankfurt am Main: Andrea.
- 1999 (1833) *Vergleichende Grammatik des Sanskrit, Zend, Griechischen, Lateinischen, Litthauischen, Gothischen und Deutschen*. I, II, III. London: Routledge.
- Brugmann, Karl
- 1911 *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen*. Vol. 2.2. *Lehre von Wortformen und ihrem Gebrauch. Nomina*. Straßburg: Trübner.
- 1913 *Grundriß der vergleichenden Grammatik der indogermanischen Sprachen. Vergleichende Laut-, Stammbildungs- und Flexionslehre nebst Lehre vom Gebrauch der Wortformen der indogermanischen Sprachen*. 2.3.1 *Lehre von den Wortformen und ihrem Gebrauch*. Straßburg: Trübner.
- 1916 *Grundriß der vergleichenden Grammatik der indogermanischen Sprachen. Vergleichende Laut-, Stammbildungs- und Flexionslehre nebst Lehre vom Gebrauch der Wortformen der indogermanischen Sprachen*. 2.3.2. *Lehre von den Wortformen und ihrem Gebrauch*. Straßburg: Trübner.
- Brugmann, Karl and Hermann Osthoff
- 1878 Vorwort. *Morphologische Untersuchungen* 1. iii-xx.
- Chantraine, Pierre
- 1953 *Grammaire homérique*. 2. *Syntaxe*. Paris: Klincksieck.
- Croft, William
- 1990 *Typology and Universals*. Cambridge: Cambridge University Press.

Delbrück, Berthold

- 1884 [1880] *Einleitung in das Sprachstudium. Ein Beitrag zur Geschichte und Methodik der vergleichenden Sprachforschung.* 2nd ed. Leipzig: Breitkopf.

Denniston, J.D.

- 1954 *The Greek Particles.* Oxford: Clarendon.

Elizarenkova, Tatyana

- 1995 *Language and Style of the Vedic R̥sis.* Albany: State University of New York Press.

Endzelīns', Jānis.

- 1971 (1948) *Comparative Phonology and Morphology of the Baltic Languages.* [Baltų valodu skaņas un formas] Translation of the original Lithuanian by William R. Schmalstieg and Benjaminš Jēgers. The Hague: Mouton.

Ernout, Alfred and Antoine Meillet

- 1959 (1932) *Dictionnaire étymologique de la langue latine.* 4^{me} éd. Paris: Klincksieck.

Ernout, Alfred and François Thomas

- 1964 *Syntaxe latine.* Paris: Klincksieck.

Friedrich, Johannes

- 1960 *Hethitisches Elementarbuch.* 1. Heidelberg: Winter.
1967 *Hethitisches Elementarbuch.* 2. Heidelberg: Winter.

Fruyt, Michèle

- 2011 Grammaticalization in Latin. In: Philip Baldi and Pierluigi Cuzzolin (eds.) *New Perspectives on Historical Latin Syntax.* 4. *Complex Sentences, Grammaticalization, Typology,* 661-864. Berlin: Mouton de Gruyter.

Gabelentz, Georg von der

- 1901 *Die Sprachwissenschaft.* Leipzig: Tauchnitz.

Gamkrelidze, Thomas V. and Vjačeslav V. Ivanov

- 1995 [1984] *Indo-European and the Indo-Europeans. A Reconstruction and Historical Analysis of a Proto-Language and a Proto-Culture.* [Indoevropskij jazyk i indoevropcejcy] Translation of the original Russian by Johanna Nichols. Berlin: Mouton de Gruyter.

Haudry, Jean

- 1973 Parataxe, hypotaxe et corrélation dans la phrase latine. *Bulletin de la Société de Linguistique* 68:147-186.

- Heine, Bernd, Ulrike Caudi, and Friederike Hünemeyer
 1991 *Grammaticalization: A Conceptual Framework*. Chicago and London: University of Chicago Press.
- Heine, Bernd and Tania Kuteva
 2002 *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Heine, Bernd and Mechthild Reh
 1984 *Grammaticalization and Reanalysis in African Languages*. Hamburg: Buske.
- Hewson, John and Vit Bubenik
 2006 *From Case to Adposition. The Development of Configurational Syntax in Indo-European Languages*. Amsterdam: Benjamins.
- Hirt, Hermann
 1904 Über den Ursprung des Verbalflexion im Indogermanischen. *Indogermanische Forschungen* 17. 36-84.
- Hopper, Paul and Elizabeth Traugott
 1993 *Grammaticalization*. Cambridge: Cambridge University Press.
- Humboldt, Wilhelm von
 1836 *Über die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluß auf die geistige Entwicklung des Menschengeschlechts*. Königl. Akademie der Wissenschaften. Berlin: Dümmler.
 1963 (1822) Über das Entstehen der grammatischen Formen, und ihren Einfluss auf die Ideenentwicklung. *Schriften zur Sprachtypologie* III, 31-63. Stuttgart: Cotta. [Paper read at the Akademie der Wissenschaften on January 17, 1822.]
- Karlsson, Keith
 1981 *Syntax and Affixation. The Evolution of Morphology in Latin and Romance*. Tübingen: Niemeyer.
- Lehmann, Christian
 1995 *Thoughts on Grammaticalization*. München: Lincom.
- Lehmann, Winfred. P.
 1958 On earlier stages of the Indo-European nominal inflection. *Language* 34. 179-202.
 1993 *Theoretical Bases of Indo-European Linguistics*. London: Routledge.
 2003 *Pre-Indo-European*. Journal of Indo-European Studies Monograph 49.
- Lindsay, Wallace M.
 1913 *Sexti Pompei Festi De Verborum significatu quae supersunt cum Pauli epitome*. Leipzig: Teubner.

Meillet, Antoine

- 1964 (1903) *Introduction à l'étude comparative des langues indo-européennes*. Tuscaloosa, AL: University of Alabama Press.
 1982 (1912) L'origine des formes grammaticales. *Linguistique historique et linguistique générale*, 130-148. Paris: Champion.
 (First published in *Scientia* 12).

Meillet, Antoine and Joseph Vendryes

- 1924 *Traité de grammaire comparée des langues classiques*. Paris: Champion.

Müller, Max F.

- 1862 *Lectures on the Science of Language Delivered at the Royal Institution of Great Britain*. London: Longman.
 1899 *The Science of Language*. Vol 2. London: Longman.

Rii, H.N.

- 1854 *Grammatical Outline and Vocabulary of the Oji-Language, with Special Reference to the Akwapim-Dialect, together with a Collection of Proverbs of the Native*. Basel: Bahnmaier.

Scherer, Wilhelm

- 1868 *Zur Geschichte der deutschen Sprache*. Berlin: Duncker.

Schlegel, August Wilhelm von

- 1818 *Observations sur la langue et littérature provençales*. Paris: Librairie Grecque - Latine- Allemande.

Senn, Alfred

- 1966 *Handbuch der litauischen Sprache*. I. *Grammatik*, Heidelberg: Winter.

Sihler, Andrew

- 1995 *New Comparative Grammar of Greek and Latin*. Oxford: Oxford University Press.

Smyth, Herbert W.

- 1956 *Greek Grammar*. Cambridge, MA: Harvard University Press.

Szemerényi, Oswald

- 1990 *Einführung in die vergleichende Sprachwissenschaft*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Van Linden, An, Jean-Christophe Verstraete, and Kristin Davidse (eds.)

- 2010 *Formal Evidence in Grammaticalization Research*. Amsterdam: Benjamins.

- Ventris, Michael and John Chadwick
1956 *Documents in Mycenaean Greek*. Cambridge: Cambridge University Press.
- Verhoeven, Elisabeth, Stavros Skopetas, Yong-Min Shin, Yoko Nishina, and Johannes Helmbrecht (eds.)
2008 *Studies on Grammaticalization*. Berlin: Mouton de Gruyter.
- Wegener, Ph.
1885 *Untersuchungen ueber die Grundfragen des Sprachlebens*. Halle: Niemeyer.
- Whitney, William D.
1875 *The Life and Growth of Language. An Outline of Linguistic Science*. New York: Dover.

Abbreviations, text references

- | | |
|----------------------|---------------------------------------|
| Paul., <i>Festus</i> | Paulus Diaconus, <i>Epitoma Festi</i> |
| Pl., <i>Aul.</i> | Plautus, <i>Aulularia</i> |
| Pl., <i>Ep.</i> | Plautus, <i>Epidicus</i> |
| Pl., <i>Men.</i> | Plautus, <i>Menaechmi</i> |
| Pl., <i>Most.</i> | Plautus, <i>Mostellaria</i> |
| Ter., <i>Heaut.</i> | Terence, <i>Heautontimorumenos</i> |