Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Discourse structure and narration

A diachronic view from Germanic

Edited by

Ulrike Demske

Barthe Bloom

Open Germanic Linguistics 503

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Open Germanic Linguistics

Editors: Michael T. Putnam, B. Richard Page, Laura Catharine Smith

In this series:

1. Lemke, Robin. Experimental investigations on the syntax and usage of fragments.
2. Larsson, Ida & Erik Petzell (eds.). Morphosyntactic change in Late Modern Swedish.
3. Hall, Tracy Alan. Velar fronting in German dialects: A study in synchronic and diachronic phonology.

ISSN (print): 2750-5588
ISSN (electronic): 2750-557X

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Discourse structure and narration

A diachronic view from Germanic

Edited by

Ulrike Demske

Barthe Bloom

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Ulrike Demske & Barthe Bloom (eds.). 2025. *Discourse structure and narration: A diachronic view from Germanic* (Open Germanic Linguistics 503). Berlin: Language Science Press.

This title can be downloaded at:
redefine\lsURL
© 2025, the authors
Published under the Creative Commons Attribution 4.0 Licence (CC BY 4.0):
http://creativecommons.org/licenses/by/4.0/
ISBN: no digital ISBN
no print ISBNs!
ISSN (print): 2750-5588
ISSN (electronic): 2750-557X
no DOI

Cover and concept of design: Ulrike Harbort
Fonts: Libertinus, Arimo, DejaVu Sans Mono
Typesetting software: XeLaTeX

redefine \publisherstreetaddress
redefine \publisherurl

Storage and cataloguing done by redefine \storageinstitution

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Contents

Preface iii

1 Grammar in narration(s)
Ulrike Demske & Barthe Bloom 1

2 The impact of narration – A diachronic view on German
Sonja Zeman 19

3 Structuring the narrative with preposed adverbial clauses: a study of the German and Dutch Ponthus adaptations
Barthe Bloom 47

4 Discourse structure and the reorganisation of the Icelandic aspectual system
Jordan Chark 85

5 Narration and clause-internal adverbs: an information-structural watershed in Old Icelandic
Hannah Booth 157

6 In consideration of Age and Genre: A diachronic corpus analysis of deverbal nouns in German verbonominal constructions and complex prepositions
Fabian Fleissner 207

7 Argument realizing verb second clauses in Early New High German narratives: A discourse-pragmatic account
Malika Reetz 249

8 Verb-initial declaratives in Old English prose: a functional analysis
Anna Cichosz 273

9 V2-Relatives in Old English
Bettelou Los & Stefano Coretta 307

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Index **344**

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Preface

The current volume is part of a research project looking into the relationship between grammatical means and discourse structure in German from a diachronic perspective. Of particular interest are word order patterns at the left edge of sentences to better understand the impact of verb placement as well as the shaping of the preverbal position on the composition of discourse and narrative structure. These relationships are being investigated using a sample of early prose novels in German written in the fifteenth and sixteenth centuries (*Wortstellung und Diskursstruktur in der Frühen Neuzeit* 'Word Order and Discourse Structure in the Early Modern Period', funded by the German Research Foundation, project number 456973946).

To enable an even deeper understanding of the relationship between grammar and narrative discourse, we invited colleagues working on similar issues to a workshop at the University of Potsdam. The present volume is the product of this workshop. It constitutes a significant extension of the research project in that it not only contains contributions on German but also addresses the relationship between grammar and narration in other Germanic languages. The spectrum of phenomena examined is furthermore significantly broadened by the volume: In addition to the influence of word order patterns, it deals with the discourse functions of adverbs, the verbalization of time, and the distribution of nominalizations in historical narrative texts, among other things. We see this volume as a stimulus for hopefully many more (diachronic) studies of narrative discourses in other (Germanic) languages, which will lead to a more profound understanding of the relationship between grammar and narration(s) in general.

Potsdam, December 2024 Ulrike Demske & Barthe Bloom

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Chapter 1

Grammar in narration(s)

Ulrike Demske[a] & Barthe Bloom[b]
[a]Universität Potsdam [b]Friedrich-Alexander-Universität Erlangen-Nürnberg

1 Introduction

The current volume deals with questions of information structuring at discourse level, focusing on narrative discourses. More precisely, it is about the contribution of grammatical devices to the organization of texts, as well as their diagnostic potential for the narrative text type. Although it is well-known that information packaging had a much greater impact on the distribution of grammatical patterns in historical stages of a language than it does today (Herring et al. 2001, Traugott 2007, van Kemenade & Milicev 2012, Bech & Salvesen 2014), so far many studies on the relationship between information structure and grammatical patterns do not go beyond the sentence level, in other words, they do not take into account the possible influence of the text type on the manifestation of certain grammatical patterns. How and to which degree changes in grammatical patterns correlate with or are affected by changes in discourse and narrative structure, how the two layers interact with each other and affect each other, and how such issues can be operationalized are still understudied. A noteworthy exception is Somers (2024), who presents a new perspective on the first literization of German. She takes into account, and thoroughly describes, the different needs of language of immediacy vs. distance and of oral vs. written communication (Koch & Oesterreicher 1985), as well as the sociolinguistic and cognitive context in which the early German texts were created. She considers the relationship of the use of oral language in context of distance to the linguistic strategies of e.g., integration, nominalizations, and deictic expressions and argues, amongst other

Ulrike Demske & Barthe Bloom. 2025. Grammar in narration(s). In Ulrike Demske & Barthe Bloom (eds.), *Discourse structure and narration: A diachronic view from Germanic*, 1–14. Berlin: Language Science Press. DOI: ??

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

things, that it is the literization of the language that drives the shift from more paratactic to hypotactic structures Somers (2024: 196).

This volume aims to shed more light on these issues by presenting eight papers, which address these questions more or less explicitly. As the research questions imply, all contributions take a historical or diachronic perspective. Another commonality between the studies is that they all focus on data from Germanic languages, as we assume that by comparing closely related languages, the relationships in question become more pronounced. Specifically, the languages in question are German, Dutch, English, and Icelandic. The introduction to this volume first presents some key concepts of narrative structure, focusing on narration as a text type or genre. This is followed by a discussion of narration in terms of discourse modes and discourse relations. In the final section of this chapter, the eight contributions are briefly presented.

2 Discourse structure and narration

From a linguistic point of view, narration can be viewed in three different ways: narration as a text type (Labov 1972), narration as discourse mode (Smith 2003), and narration as discourse relation (Asher & Lascarides 2003). The latter two will be introduced in Section 2.2, but we will first zoom in on the conception of narration as a text type, which surfaces, often implicitly, in all contributions in this volume.

2.1 Narration as text type

What makes a narrative? The definition and delimitation of the narrative as a (overarching) text type, or as a genre, is not straightforward. Not only form narratives a "fuzzy set" (Ryan 2006: 193), but even the archetype is heavily debated (e.g., Ryan 2017). What sets apart narratives from other text types is their purpose of telling stories (e.g., Labov 2013). It contrasts in this way to other genres, e.g., to argumentative and instructive genres, which are used for persuasion and teaching (Fludernik 2000). Two crucial components of narratives are put forward in the literature: (i) events, which are changes of states, and (ii) actors, or characters that engage in or are affected by these events (cf. Bal 2009[1985], Virtanen 2011: 56, among others).

Events in a story are not random, but are temporally and/or causally related to each other (Rimmon-Kenan 1983: 16-18, Sternberg 1990: 903). Whether and how these relations are presented as such varies and is a matter of *order* (Genette

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

1980[1972], 1983[1980]). Consider classical Icelandic sagas as an example, like *Brennu-Njáls saga*, which are typically organized by causally linked events: Although the settings and the characters focused upon shift and are regularly abandoned, the narrative follows the linear order from cause to action to consequent (Lönnroth 1976: 23, Korecká 2019). Often, this goes together nicely with a chronological representation of the story events. Anticipatory segments can be realized, but whenever they are, they tend to be anchored in the story-now, for example, in the form of prophecies and dreams (Lönnroth 1976: 31-32). This is illustrated by the segment from *Njála* in (1), in which the death of Þorvaldr is anticipated.

(1) *Yngvildr, móðir þeira, var hjá stǫdd; hon grét, er hon heyrði, er þeir hétu alþingisferðinni. Þorkell mælti: "Hví grætr þú, móðir?" Hon svarar: "Mik dreymði, at Þorvaldr, bróðir þinn, væri í rauðum kyrtli, ok þótti mér svá þrǫngr vera sem saumaðr væri at honum; mér þótti hann ok vera í rauðum hosum undir ok vafit at vándum dreglum. Mér þótti illt á at sjá at honum var svá óhœgt, en ek mátta ekki at gera."*

'Yngvildr, their mother, was present, and when she heard them promise to ride to the Allthing, she started weeping. "Why do you weep mother?" asked Þorkell. "I dreamed," she replied "that your brother Þorvaldr was wearing a red tunic and this tunic was so tight that it looked as if he had been sewn into it. He was also wearing red hose bound with shabby tapes. I was distressed to see him looking so uncomfortable, but I could not do anything to help him."' (Lönnroth 1976: 51-52).

In Early New High German narrative prose, it is similarly typical to follow the temporal-causal order that is representative of the order of the story events. On occasion, extra-diegetic segments may be inserted, which present retroflective or anticipatory comments. The latter is illustrated with the example in (2)

(2) *Nu laßen wir von den reden vnd komen zu Bernhart vom steynrotsche wie der zu hoffe qwame.*

'Now we leave that conversation and turn to Bernhart vom Steynrotsche, how he came to court.'

(Bloom et al. 2024: Ponthus, 42vb, translation my own)

Such anachronies are, however, quite short and do not normally intervene with the order of events. What does repeatedly re-occur is that multiple events happen simultaneously. While Middle High German narrations tended to be single-stranded narratives, i.e., they concentrated on one main narrative thread, with

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

the "now" located in the narrative act, the prose novel with multiple strands becomes more dominant in the early modern times (Zeman 2023). Such a change in narrative structure is expected to correlate with changes in the use and productivity of grammatical patterns and the freezing of expressions, like the one illustrated above, marking the transition from one strand to another.

The temporal organization of narratives may additionally exhibit variation in terms of *duration* and *frequency* (Genette 1980[1972], 1983[1980]). This is also captured by notions like *rhythm* Bal (2009[1985]) and *narrative speed* (Hume 2005). With duration, the rate at which the story progresses in relation to the narration is meant, and frequency concerns the repetition of events Genette (1980[1972]). Narrative speed is a technique that can be used to create tension or focusing attention on a particular aspect of the story (Hume 2005, Kukkonen 2020). It is by definition related to the linguistic representation: The more linguistic material is used to express a single event, the slower the story time progresses in comparison to the narrative time. Concerning frequency, an implicit relation exists to the grammatical patterns associated with the contrast between given and new information.

Besides events, the actors, or characters are a crucial component of the story. Typically, three roles are distinguished: major characters, i.e., the main protagonist(s), minor characters, and non-protagonists (Redeker 1987). Major characters re-occur regularly throughout the entire story and are the main protagonist(s) involved in the events. This tends to be the case for the Germanic literature, although Old Icelandic sagas systematically diverge from this pattern, as the main protagonist is typically introduced relatively late. *Njál* is for example only introduced in the twentieth chapter of the Brennu-Njáls saga, and *Gunnar* (a second main protagonist) does not occur until the ninetieth chapter. Main protagonists are typically introduced with a lot of linguistic material and by a proper name (Sanford et al. 1988), because they have to remain salient throughout the narration. Because main protagonists are repeatedly referred to and therefore remain highly accessible, they are often expressed by pronouns after their introduction (Ariel 2004: 93–96), at least within episodes. Minor characters interact with the main protagonist(s) and are typically only there for a part of the story, and non-protagonists do not contribute to the plot (Redeker 1987). Minor characters, therefore, tend to be introduced by less linguistic material, and their relation to the main protagonist(s) is often communicated. Moreover, they generally do not need to remain accessible for later. Therefore, they may be simply introduced by their role within the story (Sanford et al. 1988, Ariel 2004). This is illustrated in (3) from the life of Saint Æthelthryth, in which both Æthelthryth and her father

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

are introduced by a proper name in combination with a relatively long description (marked in bold). In contrast, Æthelthryth's two husbands (underlined) are only referred to by a broad gendered term ('men') and their role ('alderman') but remain unnamed.

(3) *Wē wyllað nū āwrītan, þēah ðe hit wundorliċ sȳ, be* ***ðære hālgan sancte Æðeldrȳðe þām Engliscan mǣdene, þe wæs mid <u>twām werum</u> and swā ðēah wunode mǣden,*** *swā swā þā wundra ġeswuteliað þe hēo wyrċð ġelōme.* ***Anna*** *hātte* ***hyre fǣder, Ēastengla cynincg, swȳðe Cristen man,*** *swā swā hē cȳdde mid weorcum, and eall his tēam wearð ġewurðod þurh God.* ***Æðeldrȳð*** *wearð þā forġifen <u>ānum ealdormenn</u> tō wīfe.*

'We now wish to write, although it may be surprising, about **the holy saint Æthelthryth, the English maiden, who was with <u>two men</u> and still remained a virgin**, as the miracles, which she often performed, prove. **Anna** was **her father's name**, **the king of the East-Angles, a very christian man**, as he showed with his deeds, and his entire line was honored by God. Æthelthryth was then married to <u>an alderman</u>.'

(Baker 2012: 255, translation my own)

The importance of characters in a narration does not only affect the choice of referring expression but can be connected to other types of grammatical structures as well. In her study on the Old English *þa* 'then', Wårvik (2013) argues that *þa*-clauses are primarily associated with major characters, and as such, they have a function of foregrounding and maintaining the main story line, despite their frequent co-occurrence with linguistic patterns that indicate temporal shifts.

What makes a story? A story is made up by actors and events, as we have seen above, but it is also the product of the recounting of events by an agent (Prince 1987: 58). The recounted events are filtered by, to quote Rimmon-Kenan (1983: 70), "the mediation of some 'prism', 'perspective', 'angle of vision'". This lens is what is often referred as the focalization. The focalizor ('who sees') differs from the narrator ('who speaks'), although they may coincide (Genette 1980[1972], Rimmon-Kenan 1983).

Narrative texts are thus multi-layered, containing both the descriptive content and the focalization and verbalization of that content. A wide variety of terminology has been employed to refer to the "real or fictitious events", to what is being told. We will here use the term (main) *story* (Genette 1983[1980], de Jong 2004)[1]. Two modes of presenting the story – which may not be confused with discourse

[1]Note that this may include a layer of focalization, that is, it may have been decided which events are considered relevant and which ones not.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

modes – have been traditionally distinguished, mimesis and diegesis. In essence, the idea is that there is a difference between imitation (*mimesis*) and narration (*diegesis*). In the French structural tradition, this opposition has been viewed as one between showing versus telling (Chatman 1978) and is primarily thought to be represented in the distinction between the narrative summary, expressed by the narrator, and the scene, in which characters embody and directly reflect the presupposed reality. However, in texts of the narrative genre, the entire story, including segments of direct speech, is mediated and achieved by means of diegesis (e.g., Rimmon-Kenan 1983: 108, Fludernik 1993: 28).

The presence of a narrator and focalizor makes it so that narratives lend themselves well for shifting viewpoints (see e.g., the contributions in Dancygier & Sweetser 2012). This is particularly interesting as perspective taking interacts with the use of tense and aspect (Fleischmann 1988). Schiffrin (1981) illustrates this for English, where the present tense can be used to report on past events in narrative segments, i.e., the historical present is used to represent speaker experience. German data are discussed by Zeman (2018), who argues that the preterite and the present tense contrasted in Middle High German regarding their potential of reporting the temporal sequence of events as well as their co-occurrence with *do* 'then'. Moreover, verbal tense has been exploited for many narrative structuring functions, such as distancing (Declerck 2003) and emphasizing the experience of story progression (Hoffman & Gaber 2023: 30). Furthermore, the concept of focalization may come in useful as well in relation to understanding languages of the bounded and unbounded type (Carroll & von Stutterheim 2002), because a focalizor may or may not have a particular endpoint in sight, motivating the selection of a particular verb form.

The discourse structure of a narrative is thus determined to a large extent by its temporal organization at the text level, as expressed in the rendering of the narrated events as well as in the relationship between narrated time and narrative time (i.e., duration and frequency). Likewise, linguistic means to express time and aspect are used to convey the narrative perspective.

2.2 Discourse modes and discourse relations

Temporal structuring also plays a decisive role at the level of text passages that consist of more than one discourse segment. Such sequences of discourse segments introduce situations of different kinds into the universe of the discourse. Four situation types are distinguished by Smith (2003):[2] Events, states, general

[2]The key factor in the classification of situation types is their aspectual features, which result from the interaction of the verb, its arguments and any tense adverbial that occurs.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

statives and abstract entities, giving rise to temporal discourse modes such as NARRATIVE, DESCRIPTION and REPORT as well as atemporal discourse modes such as INFORMATION and ARGUMENT, that are characteristic of texts in other genres (cf. above). Smith (2003) considers the discourse mode as an intermediate level of discourse structure between text level and discourse segment, which is made up of discourse segments exhibiting linguistic features typical of the particular mode. In her view, a narrative text consists primarily of text passages that belong to the NARRATIVE discourse mode, but will probably – due to their aspectual characteristics – also contain passages that belong to the DESCRIPTIVE discourse mode. The following excerpts from the novel *Effi Briest* by Theodor Fontane illustrate both the NARRATIVE and the DESCRIPTIVE mode of discourse: The text passage in (4a) exemplifies the NARRATIVE discourse mode, describing three situations that express bounded events (getting dressed, going there, sending a card), i.e., events having reached the end of their runtime. The second text passage, (4b), on the other hand, represents the DESCRIPTIVE discourse mode and describes Effi Briest's and her companion's lodgings in Ems. The occasional piano playing is part of the description of their sojourn in Ems, insofar as it is not a description of a particular but a recurring event. According to Smith (2003), we are dealing here with a generalizing sentence.

(4) a. *Gleich am andern Vormittag kleidete sie sich sorgfältig in ein*
Right at.the next morning dressed she herself carefully in a
dezentes Schwarz und ging auf die Linden zu, sich hier bei
discreet black and went towards the Linden PTCL herself here with
der Ministerin melden zu lassen. Sie schickte ihre Karte herein, auf
the minister report to let She sent her card in on
der nur stand: Effi von Innstetten geb. von Briest. (Effi Briest)
which only appeared: Effi von Innstetten née von Briest
'The very next morning she dressed carefully in a discreet black dress and went to the Linden to be put in touch with the minister. She sent in her card, which simply read: Effi von Innstetten née von Briest.'

b. *Effi und die Geheimrätin Zwicker waren seit fast drei*
Effi and the Privy.Counsellor Zwicker were since almost three
Wochen in Ems und bewohnten daselbst das Erdgeschoß einer
weeks in Ems and occupied there the ground.floor a
reizenden kleinen Villa. In ihrem zwischen ihren zwei Wohnzimmern
charming little villa. In their between their two living.rooms
gelegenen gemeinschaftlichen Salon mit Blick auf den
situated common drawing.room with view on the

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

> *Garten stand ein Palisanderflügel, auf dem Effi dann und wann eine*
> garden stood a rosewood.piano on which Effi now and then a
> *Sonate, die Zwicker dann und wann einen Walzer spielte;* (Effi Briest)
> sonata the Zwicker now and then a waltz played
> 'Effi and Privy Counsellor Zwicker had been in Ems for almost three weeks and lived there on the ground floor of a charming little villa. In their communal drawing room, situated between their two living rooms and overlooking the garden, stood a rosewood grand piano on which Effi occasionally played a sonata and Zwicker sporadically played a waltz;'

Only eventive descriptions in a discourse advance narrative time by updating reference time[3], stative descriptions fail to do so (Hinrichs 1986, Partee 1984, Dry 1983). Instead of narrative progress, they signal a narrative halt. According to Smith (2003: 35), the distinction between foreground and background in a narrative is established on the basis of the discourse modes that occur: The NARRATIVE discourse mode represents foreground information, the DESCRIPTIVE mode background information in a narrative text.

Discourse segments in narrative passages are typically held together by a discourse relation called NARRATION, a coordinating discourse relation which is temporal in nature, linking two eventive descriptions (Smith 2003, Altshuler 2021). Both arguments of the discourse relation must refer to particular events, thus excluding generalizing sentences which express eventive patterns. BACKGROUND is another characteristic coordinating discourse relation in narratives, which, as illustrated above, attaches stative descriptions to previous discourse segments. A hierarchical discourse structure arises when a sequence of coordinating discourse relations is interspersed with subordinating discourse relations such as the relation ELABORATION.

Temporal relations between discourse segments in a narrative manifest themselves in the use of temporal anaphora such as tense morphemes, temporal adverbials or temporal conjunctions: Regardless of the situation type involved, the newly introduced situation is always linked to the reference time of the previous discourse, establishing a temporal anaphoric relation between temporal expressions (Hinrichs 1986, Partee 1984, Smith 2003). Only the way in which the situation is related to reference time depends on its situation type and gives rise to either narrative progression or a narrative halt. As regards the narratives in

[3]Using the distinction of event time, speech time and reference time, suggested by Reichenbach (1947).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(4), the use of temporal anaphora means that the individual past tense forms of the finite verbs establish an anaphoric relation between the respective situations, thus signaling continuity across discourse segments on the temporal level. The temporal adverbial *gleich am nächsten Vormittag* 'the very next morning', on the other hand, is a frame adverbial, indicating a temporal shift and functioning as a temporal anchor for the eventive descriptions to follow. Frame adverbials therefore signal a boundary between text fragments. The temporal adverbial *schon eine Viertelstunde vorher* 'a quarter of an hour earlier' in the example below, on the other hand, shows how temporal adverbials can establish anaphoric relationships to a temporal expression in the preceding discourse. The example hence indicates temporal continuity between two discourse segments.

(5) *Das Schiff, ein leichtes Segelschiff (die Dampfboote gingen nur*
the ship a light sailing.vessel (the steamboats went only
sommers), fuhr um zwölf. Schon eine Viertelstunde vorher
in.summer) left at twelve. Already a quarter.of.an.hour earlier
waren Effi und Innstetten an Bord; auch Roswitha und Annie.
were Effi and Instetten on board likewise Roswitha and Annie
'The ship, a light sailing vessel (the steamboats only went in summer) left at twelve. Effi and Innstetten were already on board a quarter of an hour earlier, as were Roswitha and Annie.'

Despite the prominent role of temporal relations and their rendering in narratives, anaphoric relations between entities are an equally important means of indicating coherence between discourse segments in a narrative passage (von Stutterheim & Carroll 2005, Virtanen 2011). This can be demonstrated by the passages under (4), where the personal pronoun *sie* 'she' in (4a) and the proper name *Effi* in (4b) are used to express topic continuity. In terms of discourse structure, the left edge of discourse segments is of particular importance here: In a V2 language like German, the domain of entities and the temporal domain compete for the preverbal position, as von Stutterheim & Carroll (2005) observe in a series of experimental studies involving L1 and L2 speakers of German. They state that native speakers of German prefer participant-orientation over temporal orientation in the narrative discourse mode. A pertinent corpus example, taken from Theodor Fontane's novel, which llustrates this linking strategy is given below:

(6) ***Innstetten**, unbefangen und heiter, schien sich seines*
Innstetten light-hearted and cheerful seemed REFL his
häuslichen Glücks zu freuen und beschäftigte sich viel mit dem
domestic happiness to enjoy and engaged REFL much with the

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Kinde. ***Roswitha*** *war erstaunt, den gnädigen Herrn so zärtlich*
child. Roswitha was astonished the honorable master so tender
und zugleich so aufgeräumt zu sehen. ***Auch Effi*** *sprach viel*
and at the same time so tidy to see Also Effi talked a lot
und lachte viel, es kam ihr aber nicht aus innerster Seele.
and laughed a lot it came her however not from innermost soul.
Sie *fühlte sich bedrückt und wußte nur nicht, wen sie dafür*
She felt REFL depressed and knew only not whom she for it
verantwortlich machen sollte, Innstetten oder sich selber. (Effi Briest)
responsible hold should Instetten or herself

'Innstetten, light-hearted and cheerful, seemed to be enjoying his domestic happiness and was very engaged with the child. Roswitha was astonished to see the honorable master so tender and at the same time so tidy. Effi also talked and laughed a lot, but it did not come from her innermost soul. She felt depressed and did not know who to blame for it, Innstetten or herself.'

According to von Stutterheim & Carroll (2005), languages such as present-day English use other cohesive devices to signal continuity across discourse segments in narrative fragments. This is because, as Los (2012) shows, differences between V2 languages such as German and Dutch on the one hand and English on the other have emerged over the course of language history: In Old English, with its at least partial V2 syntax, the preverbal position could be used for temporal adverbials of an anaphoric nature as well as for temporal frame adverbials, just as in modern Dutch or German. In contrast, modern English favors frame adverbials at the left edge of the sentence, if temporal information appears at all at the left edge of a sentence, cf. Virtanen's (2011) example in (7). The changes that affect linking strategies in English across the temporal domain and the domain of entities are captured under the concept of boundedness: English is seen as a language that has changed from a bounded to an unbounded language (Los 2012).

(7) **In 1483**, Leonardo offered his services to Lodovico Sforza, Duke of Milan, as painter, sculptor, military engineer and architect. He remained at the Milanese court until it fell to the French in 1499.

Very briefly summarized, narratives, as related to an entire text, are the result of telling a story, in which events and characters are central. How a narrative is structured appears to interact directly with the linguistic structures, as indicated by both the temporal dimension and the tracking of characters. How and to which degree (changes in) grammatical patterns correlate with or are affected

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

by (changes in) the narrative structure remains an open question, which this volume hopes to give more insight to.

3 Outline of contents

Understandably, the contributions in this volume can only highlight some aspects of the complex relationship between grammar and narration(s). Addressing among others questions of narrative progression, temporal structure, reference tracking and discourse functions, the contributions discuss phenomena such as temporal adverbials at the left periphery as well as later in the clause, left dislocation structures, fronting of the finite verb in dependent and independent clauses, linguistic means to express aspectual and tense information, and the distribution of nominalizations across text types.

Sonja Zeman considers the implications of narration on the use of grammatical patterns, focusing on the history of German and paying empirical attention to left dislocation and the pluperfect. In her contribution, she addresses all three dimensions of narration: Narration as text type, as discourse mode, and as discourse relation. In her analysis of left dislocation in Middle High German narrative texts, she presents a picture that shows that left dislocation is not simply an oral feature, used to track important characters, but is used differently in the various text types, modes, and discourse relations. The interaction of the latter dimensions highlights a different function of left dislocation in different environments. Focusing on the interaction of temporality and narration, Sonja Zeman presents a study of the pluperfect along the same three dimensions and shows that discourse mode was the main motivation for using present or past perfect in Middle High German. The grammaticalization of the pluperfect was triggered by the non-narrative mode, as reference time plays a less prominent role in its meaning within that mode. With the two studies, Sonja Zeman showcases the role narration can play in the use and development of grammatical patterns.

Barthe Bloom addresses the role of narration on the choice between two types of clause-combining constructions within one narrative text, translated to German and Dutch in the fifteenth and sixteenth centuries. She finds a correlation between the use of a temporal *da* 'then' or a general resumptive *so* 'so' and discourse mode and duration, i.e., narrative speed. German adverbial clauses with *so* are associated with scenes and isochrony, whereas those that occur with *da* are found in narrative summaries, continuing the story. Dutch shows a similar tendency, but weaker, due to a wider spread of a third clause-combining pattern

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

of integration. Barthe Bloom addresses the interconnection between temporality and grammatical expressions and finds an important role of narration in the selection of clause-combining constructions, which may have, because of the contrastive comparison to the later translation in Dutch, diachronic implications.

Jordan Chark discusses the rise and spread of two aspectual patterns in the history of Icelandic, the *búinn*-perfect and the *að vera*-progressive. She argues that this can be better understood if one considers the changing narrative style, which moves away from an iconic, chronological reporting of events to more overt grounding by means of temporal adverbial clauses. This is related to the shift from a bounded system to unbounded system, which she argues took place in Icelandic in the early modern period in a similar way as it did in English. This means that the way of narrating shifts from being organized around a sequence of events to the reporting of events related to a topic. By corpus data, Jordan Chark shows that the amount of backgrounded temporal clauses increased, while the use of local anchoring adverbial, in particular the indicator of topic-time *þá* 'then' decreased. For the two aspectual patterns, the reanalysis is proposed to have happened in the early stages of the boundedness shift, but the circumstances of reanalysis are independent of this. The subsequent spread is, differently, highly advanced by it, since the focalized usage of the progressive allows setting up a new temporal frame, i.e., a new topic time through which events are viewed, and the *búinn*-perfect is a frame-setter strategy.

Hannah Booth examines a number of light adverbs in Old Icelandic saga narratives that have so far been well studied in particular for Old and Middle English. Depending on their position in the sentence, different discourse functions are ascribed to these adverbs in English. In the context of a corpus study of Old Icelandic – the corpus comprises two collections of saga narratives covering a period from 1150 to 1450 – it will be examined whether temporal and local adverbs such as *þá* 'then', *þar* 'there' and *nú* 'now' have the same discourse functions as their counterparts in older stages of the English language. By focusing on their clause-internal position, Hannah Booth can show that the adverbs in question actually demarcate two discourse domains in the sentence: While the position to the left of the adverb is favored by various types of topics, the domain to the right of the adverb is reserved for the assignment of information focus. Considering the decrease of clause-internal adverbs over time, Hannah Booth suggests that this development is linked to broader changes affecting the Icelandic language system and establishing the left edge of the clause as a preferred topic position.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Fabian Fleissner discusses the spread of complex prepositions in the history of German and verbonominal constructions with *in* in light of a more general development towards a more nominal style. In his paper, we evidence the effect of text type on the distribution of grammatical structures, in particular with regard to complex preposition construction. As narrative text types tend to focus on the telling of events, at least in German and Dutch (Carroll et al. 2004), the expression of events as nominals is much more restricted in narratives than in other text types, in particular scientific texts. Furthermore, the backgrounding of the agent, which is a characteristic trait of the construction, is consistent with the central role of characters in narration. As such, the construction lends itself well for condensing information without losing clarity and is suitable for the new scientific text type. For verbonominals, the relation to text type is less clear and the construction is characterized by multiple sub-schemas with different functions, showing partial overlap with complex prepositions. Besides these discourse-pragmatic considerations, Fleissner furthermore evaluates the age of nouns that occur within the constructions, and despite fluctuations, the data indicate that the increase of the use of both patterns comes from adding nouns to either construction as they are adopted into the language.

Based on the observation that argument-realizing V2 clauses occur much more frequently in Early New High German (1350–1650) than in present-day German, **Malika Reetz** investigates the question of which pragmatic factors license the occurrence of this non-canonical form of dependent clauses in Early New High German. Her corpus study is based on narrative texts from the fifteenth and sixteenth centuries which are particularly well suited to address her research question: Categories such as givenness, newness, at-issueness are easy to determine in a self-contained narrative. With the tool of conditional inference trees, she can establish that V2-clauses are preferred over V-final clauses when they are relevant for the progression of the narrative. At the same time, V2 is also used to indicate that an assertion is mediated. Only the V-final cause pattern can be used when the proposition is discourse-old and does not contribute to the main story line.

Verb-initial declaratives are associated with lively storytelling across older Germanic languages, cf. Old English (Los 2000), Old High German (Hinterhölzl & Petrova 2010) and Old Icelandic (Booth & Beck 2021). In particular, they are supposed to signal the coordinating discourse relation of NARRATION holding between two subsequent events. In a comprehensive corpus study for Old English,

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Anna Cichosz shows that the use of verb-initial declaratives is in fact not limited to a unique discourse function. As expected, the function of event-linking is most frequently attested in the corpus and seems to be conventionalized to a certain extent in Old English, as it is attested in a number of texts across different stages of Old English. Verb-initial declaratives are nevertheless also used in narratives to provide additional descriptions of given entities and, hence, to establish subordinate discourse relations such as the relation ELABORATION between discourse segments. The distribution of these verb-initial declaratives across the corpus suggests that they are an early native structure in Old English. According to Anna Cichosz, a deeper understanding of discourse functions of the syntactic pattern under question requires at least partial manual analysis of large amounts of data.

Old English also features the asymmetry of verb placement between main and subordinate clauses known from present-day German and present-day Dutch, albeit not as pronounced as in the two modern West Germanic languages (Pintzuk 1999). It is therefore not all that surprising that Old English also exhibits relative clauses with the finite verb in finite as well as in second position. **Bettelou Los and Stefano Coretta** focus in their contribution on Old English *se*-relatives exhibiting V2 word order. Compared to relative clauses with other connectors such as *þe* or *se þe* with the finite verb in final position, V2-relatives are less integrated into their parent clause because they are always non-restrictive in Old English. In their corpus study, Bettelou Los and Stefano Coretta establish that the fronting of the finite verb in V2-relatives is triggered by the fact that they take a discourse referent that is new to the discourse as their antecedent, i.e. V2-relatives primarily have a presentational function. Essentially, it remains an open question how *se*-relatives with the finite verb in second position can be distinguished from main clauses that are introduced by the demonstrative *se*. As a trial, it is proposed to decide this on the basis of the distance between the relative clause and the antecedent.

Acknowledgements

First of all, we would like to thank all the colleagues who contributed to this volume for not only engaging with the topic of the workshop, i.e., discussing discourse structures in Germanic from a diachronic perspective, but also for subsequently writing up their presentations for the present volume. Secondly, this volume would not have come about in its present form if a number of other colleagues had not supported us with their external expertise. We would like to

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

thank these colleagues for the time they have invested. Last but not least, we would like to thank Luisa Boeckmann, who supported us in preparing the print template for the volume.

References

Altshuler, Daniel. 2021. A puzzle about narrative progression and causal reasoning. In Emar Maier & Andreas Stokke (eds.), *The language of fiction*, 255–276. Oxford: Oxford University Press. DOI: 10.1093/oso/9780198846376.003.0010.

Ariel, Mira. 2004. Accessibility marking: Discourse functions, discourse profiles, and processing cues. *Discourse Processes* 37(2). 91–116. DOI: 10.1207/s15326950dp3702_2.

Asher, Nicolas & Alex Lascarides. 2003. *Logics of conversation*. New York: Cambridge University Press.

Baker, Peter S. 2012. *Introduction to Old English*. 3rd edn. Malden: Wiley-Blackwell.

Bal, Mieke. 2009[1985]. *Narratology: Introduction to the theory of narrative*. 3rd edn. Toronto: University of Toronto Press.

Bech, Kristin & Christine Meklenborg Salvesen. 2014. Preverbal word order in Old English and Old French. In Kristin Bech & Kristine Gunn Eide (eds.), *Information Structure and Language Change in Germanic and Romance Languages*, 233–269. Amsterdam: Benjamins.

Bloom, Barthe, Malika Reetz & Ulrike Demske. 2024. Romankorpus Frühneuhochdeutsch (Roko.UP): A collection of Early Modern German narrative prose. *Journal of Open Humanities Data* 10(17). 1–7. DOI: 10.5334/johd.188.

Booth, Hannah & Christin Beck. 2021. Verb-second and verb-first in the history of Icelandic. *Journal of Historical Syntax* 5(28). 1–53. DOI: 10.18148/hs/2021.v5i28.112.

Carroll, Mary & Christiane von Stutterheim. 2002. Typology and information organisation: Perspective taking and language-specific effects in the construal of events. In Anna Giacalone Ramat (ed.), *Typology and second language acquisition*, 365–402. Berlin: De Gruyter.

Carroll, Mary, Christiane von Stutterheim & Ralf Nuese. 2004. The language and thought debate: A psycholinguistic approach. In Thomas Pechmann & Christopher Habel (eds.), *Multidisciplinary approaches to language production*, vol. 157 (Trends in Linguistics: Studies and Monographs), 183–218. Berlin: De Gruyter. DOI: 10.1515/9783110894028.183.

Chatman, Seymore. 1978. *Story and discourse*. Ithaca, NY: Cornell University Press.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Dancygier, Barbara & Eve Sweetser (eds.). 2012. *Viewpoint in language: A multimodal perspective*. Cambridge: Cambridge University Press.

Declerck, Renaat. 2003. Distancing vs. backshifting: Modal indicative forms in English. *Leuvense Bijdragen: Tijdschrift voor Germaanse Filologie* 92(3–4). 13–22. DOI: 10.2143/LB.92.3.2022871.

Dry, Helen Aristar. 1983. The movement of narrative time. *Journal of Literary Semantics* 12(2). 19–53. DOI: 10.1515/jlse.1983.12.2.19.

Fleischmann, Suzanne. 1988. Discourse functions of tense-aspect oppositions in narrative: Toward a theory of grounding. *Linguistics* 23(6). 851–882. DOI: 10.1515/ling.1985.23.6.851.

Fludernik, Monika. 1993. *The fictions of language and the languages of fiction. The linguistic representation of speech and consciousness*. London: Routledge.

Fludernik, Monika. 2000. Genres, text types, or discourse modes? Narrative modalities and generic categorization. *Style* 34(2). 274–292.

Genette, Gérard. 1983[1980]. *Narrative discourse revisited*. Ithaca: Cornell University Press.

Genette, Gérard. 1980[1972]. *Narrative discourse: An essay in method*. Ithaca: Cornell University Press.

Herring, Susan C., Pieter van Reenen & Lene Schøsler. 2001. On textual parameters in older languages – theoretical overview. In Susan C. Herring, Pieter van Reenen & Lene Schøsler (eds.), *Textual parameters in older languages*, 1–31. Amsterdam: Benjamins.

Hinrichs, Erhard. 1986. Temporal anaphora in discourses of English. *Linguistics and Philosophy* 9(1). 63–82. DOI: 10.1007/BF00627435.

Hinterhölzl, Roland & Svetlana Petrova. 2010. From V1 to V2 in West Germanic. *Lingua* 120(2). 315–328. DOI: 10.1016/j.lingua.2008.10.007.

Hoffman, Yonina & Alice Gaber. 2023. First-person narration in Ancient Greek and Modern English literature. In Peter Hühn, John Pier & Wolf Schmid (eds.), *Handbook of diachronic narratology*, 16–40. Berlin: De Gruyter. DOI: 10.1515/9783110617481.

Hume, Kathryn. 2005. Narrative speed in contemporary fiction. *Narrative* 13(2). 105–124.

de Jong, Irene. 2004. *Narrators and focalizers: The presentation of the story in the Iliad*. 2nd edn. Bristol: Bristol Classical Press.

van Kemenade, Ans & Tanja Milicev. 2012. Syntax and discourse in Old and Middle English Word Order. In Dianne Jonas, John Whitman & Andrew Garrett (eds.), *Grammatical Change: Origins, Nature, Outcomes*, 239–255. Oxford: Oxford University Press.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Koch, Peter & Wulf Oesterreicher. 1985. Sprache der Nähe—Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. *Romanistisches Jahrbuch* 36(1985). 15–43. DOI: 10.15496/publikation-20410.

Korecká, Lucie. 2019. Narrative modes, narrative space, and narrative play in the Post-Classical sagas and Þættir of Icelanders. *Acta Universitatis Carolinae Philologica* 3. 9–26. DOI: 10.14712/24646830.2019.26.

Kukkonen, Karin. 2020. The speed of plot: Narrative acceleration and deceleration. *Orbis Litterarum* 75(2). 73–85. DOI: 10.1111/oli.12251.

Labov, William. 1972. *Language in the inner city*. Philadelphia: University of Pennsylvania Press.

Labov, William. 2013. *The language of life and death: The transformation of experience in oral narrative*. Cambridge: Cambridge University Press. DOI: 10.1017/cbo9781139519632.

Lönnroth, Lars. 1976. *Njáls saga: A critical introduction*. Berkeley: University of Californa Press.

Los, Bettelou. 2000. *Onginnan/beginnan* with bare and *to*-infinitive in Ælfric. In Olga Fischer, Anette Rosenbach & Dieter Stein (eds.), *Pathways of change: Grammaticalization in English*, 251–274. Amsterdam: Benjamins.

Los, Bettelou. 2012. The loss of Verb Second and the switch from bounded to unbounded systems. In Anneli Meurman-Solin, María José López-Couso & Bettelou Los (eds.), *Information structure and syntactic change in the history of English*, 21–46. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199860210.003.0002.

Partee, Barbara H. 1984. Nominal and temporal anaphora. *Linguistics and Philosophy* 7(3). 243–286. DOI: 10.1007/bf00627707.

Pintzuk, Susan. 1999. *Phrase structures in competition: Variation and change in Old English word order*. New York: Garland Publishing.

Prince, Gerald. 1987. *Dictionary of narratology*. Lincoln: University of Nebraska Press.

Redeker, Gisela. 1987. Introductions of story characters in interactive and non-interactive narration. In Jef Verschueren & Marcella Bertuccelli Papi (eds.), *The pragmatic perspective*, chap. 17, 339–355. Amsterdam: Benjamins. DOI: 10.1075/pbcs.5.26red.

Reichenbach, Hans. 1947. *Elements of symbolic logic*. New York: The Free Press. DOI: 10.2307/2268978.

Rimmon-Kenan, Schlomith. 1983. *Narrative fiction: Contemporary poetics*. London: Routledge.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Ryan, Marie-Laure. 2006. Semantics, pragmatics, and narrativity: A response to David Rudrum. *Narrative* 14(2). 188–196.

Ryan, Marie-Laure. 2017. Narrative. In Imre Szeman, Sarah Blacker & Justin Sully (eds.), *A companion to critical and cultural theory*, 517–530. Hoboken: Wiley Blackwell.

Sanford, Anthony, Kanwalraj Moar & Simon Garrod. 1988. Proper names as controllers of discourse focus. *Language and Speech* 31(1). 43–56. DOI: 10.1177/002383098803100102.

Schiffrin, Deborah. 1981. Tense variation in narrative. *Language* 57(1). 45–62. DOI: 10.1353/lan.1981.0011.

Smith, Carlota S. 2003. *Modes of discourse: The local structure of texts*. Cambridge: Cambridge University Press.

Somers, Katerina. 2024. *How to create an early German scriptus: The literization approach to historical German syntax* (Open Germanic Linguistics 9). Berlin: Language Science Press. DOI: 10.5281/zenodo.13836979.

Sternberg, Meir. 1990. Telling in time (I): Chronology and narrative theory. *Poetics Today* 11(4). 901–948. DOI: 10.2307/1773082.

von Stutterheim, Christiane & Mary Carroll. 2005. Subjektwahl und Topikkontinuität im Deutschen und Englischen. *Zeitschrift für Literaturwissenschaft und Linguistik* 35(139). 7–27. DOI: 10.1007/BF03379441.

Traugott, Elizabeth. 2007. Old English left-dislocations. *Folia Linguistica* 41. 405–441. DOI: 10.1515/flin.41.3-4.405.

Virtanen, Tuija. 2011. Variation across texts and discourses: Theoretical and methodological perspectives on text type and genre. In Heidrun Dorgeloh & Anja Wanner (eds.), *Syntactic variation and genre*, 53–84. Berlin: De Gruyter. DOI: 10.1515/9783110226485.1.53.

Wårvik, Brita. 2013. Participant continuity and narrative structure: Defining discourse marker functions in Old English. *Folia Linguistica Historica* 34(1). 209–242. DOI: 10.1515/flih.2013.008.

Zeman, Sonja. 2018. What is a narration – and why does it matter? In Markus Steinbach & Annika Hübl (eds.), *Linguistic foundations of narration in spoken and sign language* (Linguistics Today 247), 173–206. Amsterdam: Benjamins. DOI: 10.1075/la.247.08zem.

Zeman, Sonja. 2023. Temporality in ancient and medieval literatures. In Peter Hühn, John Pier & Wolf Schmid (eds.), *Handbook of diachronic narratology*, 371–395. Berlin: De Gruyter. DOI: 10.1515/9783110617481-020.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Chapter 2

The impact of narration – A diachronic view on German

Sonja Zeman[a]
[a]Universität Augsburg

With respect to the question of how grammatical patterns correlate with changes in discourse structure, the paper investigates the influence of narration on the usage of grammatical means in older stages of German and their diachronic development. The first part of the paper provides an overview on previous literature on the different dimensions of narration and their interaction with different grammatical means. Against this background, two empirical case studies show that these different dimensions of narration interact differently with the distribution and development of grammatical elements. Narration has thus a relevant impact and is an important aspect to consider when describing the diachronic relation between grammar and discourse structure.

1 Narration and discourse structure

It is general consensus that the use of grammatical forms is strongly influenced by discourse structure. It is therefore undisputed that grammatical elements should not be investigated in isolation but rather in their respective communicative contexts. This naturally also applies to studies of grammar in historical stages of language (see Ziegler 2010 for a plea for a historical discourse grammar that goes beyond the sentence level). In the historical linguistics of German, it has been shown by numerous studies that the use of grammatical means depends on various contextual factors such as text types (e.g. Gloning 2010), the oral vs. written distinction (Elspaß 2005, Ágel & Hennig 2006), the prose vs. verse distinction (e.g. Prell 2005; Speyer 2013), information structure (e.g. Hinterhölzl et al. 2005; Speyer 2013; Demske 2018), and regional variation (e.g. Elspaß 2010; Werth 2020).

Sonja Zeman. 2025. The impact of narration – A diachronic view on German. In Ulrike Demske & Barthe Bloom (eds.), *Discourse structure and narration: A diachronic view from Germanic*, 15–36. Berlin: Language Science Press. DOI: ??

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

In addition, the relationship between grammatical means and discourse structure has been studied across languages, particularly in historical register analysis (see for an overview Kytö 2019 and Goulart et al. 2020).

Recent investigations have emphasised that text and discourse parameters are also key factors in language change (see e.g. Hennig 2009 for German; Biber & Gray 2013, Biber et al. 2021, for English). Consequently, it has been claimed that registers are a mediating factor and "act as powerful vehicles promoting or retarding language change, or even contributing to stability" (Kytö 2019: 155).

One specific aspect of discourse that needs to be considered in terms of its influence on grammatical devices and their diachronic development is narration. There are two reasons for this. First, many of the texts that have survived from the earlier stages are narrative texts. Second, many texts contain narrative passages. Narration is therefore an important parameter in many studies and corpus annotations. In Biber's multimodal discourse analysis, for example, the narrative vs. non-narrative distinction is one of the listed parameters (e.g. Biber & Conrad 2019). "Narrative" is (next to expository, argumentative, and instructive) also one of the categories for defining the prototypical text classes in the Helsinki Corpus of English texts (Kytö 2019: 138). In addition, various studies have shown that the distribution of grammatical features varies within the intra-textual registers of dialogue and narration (see, for example, Padučeva 2011; Mazziotta & Glikman 2019, Egbert & Mahlberg 2020, and Larrivée 2022: 2 for an overview with further references), and it has been discussed whether dialogue or narration are more often the initiating context for innovation in language change (see Brown & Tagliamonte 2012; Mazziotta & Glikman 2019).

Although the influence of narration has been proven across languages and for quite specific aspects of narration (e.g. on intensifiers in storytelling in Brown & Tagliamonte 2012), we still do not know exactly which aspects are responsible for the different distribution of grammatical elements in narrative vs. dialogue, and the influence of narration on language change has remained vague (see also Larrivée 2022: 2).

This applies particularly to German linguistics, in which narration itself has not often been a topic in its own right. Rather, the focus of diachronic discourse studies has been on the development from orality to writing and the influence of different text types and communicative situations (see, for an example, the various papers in Ziegler (2010)'s volume on historical text grammar in German).

The aim of this article is to take a further step towards a systematic investigation of the influence of narration in the history of German and in language change from both a theoretical and empirical point of view by (i) clarifying the

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

theoretical concept of narration and its interaction with discourse structure, orality, and grammar, and (ii) presenting two exemplary case studies on the development of the pluperfect and left dislocations.

In this respect, §2 provides the theoretical background focusing on the interaction between narration, discourse structure, and grammar. This is based on the assumption that three dimensions of narration have to be distinguished, namely (i) narrative *texts* that present a story by following a narrative macrostructure (in the sense of Labov & Waletzky 1967) and prototypically integrate the viewpoints of different characters and narrators (Zeman 2020), (ii) narrative *discourse mode*, i.e. narrative segments that can occur in (narrative and non-narrative) texts and are characterised by a specific pattern of grammatical use (Smith 2003), and narrative *discourse relation* that holds between two utterances in discourse (in the sense of Asher & Lascarides 2003). Against this background, two empirical case studies will show that these three dimensions of narration have an impact on the distribution and grammaticalisation of linguistic elements (§3). The case studies address left dislocations in Middle High German (§3.1) and the grammaticalisation of the pluperfect (§3.2). In sum (§4), the paper argues that the different dimensions of narration interact differently with the distribution and development of grammatical elements. Narration is therefore an important aspect to consider when describing grammatical elements in their context in the discourse structure and should be recognised as an essential factor in language change.

2 Dimensions of narration

2.1 Macrostructure of narration

To address the question of which properties of narration have an influence on the distribution and development of grammatical elements, it is first necessary to specify what a narration is and how it can affect the use of grammatical devices. In this respect, narration has often been seen as a property of narrative *texts*. Fairytales such as "Little Red Riding Hood" or novels are seen as prototypical examples of narrative texts, as they tell a story. Although there is no generally accepted definition of narration in either narratology or linguistics (see Zeman 2018 for an overview), there is broad agreement in linguistics that narrative texts are characterised by sequences of specific events in temporal and causal order and follow an underlying macrostructure in the sense of Labov & Waletzky (1967), which contains a setting, a complication and a resolution as main components. According to this definition, many text sources from older stages of language are prototypical narrative texts, such as Middle High German epic poems (e.g.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Nibelungenlied, *Tristan*, *Parzival*) and Early New High German novels (e.g. *Fortunatus*, *Melusine*).

This is an important fact as the specific discourse structure of narration can influence both the distribution and the development of grammatical elements in a relevant way. This can be seen, for example, in the discourse organisation of a narrative text through scene shifts, as shown for Old English. Fludernik (2003) has shown that in the history of English there has been a development of "chapterification" and an "attendant emphasis on the chapter beginning as a salient point for temporal or locative shifts" (Fludernik 2003: 337), as it is common in the nineteenth-century novel. In contrast, episodic narratives in Old English were organised by means of macroepisodic markers, i.e. scene settings were merely indicated by the change of protagonists or the next stage of the hero's adventure (Fludernik 2003: 337). This development had a considerable influence on the use of metanarrative and metafictional comments and formulae.

For the history of German, systematic investigations on scene shift from a linguistic perspective are missing. However, it can be assumed that there has been a similar development, as the increase in explicit and conventionalised scene shift markers (e.g. "Now let us leave X and Y (in location A) and turn to O and P"; Fludernik 2003: 334) seems to correlate with the emergence of stories with multiple subplots such as the *Prosalancelot* from Late Middle High German onwards (see Zeman 2023b: 699f.). In contrast, the discourse organisation of earlier Middle High German verse epics is characterised by the frequent use of the particle *dô* ('then'), which indicates the next step in the sequence of events. In this regard, the organisation of the narrative structure is not very explicit or elaborate but comparable to patterns that are characteristic of everyday oral storytelling (e.g. "and then she said, and then I said [...]").

In addition to such discourse-organizing patterns that are specific to the narrative macrostructure, another important difference between narrative and non-narrative *texts* is that they are not objective representations of a sequence of events but present a story from a particular point of view. The narrator's perspective is therefore a crucial aspect that distinguishes narrative from non-narrative texts with corresponding consequences for the discourse structure (see Zeman 2018). Since utterances in narrative discourse can contain both the narrator's and the protagonists' perspectives, it allows for complex viewpoint constructions such as Free Indirect Discourse (*Tomorrow was Monday.*) and the Future-of-Fate construction (*He was never to see her again.*), which for this reason are restricted to narrative discourse only (see in detail Zeman 2018).

Both Free Indirect Discourse and Future-of-Fate are specific constructions of narrative discourse and do not occur in all stages of language. There are, how-

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

ever, other more common grammatical elements that interact with the narrator's voice. One example is interjections. Middle High German verse epics are characterised by the fact that interjections are not only used in the characters' dialogues but are also frequently linked to the narrator who is "an active participant in the act of narration" (Fludernik 2003: 338). In this regard, the epics reflect the fact that storytellers in medieval times recited the stories in front of an audience. In this shared communication situation between narrator and listeners, the narrator addresses the audience and comments on the story. The story told in the distant world is often presented as if the audience could look at it and see the protagonists before their eyes. Consequently, medieval storytelling is characterised by a combination of deictic markers of distance and proximity. In addition to the use of past tenses that mark the distance of the story world, the narrator's discourse is often characterised by deictic adverbials that indicate the "here" and "now" of the story, verbs of perception and expressive interjections that "intrude" into the plot level (e.g. 'alas how many chainmails the bold Dancwart broke there!'; *Nibelungenlied* B, 212.4).

Such a conflation of proximity and distance is seen as a general feature of oral storytelling. With the transition from metrically bound verse epics in Middle High German to the increase of written stories and the development of prose novels with multiple narrative strands (such as the *Prosalancelot*), the role of the narrator and the deictic conceptualisation of place and time have changed (see for an overview Zeman 2023b). The changes in narrative macrostructure thus correlate with a different use of grammatical devices such as of interjections, personal pronouns, exclamative sentences etc.

In conclusion, narrative macrostructure is an important factor to consider when investigating the interaction between discourse structure and grammatical means in the history of German. However, it is not the only dimension of narration that has an impact on the distribution and development of grammatical means. In addition to the macrostructure of narrative *texts*, it is also important to distinguish between two narrative dimensions *within* narrative texts, i.e. narrative *discourse modes* and narrative *discourse relations*.

2.2 Microstructure of narration: Discourse mode

As seen above, narration can refer to the macrostructure of discourse and, as such, to whole texts. A text telling the story of Little Red Riding Hood would thus be considered as "narrative". In the case of a section of text such as (1), the whole section would be narrative because it belongs to a narrative text. In the tradition of distinguishing between discourse modes, however, narrative does not refer to

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

entire texts, but to *segments* within a – narrative or non-narrative – text. In (1), only those utterances would count as narrative that refer to the events of the plot (marked in bold) as opposed to the dialogic passages that represent the speech and thought of the protagonists.

(1) ***But the wolf ran straight to the grandmother's house and knocked on the door.*** *"Who's there?" – "Little Red Riding Hood. I'm bringing you some cake and wine. Open the door for me." – "Just press the latch,"* ***called out the grandmother.*** *"I'm too weak to get up."* ***The wolf pressed the latch, and the door opened. He stepped inside, went straight to the grandmother's bed, and ate her up.***
(https://www.grimmstories.com/language.php?grimm=026&l=en&r=de; access 01-11-2023)

Regarding the interaction between grammar and discourse, the distinction between narrative and dialogic segments is crucial, as it is associated with different behaviour of grammatical devices such as pronouns, aspect, and tense (see e.g. Mazziotta & Glikman 2019, Egbert & Mahlberg 2020, and Larrivée 2022: 2 with further references). The grammatical differences are particularly evident in the textual distribution of tenses. Benveniste (1974[1972]: 264), for example, has shown that the French tenses are distributed differently in the two text modes "histoire" vs."discourse". In the "histoire" only the aorist, imperfect, pluperfect, and prospective tenses are used, whereas the prototypical tenses of "discours" – i.e. present, present perfect, and future – are excluded. Benveniste sees this distribution as an argument for the fact that the two levels of enunciation "discours" and "histoire" are "two different and complementary systems" (Benveniste 1974[1972]: 266). Similarly, Weinrich (1964) and Hamburger (1968[1957]) assume a complementary distribution of tenses in narrative and non-narrative contexts.

To date, the close interdependence between verbal categories and narrative context has been confirmed in various studies (e.g. Dahl 1985; Fleischman 1990; Smith 2003; Zeman 2010; Padučeva 2011; Fischer 2020), albeit sometimes using different terminology (cf. for an overview Zeman 2020: 461–466). As a result, narrative discourse is characterised by "a special subset of rules, conditions and restrictions" (van Dijk 2015: 289) and by "its own special features" (Padučeva 2011: 9). According to Smith (2003), the narrative discourse mode has the status of a "covert" category, i.e. a category characterised by a specific pattern in terms of the distribution of grammatical devices. Its basic structure is characterised by the fact that narrative events are anaphorically related to each other and are not directly related to speech time but oriented to a displaced reference point (Smith 2003: 13).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Regarding the discourse structure in the historical stages of German, the differences between narrative and non-narrative discourse modes are also evident in the tense distribution. In Old and Middle High German, the present perfect cannot denote events in the narrative past (Zeman 2010) and is therefore primarily used in dialogic passages. This preference is also reflected in the grammaticalisation process (see Fischer 2020). As can be seen for other examples of perfect expansion in the Romance languages (Drinka 2017: 9), the present perfect increasingly occurs in non-narrative texts, while the main domain of the past tense remains the narrative discourse mode.

In addition to tense, other grammatical devices are also sensitive to the distinction between narrative vs. non-narrative. This can be seen in a comparison between text samples from the *Nibelungenlied* and *Tristan*. Table 1 shows the results of a study comparing two different historical texts based on the method of Ágel & Hennig (2006). On the basis of a sample of 2000 words, different linguistic features are annotated, which are considered to be markers of the language of proximity and parameters of syntactic complexity (sentence length, relationship between main clauses and subordinate clauses). In addition, a distinction is made between narrative and dialogical passages. All statements within direct speech (but without any embedded narrative passages) as well as narrator comments in the text, which are oriented towards the time of speaking, were considered "dialogical". All passages that represent the story of the narrated world were considered narrative.

Table 1 shows a significant difference between narrative and dialogical passages. This is particularly true for address-oriented features such as interjections, imperatives, vocatives etc. and deictic elements which occur significantly more often in dialogical passages than in narrative passages. Due to the sensitivity of several features to the narrative vs. non-narrative distinction, discourse modes must be taken into account with regard to the distribution and development of grammatical features in the history of German.

2.3 Microstructure: Discourse relations

In addition to narrative macrostructure and narrative discourse mode, there is a third dimension of narration, namely the dimension of discourse *relations*. discourse relations are semantic connections that link two utterances within a text (cf. Hobbs 1979; Asher & Lascarides 2003). For example, a sentence may be an explanation or an elaboration of the previous utterance. Since a coherent discourse is characterised by the fact that each utterance must be semantically and structurally connected in some way to the preceding and the following utterances, it

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Table 1: Comparison of selected features between *Nibelungenlied* B and *Tristan* (Zeman 2022: 188)

Parameter (Ágel/Hennig 2006)	Nibelungenlied B narrative	Nibelungenlied B dialogical	Tristan narrative	Tristan dialogical
MICROSTRUCTURE				
code (interjections, emotive	4	4	2	58
role (imperative, vocative)	0	52	0	55
situation (personal deixis;	5	268	8	264
temporal & local deixis)	4	24	9	24
time (left-placement)	17	4	11	15
MACROSTRUCTURE				
matrix sentences	180	200	140	163
subordinate sentences	65	97	117	125
length on average (in words)	11.1	10	14.3	12.3
(Microstructural) features in total	30	352	30	416
words in total	2000	2000	2003	2004
	1.5%	17.6%	1.5%	20.8%

can be assumed that the entire discourse is structured by underlying discourse relations that either remain implicit or are explicitly marked by connectors such as conjunctions, conjunctional adverbs etc.

According to Asher & Lascarides (2003), a narrative discourse relation is characterised by the fact that two denoted events are linked by temporal sequence and correspond to their natural order. Narrative connections can be overtly marked by instances of *and then* (Asher & Lascarides 2003: 200). Narrative texts and discourse modes often contain narrative connections, as indicated by [then] in example (2). However, not every connection between two propositions in a narrative text or in narrative discourse mode is necessarily a narrative connection. The first sentence in (2), for example, stands in a contrastive relation to the preceding text segment, as indicated by the connector *but*.

(2) ***But** the wolf **[in contrast to Little Red Riding Hood]** ran straight to the grandmother's house and **[then]** knocked on the door. "Who's there?" -*

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

> *"Little Red Riding Hood. I'm bringing you some cake and wine. Open the door for me." - "Just press the latch," called* ***[then]*** *out the grandmother. "I'm too weak to get up." The wolf pressed the latch* ***[then]****, and* ***[then]*** *the door opened. He* ***[then]*** *stepped inside,* ***[then]*** *went straight to the grandmother's bed, and* ***[then]*** *ate her up.*
> (https://www.grimmstories.com/language.php?grimm=026&l=en&r=de; access 05-11-2023)

In terms of the interplay between grammar and discourse structure, discourse relations have also been relevant in the history of German. A recent study by Speyer (2022), for example, investigates the dependency between syntactic serialisation and the semantic connections between propositions. It shows that the ranking of informational constituents in the prefield differs depending on the discourse relation (Speyer 2022: 225). Furthermore, it is shown that this ranking is affected by language change. In Early New High German, the preferences of the information structure in the prefield differ from those in Modern High German (Speyer 2022: 227). Studies of discourse relations are therefore promising for shedding more light on the diachronic development of connective markers and word serialisation.

3 Case studies

Against the background of section 2, it becomes evident that all three different dimensions of narration – (i) narrative *texts* that exhibit a narrative macrostructure, (ii) narrative *discourse mode*, i.e. narrative segments that can occur in (narrative and non-narrative) texts, and (iii) narrative *discourse relations* between propositions – can have effects on linguistic changes. The distinction is therefore crucial to systematically trace the development of grammatical means in relation to discourse structure. This is illustrated in the following by two case studies that will take a closer look at the effects of narration on the distribution and development of left dislocations and the pluperfect in the history of German.

3.1 Left dislocations (LD)

"Left dislocations" (LD), such as e.g. *My aunt, she used to play the guitar*, are traditionally described as constructions in which a syntactic constituent is "displaced" to the front (i.e. "to the left") of a sentence and is subsequently taken up by an anaphoric pronoun. As such, the constituent is seen as located "outside" the actual sentence. Consequently, LDs have a highlighting effect because they divide

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

the sentence into a two-part structure (cf. Bally 1932), with the first part containing the referencing element, the second part containing the statement about the respective referenced object.

The two-part structures of LDs are considered to be typical features of spoken language cross-linguistically (Chafe 1994: 67). It has been claimed that the division of a sentence into reference and statement provides a cognitive processing advantage, because it breaks the sentence down into smaller information or processing units, thus facilitating production and comprehension in spoken language (cf. Lötscher 1994: 39; Fiehler 2009: § 2015). They are therefore also seen as evidence of a more "oral syntax" in historical stages of language, i.e. a syntax that is structured in information units and characterised by a stronger "aggregative" (as opposed to "integrative") organisation (e.g. Lötscher 1994: 48; Ágel & Hennig 2006). One might therefore assume that LDs are mainly used in dialogic passages. However, this is not the case. As Table 1 above shows, the LDs in the *Nibelungenlied* show a preference for narrative passages. This leads to the questions of whether syntactic structures depend on the difference between discourse modes and whether LDs display different functions in narrative vs. non-narrative passages – and how the interaction between syntactic structure and discourse mode affects the development of LDs in German.

3.1.1 LD in the history of German

To pursue these questions, it is first necessary to take a closer look at LDs in the history of German. With regard to their general diachronic development, Lötscher (1994: 42) has shown that LDs were particularly common in the 13th and 14th centuries, but that their frequency decreased significantly in the period from 1200 to 1600. This development is usually seen in parallel with a process of literacy. It is assumed that as literacy increases, the characteristics of spoken language in written texts diminish and the linguistic norm becomes more determined by the laws of writing. This assumption is supported by the fact that LDs occur particularly frequently in texts that are close to oral language (Lötscher 1994: 48).

With regard to the use of LDs in Middle High German, two distinctions have to be made. First, it should be noted that LDs are not a homogeneous group but can occur in different constellations, cf. (3).

(3) a. *Den Schatz, den kennt jetzt niemand.*
'The treasure that knows now nobody'

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

b. *Der Schatz – den / ihn kennt niemand*
niemand kennt ihn.
'The treasure – that / it knows nobody'
nobody knows it

(3a) is characterised by the fact that the resumptive pronoun is a *d*-pronoun and that pronoun and NP are both coreferential and case-congruent. In (3b), on the other hand, there is no case congruence, the resumptive pronoun does not necessarily occur at the beginning of the sentence, and the NP can be resumed by either a d- or p-pronoun. In addition, there is often a pause in speech between the preceding NP and the referential statement. In cases like (3b), the first constituent forms a syntactically more independent unit and is less integrated into the sentence than in (3a).

These two syntactic variants are also associated with different discourse-pragmatic functions (cf. Cinque 1977; Selting 1993; Averintseva 2007; Dewald 2012). Although both (3a) and (3b) are used to mark topical referents, they differ in the way they mark them. While LDs such as (3a) are locally linked to the immediately preceding context, LDs such as (3b) are assumed to be a global resumption strategy by introducing a topic for the entire following discourse section (Selting 1993: 309).

There are different terms for these two variants used in the literature. In the German-language literature, a distinction is usually made between "Linksversetzung" ('left dislocation') in the narrow sense for examples such as (3a) and "Freies Thema" ('free topic') for (3b) (Altmann 1981). The latter have also been called 'hanging topics'. Frey (2004), for example, distinguishes between "German left dislocation" and "hanging topic left dislocation" (Frey 2004). Older publications have used terms such as "nominative pendens" (Behaghel 1932; Zäch 1931; Paul 2007) and "isolated nominative" (Paul 2007: § S 56).

For historical stages of language, it is difficult to distinguish clearly between the two variants of LDs, as one of the key criteria for evaluating the degree of integration is the phonetic level. In the works of Selting (1993) and Fiehler (2009: § 2015), intonation and prosody are identified as the most reliable criteria for distinguishing whether the referential expression and the subsequent statement are integrated in a common structure or represent two independent phrases. With regard to the written Middle High German data, a distinction can only be made if the syntactic features such as the absence of case congruence, distance position, type of pronoun, etc., deviate from the common pattern set out in (3a). Consequently, all occurrences of LDs have been analysed with respect to these parameters within the case study. This procedure has shown that only a few instances deviate from pattern (3a). This result corroborates Lötscher's (1994) observation

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

that 'free topics' do not play a significant role in Middle High German literature. Consequently, the subsequent analysis will focus on LDs in the narrow sense (i.e. pattern 3a).

A second distinction that must be considered is the difference between "simple" and "complex" LDs, as illustrated in examples (4a) and (4b). In comparison to example (4a), where a simple NP (*den hort*) is resumed by a d-pronoun right after the referential object, example (4b) is more complex since the referential object in the prefield is further attributed by relative clauses.

(4) a. "simple" LDs
den hort den weiz nu niemen wan got unde mîn
the treasure that knows now nobody except god and myself
[*Nibelungenlied* C, 2431, 3-4]

b. "complex" LDs
Strît den aller hôhsten der inder dâ geschach,
fight the biggest that happened there to them
ze jungest und zem êrsten, den ieman gesach den tet vil
last and first, that somebody saw **that** did very
willeclîche diu Sîfride hant
intentionally the hand of Siegfried
[*Nibelungenlied* C, 237, 1–3]

This distinction is of particular importance in the context of Middle High German, given that the resumptive pronoun is obligatory for complex LDs in the 13th and 14th century. This implies that in the *Nibelungenlied*, complex NPs are always resumed by a pronoun and thus obligatorily display left dislocation. In contrast, the LDs of simple NPs as exemplified in (4), are never obligatory (Lötscher 1994: 44). The following analysis will therefore be restricted to the examination of 'simple' LDs. Only in these cases will it be possible to examine the contexts in which they occur and their functions in comparison to non-LD constructions.

3.1.2 LDs in narrative and non-narrative contexts

To clarify whether the preference for narrative passages exhibited by the LDs in the *Nibelungenlied* (as evidenced in Table 1) can be regarded as representative of the entire text or whether it is merely a random occurrence in the sample, all LD occurrences of simple noun phrases in the *Nibelungenlied* were analysed. To facilitate comparison, an equal number of LDs were analysed (verses 1 to 5469)

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

for the *Tristan* (i.e. 28% of the entire text of 19545 verses). As "dialogical" are considered those passages in direct speech (without embedded narrations) and the narrator's speech, which are both oriented to speech time. In contrast, "narrative" are text passages that represent the story in the narrated world.Since the distribution of LDs in the dialogue and narrative passages can only be compared if the proportion of dialogue passages in the epic is taken into account, the speech density (i.e. the proportion of direct speech in relation to the entire text) was also determined for both texts,[1] see Table 2.

Table 2: Comparison of LDs in *Nibelungenlied* C and *Tristan* (Zeman 2023a: 77)

	Nibelungenlied C				Tristan (verses 1-5469)			
	narr	dia	narrator	sum	narr	dia	narrator	sum
LD	136	33	1	170	120	37	13	170
	80%	19.4%	0.6%		70.6%	21.8%	7.6%	
total	60.2%	39.8%			71.9%	28.1%		

If LDs were evenly distributed in both discourse modes, one would expect that the percentage in the dialogic passages would correspond to the percentage of direct speech in the whole text. In the *Nibelungenlied*, however, the frequency of LDs is higher in the narrative passages than in the dialogic passages. These quantitative findings correspond to qualitative differences.

In the narrative discourse mode, LDs usually serve to actualise or re-actualise a protagonist as a prominent agent in the discourse. It is particularly often used to select a referent from an established set of protagonists, cf. (5).

(5) *Der künic von Tenemarke der sprach sâ zestunt.*
The king of Denmark, the spoke there suddenly
[*Nibelungenlied C*, 300, 1]

In the previous context of (5), the first meeting between the hero Siegfried and his future wife Kriemhild, princess of Burgundy, is told. The king of Denmark is not mentioned throughout the chapter (Aventiure 5). However, it is mentioned

[1]The speech density was counted automatically. Therefore, the number of narrator's comments and embedded narrations could not be tracked for the entire text. Since narrator comments and embedded narrations are very rare for the Nibelungenlied, this should not have an impact on the ratios.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

in the immediately preceding section (*Nibelungenlied C*, 299) that guests from all over the world attend the court's celebration of its victory over Denmark and the Saxons. Although the king of Denmark is not mentioned since verse 238,2, he is part of the group of guests and, as such, an (implicitly) available discourse referent.

This functional pattern is well known for the use of LDs in spoken language. However, it is also characteristic of oral storytelling. According to Chafe (1994: 67), oral narratives are characterised by the fact that the intonation unit often verbalises only a single isolated referent who is (re-)actualised or (re-)focused as the protagonist. The importance of this function in the *Nibelungenlied* is reflected in the high proportion of animated (70.5%) and definite subjects (93.3%) and the frequent use of proper names (61.2%). When unanimated objects are displaced, they are often objects of particular relevance to the plot, such as the stone that Brünhild throws in competition with Gunther (verses 473.2; 474.1; 475.2) and the treasure of the Nibelungs (89.1; 2431.3).

(6) *Hort der Nibelunges der was gar getragen*
treasure of the Nibelungs the was wholly carried
[*Nibelungenlied* C, 89, 1]

Furthermore, LDs often occur at the beginning of an episode. In the manuscript, these are often marked by coloured initials. In these contexts, LDs contribute to the organisation of the text. This function can also be observed in the *Tristan*, where the LDs are often combined with the discourse marker *nu*, which sets a new temporal reference point for the story, cf. (7), similar also *Tristan* 3379, 3576.

(7) *Nû Tristan der was aber bereit.*
now Tristan the was again willing
[*Tristan*, 2970]

In comparison with the *Nibelungenlied*, the scene shift is thus marked more explicitly. This observation would be in line with the diachronic development of "chapterification" (Fludernik 2003) and clearer marking of textual organisation, as described in §2.

LDs thus serve several functions in narratives, which can be described more systematically when considering the distinction between the three dimensions of narration. First, LDs contribute to the macrostructure of narration by (re-)actualizing and highlighting the story's protagonists. They also contribute to the organisation of narrative by marking episode shifts. As such, it is understandable why they are frequently used in narrative discourse mode. In terms of discourse

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

relations, however, it is interesting to note that the LDs do not only (and not primarily) occur in narrative discourse relations. While 60/140 LDs in narrative discourse mode are used in utterances that are linked to the previous discourse by a narrative discourse relation, 58/140 LDs are used in utterances that provide a description or background to the previous discourse segments, cf. (7) above and (8).

(8) *ir schilde wol getân,* ***die*** *lûhten von den handen den vil waetlîchen man.*
their beautiful shields, they shone from their hands for the very splendid men

[*Nibelungenlied* C, 408, 3f.]

(7) above sets a new reference point for the story now. However, it does not indicate the next step of action, but gives the background (*being ready*) for Tristan's subsequent actions, which are recounted from the immediately following sentence. In (8), there is no narration of an event. Rather, the story comes to a pause and the Nibelungs' weapons are described. This shows that the LDs are not restricted to a single function within a narrative but can be used in very different ways.

In the dialogic discourse mode, LDs only rarely (1/30) occur in the context of narrative discourse relations. Most often they are used to evaluate a situation or to indicate a result. An example is (5) above, repeated here as (9).

(9) *den hort den weiz nu niemen wan got unde mîn*
the treasure that knows now nobody except god and myself

Again, (9) does not refer to the next step of an action or a specific event. Hagen as one of the heroes of the story comments on the treasure of the Nibelungs as an object that is relevant within the whole story and emphasises the result of the preceding actions. Since all the others are dead, Hagen is the only one how who knows where he has sunk the treasure in the Rhine. The analysis of discourse relations thus shows that narrative and non-narrative discourse modes correlate with different functional patterns of the LDs.

This assumption is supported by a look at the use of LDs in sermons. Sermons are known for their high frequency of LDs occurrences (Lötscher 1994). So far, this has been explained by the fact that sermons are a text type that is relatively close to spoken language. Since sermons are also characterised by narrative passages, it could have been assumed that the LDs would also be found primarily in narrative functions. However, this is not the case in Berthold von Regensburg's sermon. LDs are mainly used in explanatory descriptions, cf. (10).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(10) *Swaz daz firmament begriffen hât – daz ist der himel, den wir dâ sehen, dâ die sternen ane stênt –, swaz der umbe sich begriffen hât, daz ist geschaffen als ein ei. Diu ûzer schale* ***daz*** *ist der himel den wir dâ sehen. Daz wîze al umbe den tottern* ***daz*** *sint die lüfte. Sô ist der totter enmitten drinne, daz ist diu erde.*

'What the firmament encompasses – that is the sky that we see, where the stars are –, what it encompasses, that is built like an egg. The outer shell, **that** is the sky that we see there. The white all around the yolk, **that** is the air. So is the yolk in the middle, that is the earth.'

[Berthold von Regensburg, sermo XXV: *Sælic sint die reines herzen sint*]

The passage is not narrative but instructional as it explains how things in the world are according to the knowledge of the time. In this context, there are several sentences beginning with *that* in resumptive function. The two LDs in the text (marked in bold) reflect the same pattern as the surrounding sentences. The sentence with *that* offers a description of the previous statement or referent.

Considering the narrative discourse mode and the discourse relations, it seems that the LDs in Berthold's sermon are mostly used in the non-narrative discourse mode.

In conclusion, LDs show a great variety both in terms of their frequency and their range of functions in different text types. Further research is needed in order to gain a more detailed insight into the discourse functions. This is also relevant with regard to the diachronic line of development. As the analysis has shown, the development of LDs from 1200 to 1600 is not linear. While there is a general decrease of LDs in the history of German, there are also micro-changes in dependence of discourse modes and discourse relations.[2] The combination of LDs with other markers of discourse organisation suggests that the development of the LD is connected to a more explicit strategy of marking discourse structure in narrative texts. Further research would be needed to test this hypothesis. It is clear from the observations in this section that the three dimensions of narration are useful for adequately describing the distribution and functions of LDs and their development.

3.2 The grammaticalisation of the pluperfect

It has already been mentioned in §2.2 that the interaction between narration and grammar is particularly evident in relation to the category of tense. Again,

[2]Nor is there a simple dependence to the degree of orality: LDs are more frequent in the courtly epic *Tristan* than in the *Nibelungenlied*, which is considered closer to oral poetry (see in detail Zeman 2023a).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

this interaction involves all three dimensions of narration. In terms of narrative macrostructure, it has been noted that narration "displays some specifically narrative uses of tense" (Fludernik 2012: 94), such as the Epic Preterite and the Historical Present. Both uses of tense – the use of the present tense to denote past events and the use of the preterite in combination with speaker-oriented adverbials such as *now* or *tomorrow* (e.g. *Tomorrow was Christmas day*) are restricted to narratives only, since they presuppose the distinction between the narrator's and the character's voice and hence the double-layered structure of narration (see in detail Zeman 2010, Zeman 2018).

The distribution of tenses has also been the main argument for distinguishing between different discourse modes. Benveniste (1974[1972]), Weinrich (1964) and Hamburger (1968[1957]) have argued that there are two different systems of tenses which are distributed complementarily within the text. Although their approaches are conceptually different (see Fludernik 2012 for a comparison), more recent literature has confirmed that the different modes of discourse show different patterns of tense use (e.g. Smith 2003).

This is also evident for the tense distribution in Middle High German (Zeman 2010). Table 3 shows that the present and the present perfect have a strong preference for the speaker-oriented discourse mode (SR), while the preterite and the pluperfect are mainly attested in narrative discourse mode (NR). As such, the tenses in Middle High German are distributed in a complementary way.

Table 3: Tense distribution in the verse epic *Herzog Ernst* / Middle High German (according to Zeman 2010: 117)

	narrative		speaker-oriented		general		sum
	n	%	n	%	n	% n	
PRES	2	0.24%	820	97.74%	17	2.03%	839
PRET	2586	97.47%	66	2.49%	1	0.04%	2653
PERF (*hân*)	1	0.89%	108	96.43%	3	2.68%	112
PERF (*sîn*)	0	0%	18	100%	0	0%	18
PLUP (*hân*)	80	100%	0	0%	0	0%	80
PLUP (*sîn*)	43	97.73%	1	2.27%	0	0%	44
sum	2712		1013		21		3746

While there have been many studies on the functional difference between the present tense, the preterite and the present perfect, the pluperfect has not been the focus of attention. This may be because the development of the pluperfect

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

has been seen as parallel to that of the present perfect. It is assumed that both perfect constructions follow the same grammaticalisation path from a resultative construction to a tense form with eventive meaning. Furthermore, its semantics is assumed to be straightforward, signalling "that there is a reference point in the past, and that the situation in question is prior to that reference point, i.e. the pluperfect can be thought of as 'past in the past'" (Comrie 1985: 65). In terms of Reichenbach (1947), its event time (te) is set before a reference time (tr), which in turn is set before the speech time (ts), i.e.: te < tr < ts. As such, its prototypical meaning is seen in the designation of the temporal order of events within the past, such as in (11).

(11) *Dô sî ditz hâten vernomen, dô sprach der rîter mittem leun*
'When they had heard that, then the knight with the lion spoke.'
[*Iwein* 6108]; (example taken from Paul 2007: § S10)

The pluperfect seems to have the same semantics as the perfect, with the only difference being the past meaning of the auxiliary.

In terms of their relations to discourse mode, however, Table 3 evidently shows that the present perfect and the pluperfect differ in their preferences for narrative vs. non-narrative discourse mode. While the present perfect in Old and Middle High German cannot be used to denote a sequence of past events and is therefore restricted to non-narrative discourse modes, the pluperfect is used almost exclusively in narrative discourse mode. This preference runs through the entire history of language. To this day, the pluperfect is seen as a tense of written language and a form of "distance" (e.g. Luscher 2011: 24). As will be shown below, this impression is related to the preference of the pluperfect for the narrative discourse mode, which in turn leads to a different semantic development of the past perfect in comparison to the present perfect.

A major methodological problem in examining the use of the pluperfect in narrative vs. non-narrative discourse modes is the fact that the pluperfect is not used very frequently. This is particularly true for non-narrative modes of discourse. In Table 3, there is only one instance of the pluperfect that is not documented in the narrative discourse mode. Looking at such rare instances, however, it is striking that the early uses in the non-narrative discourse mode do not seem to have the prototypical meaning of the pluperfect as described in (11), cf. (12) for comparison.

(12) *Was hettet ir vernomen, gut öheim, das ir des wondent?*
'What had you heard, good uncle, that you could believe that?'
[*Prosalancelot*, 2, 25]

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

In (12), the speaker is referring to an event that happened in the past as seen from the time of speech. The example shows the semantics of a pluperfect by indicating that the hearing of the news was already completed at a certain time in the past, i.e. te < tr < ts. However, due to the dialogue situation, there is a strong reference to the actual speech time. This has the effect that the reference point tr is less dominant than in cases such as (11), where the past reference point is given by the context. Consequently, one event serves as the reference time for the other one, while both events are located in a past narrative frame before the speech time.

The weaker status of tr in (12) also depends on the interaction between tenses and narrative discourse mode. As already mentioned, the narrative discourse mode is characterised by a different deictic pattern than non-narrative discourse modes, cf. Smith (2003):

> In Narrative, situations are related to each other and dynamic Events advance narrative time. In Reports, situations are related to Speech Time and time progresses forward and backward from that time. (Smith 2003: 13)

The narrative discourse mode is characterised by the temporal sequence of events in a deictically displaced story world. As such, the (dynamic) *story now* provides a reference point for anchoring the events. In non-narrative discourse modes, such as reports, the deictic anchor is not localised within a story world but is set by the deictic origo of speech time. If a third reference point is not contextually given within the text, the impression arises that tr – and hence the semantics of the pluperfect – seems "weaker", since there is no reference point that distances the time of the event and the time of speech.

This observation is particularly noteworthy since Hennig (2000) and Bertinetto (2010) have described similar cases of "r-weakening" (term by Bertinetto 2010) for the pluperfect in Modern German. Based on their investigations, they suggest that the semantics of the pluperfect has changed from past perfect (te < tr < ts) to a general past tense (te < ts) – which would be in line with the cross-linguistic grammaticalisation path of the pluperfect as described by Squartini (1999). Strikingly, the examples discussed by Hennig are attested in instances of spoken language (Hennig 2000: 67). In the light of the foregoing, it might be assumed that the development was triggered by non-narrative discourse mode.

To test this hypothesis, a pilot study was conducted. The data sample consists of 982 instances of the pluperfect from Middle High German texts. The analysis was based on Middle High German epic texts (*Herzog Ernst*, Konrad von Würzburg: *Trojanerkrieg*), a verse epic with a first-person narrator (Rudolf von

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Ems: *Der Guote Gerhart* (1220)) and verse narratives by *Der Stricker* which are considered to be close to orality. All occurrences were analysed with regard to discourse mode, text type, 1st, 2nd and 3rd person, matrix clause vs. subordinate clause as well as further contextual features such as adverbials, negation, etc. Table 4 shows the distribution of the pluperfect (i) in the entire text, (ii) in first and second person, and (iii) in non-narrative contexts (as laid out above, passages of narrative passages within direct speech were counted as 'narrative'). Furthermore, the occurrences were semantically analysed with respect to a possible r-weakening. Although the absolute number of the pluperfect in non-narrative discourse mode is low (n = 15), it is obvious that the pluperfect in non-narrative contexts tend to have a meaning te < (tr) < ts. It is also important to note that the non-narrative context has more influence than the use of the pluperfect in the first or second person.

Table 4: R-weakening of the pluperfect in Middle High German.

	in total	1st/2nd person	non-narrative
r-weakening	(12/982) 1.22%	(7/28) 25.00%	(11/15) 73.33%

Table 4 supports the above hypothesis that the semantic change occurs primarily in non-narrative contexts. Although the instances in the Middle High German corpus do not show cases of general past semantics (as assumed for Modern German by Hennig 2000), the high amounts of r-weakening suggest that narrative and non-narrative contexts have an impact on the meaning of grammatical forms. This can be said for the synchronic state of Middle High German. In comparison with the more recent development in Modern German, it can be assumed that the motivation for this development lies in the distinction between discourse modes.

This hypothesis should be tested with further diachronic data. If confirmed, it would suggest that narrative and non-narrative discourse modes are not only relevant in terms of the synchronic distribution of tense forms. They also seem to be able to trigger different grammaticalisation pathways. In contrast to the present perfect, which is initially restricted to non-narrative contexts and then expands into the narrative discourse mode in the course of linguistic history, the pluperfect initially appears only in the narrative discourse mode. But it is not restricted to this context. When the pluperfect is used in non-narrative contexts, there is a shift in meaning that triggers a development that is different from the

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

development of the pluperfect forms used in narrative discourse. Because of the rare use of the pluperfect in non-narrative discourse context, one might assume that this shift is a recent development related to the qualitative spread of the pluperfect into non-canonical contexts of use. The observations on Middle High German, however, suggest that the development is already mapped out in the history of German. More precisely, in the distinction between narrative and non-narrative discourse modes.

4 The impact of narration

The aim of the article was to examine narration as an important factor in the interaction between grammar and discourse structure in relation to language change. In order to allow for a systematic investigation of the effects of narration, it has been argued that there are three different dimensions of narration that need to be distinguished – (i) narrative *texts* that exhibit a narrative macrostructure, (ii) narrative *discourse mode*, i.e. narrative segments that can occur in (narrative and non-narrative) texts, and (iii) narrative *discourse relations* between propositions. The effects of these different dimensions of narration have been illustrated by means of two empirical case studies, leading to the following conclusions:

1. The analysis of left dislocations in Middle High German has shown that (i) they have functions specific to narrative texts, (ii) that their distribution and textual functions differ in narrative and non-narrative discourse modes and (iii) that this is linked to different preferences for discourse relations. Against this background, it has also become clear that the line of development of LDs from 1200 to 1600 is not straight. While there is a general decline of LD in the history of German, there are also micro-changes in the dependence of discourse modes and discourse relations. This supports the claim that all three dimensions of narration must be taken into account in order to adequately investigate the development of LD in the history of German.

2. The analysis of the pluperfect in Middle High German has shown that present and past perfect constructions are used complementarily in non-narrative and narrative discourse modes. With regard to the development of the pluperfect, it has been shown that the occurrences in the non-prototypical context of non-narrative discourse modes tend to show "r-weakening" and thus a different semantics than the pluperfect forms used in narrative contexts. This suggests that the development of the pluperfect is different

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

in narrative and non-narrative discourse modes and that the underlying discourse mode may trigger different grammaticalisation pathways.

In sum, it can be said that narration – in its various dimensions – is an important factor in the interaction between grammar and discourse structure. In this context, it is also important to note that both the distinction between narrative and non-narrative discourse modes and the distinction between narrative and non-narrative discourse relations allow for a finer and more systematic distinction than the distinction between text types and text genres, which has so far been one of the central parameters of investigation in many studies of the structure of discourse in historical languages. The distinction between text types and text genres, however, faces considerable difficulties. Apart from the problem of inconsistent classifications, diachronic comparisons are difficult because not every text type or genre occurs in every period. The differentiation of discourse modes and discourse relations, on the other hand, allows not only a finer distinction within the text, but also comparisons across centuries. In this respect, further research into discourse modes and discourse relations could help to answer the question of whether the effects of text type observed in many studies are a causal effect or rather an epiphenomenon of the underlying discourse structure (see also Speyer 2022: 2224).

Primary sources

Das *Nibelungenlied* nach der Handschrift C. Edited by Ursula Hennig. Tübingen: Niemeyer 1977.

Gottfried von Straßburg: *Tristan*. Edited by Karl Marold. Berlin: De Gruyter 2016.

Berthold von Regensburg. Sælic sint die reines herzen sint. In *Deutsche Neudrucke*. Edited by Franz Pfeiffer. Berlin: De Gruyter 2021.

Herzog Ernst. Ein mittelalterliches Abenteuerbuch. In der mittelhochdeutschen Fassung B nach der Ausgabe von Karl Bartsch mit den Bruchstücken der Fassung A. Edited and translated by Bernhard Sowinski. Stuttgart: Reclam 1979.

Prosalancelot. Heidelberger Handschrift Cod. Pal. Germ. 147. Edited by Reinhold Kluge. Berlin: Deutscher Klassiker Verlag 1948.

Der Stricker. *Verserzählungen I und II*. Edited by Hanns Fischer & Johannes Janota. Berlin: De Gruyter 1997/2000.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

References

Ágel, Vilmos & Mathilde Hennig (eds.). 2006. *Grammatik aus Nähe und Distanz: Theorie und Praxis am Beispiel von Nähetexten 1650-2000*. Tübingen: Niemeyer. DOI: 10.1515/9783110944709.

Altmann, Hans. 1981. *Formen der Herausstellung im Deutschen: Rechtsversetzung, Linksversetzung, Freies Thema und verwandte Konstruktionen*. Tübingen: Niemeyer.

Asher, Nicolas & Alex Lascarides. 2003. *Logics of conversation*. New York: Cambridge University Press.

Averintseva, Maria. 2007. Links und rechts vom Satz: Satzperipherien im Deutschen und ihre Rolle im Diskurs. In Mari Tarvas, Sonja Pachali, Aigi Heero, Merle Jung, Helju Ridali & Maris Saagpakk (eds.), *Linguistik und Didaktik. Beiträge der Tagung „Tradition und Zukunft der Germanistik"*, vol. 2, 137–149. Tallinn: TLÜ Kirjastus.

Bally, Charles. 1932. *Linguistique générale et linguistique française*. Paris: Librairie Ernest Leroux.

Behaghel, Otto. 1932. *Deutsche Syntax: Eine geschichtliche Darstellung. Bd. IV: Wortstellung. Periodenbau*. Heidelberg: Winter.

Benveniste, Émile. 1974[1972]. *Probleme der allgemeinen Sprachwissenschaft*. München: List.

Bertinetto, Pier Marco. 2010. Non-conventional uses of the Pluperfect in the Italian (and German) literary prose. In Emmanuelle Labeau & Jacques Bres (eds.), *Evolution in Romance verbal systems*, 145–170. Frankfurt am Main: Peter Lang. DOI: 10.3726/978-3-0352-0233-5.

Biber, Douglas & Susan Conrad. 2019. *Register, genre, and style*. 2nd edn. Cambridge: Cambridge University Press.

Biber, Douglas & Bethany Gray. 2013. Being specific about historical change: The influence of sub-register. *Journal of English Linguistics* 41(2). 104–134. DOI: 10.1177/0075424212472509.

Biber, Douglas, Bethany Gray, Shelley Staples & Jesse Egbert (eds.). 2021. *The register-functional approach to grammatical complexity: Theoretical foundation, descriptive research findings, application*. London: Routledge.

Brown, LeAnn & Sali Tagliamonte. 2012. A really interesting story: The influence of narrative in linguistic change. *University of Pennsylvania Working Papers in Linguistics* 18(2).

Chafe, Wallace. 1994. *Discourse, consciousness and time: The flow and displacement of conscious experience in speaking and writing*. Chicago: University of Chicago Press.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Cinque, Guglielmo. 1977. The movement nature of left dislocation. *Linguistic Inquiry* 8(2). 397–412.

Comrie, Bernard. 1985. *Tense*. Cambridge: Cambridge University Press.

Dahl, Östen. 1985. *Tense and aspect systems*. Oxford: Blackwell.

Demske, Ulrike. 2018. Syntax and discourse structure: Verb-final main clauses in German. *Linguistische Berichte* 25. Special issue: Non-canonical verb positioning in main clauses, 135–159.

Dewald, Anika. 2012. *Versetzungsstrukturen im Deutschen: Zu ihrer Syntax, Prosodie und Diskursfunktion*. Köln: Universität zu Köln. (Doctoral dissertation).

Dijk, Teun van. 2015. *Some aspects of text grammars: A study in theoretical linguistics and poetics*. Berlin: De Gruyter.

Drinka, Bridget. 2017. Romance Perfects, Aorists, and the Role of 'Aoristic Drift'. In Marc Fryd & Pierre-Don Giancarli (eds.), *Aorists and perfects*, 5–24. Leiden: Brill. DOI: 10.1163/9789004326651_003.

Egbert, Jesse & Michaela Mahlberg. 2020. Fiction – one register or two? Speech and narration in novels. *Register Studies* 2(1). 72–101. DOI: 10.1075/rs.19006.egb.

Elspaß, Stephan. 2005. *Sprachgeschichte von unten: Untersuchungen zum geschriebenen Alltagsdeutsch im 19. Jahrhundert*. Tübingen: Niemeyer.

Elspaß, Stephan. 2010. Zum Verhältnis von ‚Nähegrammatik' und Regionalsprachlichkeit in historischen Texten. In Vilmos Ágel & Mathilde Hennig (eds.), *Nähe und Distanz im Kontext variationslinguistischer Forschung*, 63–84. Berlin: De Gruyter. DOI: 10.1515/9783110220872.63.

Fiehler, Reinhard. 2009. Gesprochene Sprache. In Peter Eisenberg & Kathrin Kunkel-Razum (eds.), *Duden. Die Grammatik*, 8th edn., 1165–1244. Berlin: Dudenverlag.

Fischer, Hanna. 2020. Diskursmodus und Tempusformen. Zum Tempusgebrauch in den frühneuhochdeutschen Hexenverhörprotokollen. In Renata Szczepaniak, Lisa Dücker & Stefan Hartmann (eds.), *Hexenverhörprotokolle als sprachhistorisches Korpus* (Reihe Germanistische Linguistik 322), 211–238. Berlin: De Gruyter. DOI: 10.1515/9783110679649.

Fleischman, Suzanne. 1990. *Tense and narrativity: From medieval performance to modern fiction*. London: Routledge. DOI: 10.7560/780903.

Fludernik, Monika. 2003. The diachronization of narratology. *Narrative* 11(3). 331–348. DOI: 10.1353/nar.2003.0014.

Fludernik, Monika. 2012. Narratology and literary linguistics. In Robert I. Binnick (ed.), *The Oxford handbook of tense and aspect*, 75–101. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780195381979.013.0002.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Frey, Werner. 2004. Notes on the syntax and the pragmatics of German left dislocation. In Horst Lohnstein & Susanne Trissler (eds.), *The syntax and semantics of the left periphery*, 203–234. Berlin: De Gruyter. DOI: 10.1515/9783110912111.203.

Gloning, Thomas. 2010. Funktionale Textbausteine in der historischen Textlinguistik: Eine Schnittstelle zwischen der Handlungsstruktur und der syntaktischen Organisation von Texten. In Arne Ziegler & Christian Braun (eds.), *Historische Textgrammatik und Historische Syntax des Deutschen*, 173–194. Berlin: De Gruyter. DOI: 10.1515/9783110219944.173.

Goulart, Larissa, Bethany Gray, Shelley Staples, Amanda Black, Aisha Shelton, Douglas Biber, Jesse Egbert & Stacey Wizner. 2020. Linguistic perspectives on register. *Annual Review of Linguistics* 6. 435–455. DOI: 10.1146/annurev-linguistics-011718-012644.

Hamburger, Käte. 1968[1957]. *Logik der Dichtung*. Stuttgart: Klett.

Hennig, Mathilde. 2000. *Tempus und Temporalität in geschriebenen und gesprochenen Texten*. Tübingen: Niemeyer.

Hennig, Mathilde. 2009. *Nähe und Distanzierung: Verschriftlichung und Reorganisation des Nähebereichs*. Kassel: University Press. http://nbn-resolving.de/urn:nbn:de:hebis:34-2010011231675.

Hinterhölzl, Roland, Svetlana Petrova & Michael Solf. 2005. Diskurspragmatische Faktoren für Topikalität und Verbstellung in der ahd. Tatianübersetzung (9. Jh.) In Michaela Schmitz Shinichiro Ishihara & Anne Schwarz (eds.), *Approaches and findings in oral, written and gestural language*, 143–182. Potsdam: Universitätsverlag.

Hobbs, Jerry R. 1979. Coherence and coreference. *Cognitive Science* 3(1). 67–90.

Kytö, Merja. 2019. Register in historical linguistics. *Register Studies* 1(1). 136–167. DOI: 10.1075/rs.18011.kyt.

Labov, William & Joshua Waletzky. 1967. Narrative Analysis. In June Helm (ed.), *Essays on the verbal and visual arts*, 12–44. Seattle: University of Washington Press. DOI: 10.1075/jnlh.7.1-4.02nar.

Larrivée, Pierre. 2022. Is Medieval French diglossic? New evidence on remnant V2 and register. *Isogloss. Open Journal of Romance Linguistics* 8. 1–16. DOI: 10.5565/rev/isogloss.114.

Lötscher, Andreas. 1994. Herausstellung nach links in diachroner Sicht. *Sprachwissenschaft* 24. 32–63.

Luscher, Renate. 2011. *Übungsgrammatik für Anfänger: Deutsch als Fremdsprache*. München: Hueber.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Mazziotta, Nicolas & Julie Glikman. 2019. Oral représenté et narration en ancien Français. Spécificités syntaxiques dans trois textes de genres distincts. *Linx*. DOI: 10.4000/linx.3151.

Padučeva, Elena. 2011. *The linguistics of narrative: The case of Russian*. Saarbrücken: LAP LAMBERT.

Paul, Hermann. 2007. *Mittelhochdeutsche Grammatik*. Thomas Klein, Hans-Joachim Solms & Klaus-Peter Wegera (eds.). Tübingen: Niemeyer. DOI: 10.1515/9783110942354.

Prell, Heinz-Peter. 2005. Zur Neubearbeitung der mittelhochdeutschen Syntax von Ingeborg Schröbler. Mit einem Vergleich von Vers- und Prosatexten am Beispiel der Nominalphrase. In Franz Simmler (ed.), *Syntax. Althochdeutsch - Mittelhochdeutsch. Eine Gegenüberstellung von Metrik und Prosa. Akten zum Internationalen Kongress an der Freien Universität Berlin 26. bis 29. Mai 2004*. Weidler Buchverlag.

Reichenbach, Hans. 1947. *Elements of symbolic logic*. New York: The Free Press. DOI: 10.2307/2268978.

Selting, Margret. 1993. Voranstellungen vor den Satz: zur grammatischen Form und interaktiven Funktion von Linksversetzung und Freiem Thema in Deutschland. *Zeitschrift für Germanistische Linguistik* 21. 291–319.

Smith, Carlota S. 2003. *Modes of discourse: The local structure of texts*. Cambridge: Cambridge University Press.

Speyer, Augustin. 2013. Performative Mündlichkeitsnähe als Faktor für die Objektstellung im Mittel- und Frühneuhochdeutschen. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 135(3). 342–377. DOI: 10.1515/pbb-2013-0089.

Speyer, Augustin. 2022. Discourse relations and the German prefield. In Nicholas Catasso, Marco Coniglio & Chiara de Bastiani (eds.), *Language change at the interfaces. Intrasentential and intersentential phenomena*, 215–233. Amsterdam: Benjamins.

Squartini, Mario. 1999. On the semantics of the Pluperfect: Evidence from Germanic and Romance. *Linguistic Typology* vol. 3(1). 51–90. DOI: 10.1515/lity.1999.3.1.51.

Weinrich, Harald. 1964. *Tempus: Besprochene und erzählte Welt*. München: C. H. Beck.

Werth, Alexander. 2020. Klisen in frühneuzeitlichen Hexenverhörprotokollen. In Renata Szczepaniak, Lisa Dücker & Stefan Hartmann (eds.), *Fallstudien zur Erschließung der frühneuzeitlichen Schriftsprache*, 177–210. Berlin: De Gruyter. DOI: 10.1515/9783110679649-007.

Zäch, Alfred. 1931. *Der Nominativus pendens in der deutschen Dichtung des Hochmittelalters*. Bern: Paul Haupt.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Zeman, Sonja. 2010. *Tempus und „Mündlichkeit“ im Mittelhochdeutschen: Zur Interdependenz grammatischer Perspektivensetzung und „Historischer Mündlichkeit“ im mittelhochdeutschen Tempussystem*. Berlin: De Gruyter. DOI: 10.1515/9783110227697.

Zeman, Sonja. 2018. What is a narration – and why does it matter? In Markus Steinbach & Annika Hübl (eds.), *Linguistic foundations of narration in spoken and sign language* (Linguistics Today 247), 173–206. Amsterdam: Benjamins. DOI: 10.1075/la.247.08zem.

Zeman, Sonja. 2020. Grammatik der Narration. *Zeitschrift für Germanistische Linguistik* 48(3). 457–494. DOI: 10.1515/zgl-2020-2011.

Zeman, Sonja. 2022. Paradoxes of ‘Orality’: A Comparison between Homeric Oral Poetry and the Heroic and Courtly Epics in Middle High German. In Andrea Ercolani & Laura Lulli (eds.), *The mechanisms of the oral communication system in the case of the archaic epos*, 177–206. Berlin: De Gruyter. DOI: 10.1515/9783110751963-011.

Zeman, Sonja. 2023a. ‚Linksversetzungen‘ im Mittelhochdeutschen als Beispiel narrativer Syntax. *Jahrbuch für Germanistische Sprachgeschichte* 14(1). 68–84. DOI: 10.1515/jbgsg-2023-0007.

Zeman, Sonja. 2023b. Oral storytelling in Ancient Greek and German medieval literature. In Peter Hühn, John Pier & Wolf Schmid (eds.), *Handbook of diachronic narratology*, 684–709. Berlin: De Gruyter. DOI: 10.1515/9783110617481-034.

Ziegler, Arne. 2010. Historische Textgrammatik und Historische Syntax des Deutschen – Eine kurze Einleitung. In Arne Ziegler & Christian Braun (eds.), *Historische Textgrammatik und Historische Syntax des Deutschen*, 1–8. Berlin: De Gruyter. DOI: 10.1515/9783110219944.1.

Acknowledgements

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Chapter 3

Structuring the narrative with preposed adverbial clauses: a study of the German and Dutch Ponthus adaptations

Barthe Bloom [a]
[a]Universität Potsdam

This study investigates the use of preposed adverbial clauses in the Early New High German and Dutch adaptations of the Ponthus et la belle Sidoine. The focus lies on so-called resumptive constructions, that is, on adverbial clauses followed by *so/soo* or by a temporal adverb (e.g., *da* 'then'). The study shows that the choice between these resumptive constructions correlates with a distinction between mimesis and diegesis and they are associated with different narrative speed.

For German, adverbial clauses that are followed by *so* tend to occur in scenes and are associated with isochrony, whereas those followed by the temporal adverb *da* 'then' occur in summaries and are used to continue the story. This is correlated with the type of adverbial clauses that occur in the constructions: The prototypical protasis-apodosis construction – V1 clause + *so* – is more dominant in dialogues, whereas a temporal adverbial clause with *da* tends to occur in summary.

The same tendency is visible for Dutch, albeit weaker. This is explained, as the temporal resumptive construction, which was dominant in German, is primarily replaced by non-resumed adverbial clauses. These constructions are similarly preferred in in diegetic parts of the discourse that progress the story.

1 Introduction

This study investigates the use of adverbial V3 patterns in the Early New High German (GE) and Dutch (DU) Ponthus. The primary focus lies on preposed adverbial clauses (henceforth PAC) followed by *so/soo* (1a) or by a temporal adverb

Barthe Bloom. 2025. Structuring the narrative with preposed adverbial clauses: a study of the German and Dutch Ponthus adaptations. In Ulrike Demske & Barthe Bloom (eds.), *Discourse structure and narration: A diachronic view from Germanic*, 37–71. Berlin: Language Science Press. DOI: ??

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(e.g., *da* 'then') (1b), known as or correlative constructions (e.g., Pittner 1999, Meklenborg 2020, Axel-Tober 2023).

(1) a. *ob sie noch kein man haitt* ***so*** *wil ich einen man mittbringen*
if she not.yet no man has.SBJV SO want I a man with.bring
'If she does not yet have a man, I will bring a man with me.'
[GE, 119va]

b. *Da ir glenen waren zu Brochen* ***da*** *zugen sie die swert*
then.CNJ their lances were to broken TEMP drew they the swords
'When their lances had broken, they drew the swords.' [GE, 66va]

A terminological note: I will refer with *so* and *da* to the preverbal use of *so* and *da* as illustrated above, viz. following adverbial clauses and preceding the finite verb of the host sentence, unless otherwise specified. With TEMP, I refer to the use of temporal adverbs in the prefield following PACs, including *da*.

In Present-Day German, *so* and *da* are associated with different registers: *so* after adverbial clauses is found in formal, written register, whereas *da* is found after adverbial clauses in spoken, colloquial language (Condoyannis 1944, Weinert 2007, Catasso 2021). If this distribution is a continuation of a historical situation, it would be expected that *da* would be found primarily in parts of a text in which spoken language is reflected in earlier stages of German, whereas *so* would be the more prevalent construction in more formal segments. What we find is that the constructions do statistically correlate with a distinction between mimesis (*μίμησις* from *μιμεῖσθαι* 'to imitate'), where characters' speech is imitated, and diegesis (*διήγησις* from *διηγεῖσθαι* 'to narrate'), narrated segments. However, the results point in the opposite direction than what is expected: PACs that are followed by *so* tend to occur in mimetic segments, whereas constructions with *da* occur primarily in diegetic parts of the text. To understand this and to explain the distribution of the patterns in the text, the study evaluates the role of narrative speed (e.g., Genette 1983[1980], Packard 2008). Specifically, it is argued that, although constructions statistically correlate with a distinction between mimesis and diegesis, this is better accounted for in terms of narrative speed (e.g., Genette 1983[1980], Packard 2008).

Mimesis in narrative prose is found in segments of direct speech, in which the words of characters are presented as if the characters utter them directly. These sections are associated with isochrony, viz., the simultaneous progression of story time (i.e., the events that are reported) and narration time (i.e., the reporting of the events). Differently, in diegetic segments, the events of the story are often presented by the narrator as narrative summaries. The events of the

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

story unravel more rapidly than the narration event in these sections. As such, narrative summaries continue the story at a higher pace.

The results show that *da* is associated with narrative summary, whereas *so* occurs in segments in which the story progresses more slowly. This is correlated with the type of PACs that occur in the constructions: The V1-conditional clause, introducing alternative events, is more dominant in dialogues and strongly associated with *so*, whereas temporal clauses – normally combined with *da* – tend to occur in narrative summary.

A comparison with the Dutch adaptation of the Ponthus finds a similar tendency in Dutch, although the effect is weaker: *Soo* occurs with sentences that introduce alternative events and is associated with isochrony, whereas temporal adverbs are found in narrative summaries. This is illustrated in (2a) and (2b).

(2) a. *kinderen is dat sake dat ghi niet steruen en wilt* ***soo*** *blijft hier binnen*
children is that sake that you not die NEG want SO stay here inside
'Children, if you do not want to die, stay here inside.' [DU, A3ra]

b. *Als hi dat hoorde* ***terstont*** *dede hy hem brenghen eten ende drincken.*
when he that heard TEMP did he them bring food and drink
'When he heard that, he immediately let food and drink be brought for them.' [DU, A3vb]

The weaker effect is primarily due to a decreased use of the correlative structure with a temporal adverb that is highly frequent in German. This becomes apparent once other configurations with PACs are taken into consideration. Particularly relevant are PACs that are followed by a subject (viz. the juxtaposition of the PAC) and the immediate adjacency of the PAC and the finite verb of the host (typically the integration of the PAC into the host). These patterns are illustrated for Dutch in (3a) and (3b) respectively.

(3) a. *ist dattet pontus verneemt* ***hi salt*** *wel beletten*
is.it that.it Ponthus perceived he will.it DISC prevent
'If Ponthus learns from it, he will prevent it.' [DU, j2ra]

b. *Als landri dat vernam* ***hief*** *hy op sijn swaert*
when Landry that perceived heaved he up his sword
'When Landry learned of this, he raised his sword.' [DU, f1vb]

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

What we will see is that the dominant realizations of PACs in the German Ponthus are not the same as in the Dutch text, where juxtaposition and integration are more prevalent. In particular, where German uses *da* in narrative summaries, integration is the dominant strategy to realize PACs in narrative summaries in Dutch.

The paper is structured as follows. §2 discusses in general the role of diegesis and mimesis in narrative prose and the connection to narrative speed. In §3, I introduce the texts from which the data is extracted, the data collection process, the annotation, and the statistical methods that are used in this paper. §4 presents the analysis of the two correlative constructions in German and in Dutch, discussing first the relation to diegesis and mimesis and subsequently the relation to narrative speed. §5 includes other constructions with PACs. §6 concludes this contribution.

2 Style, voice, and speed in narrative prose

This section first introduces mimesis, diegesis, and extradiegesis as different narrative styles in §2.1 and relates this to the concept of voice. Thereafter, narrative style is related to narrative speed in §2.2.

2.1 Mimesis, diegesis, and extradiegesis

The distinction between mimesis and diegesis, originally introduced by Plato, has been subject to a long tradition in which the concepts themselves and the relation between them have been interpreted and analyzed in many a way. In essence, the idea is that there is a difference between imitation (mimesis) and narration (diegesis). This distinction in style applies beyond prose text across other media, as for example evident in film (Chatman 1978). I will focus on narrative prose, focusing on and adopting de Jong's (2004) interpretation of Plato's classification as it relates to narrative prose.

First, I deem it useful to define what is meant with *story*, *narration* and *text* and to introduce these layers, as these terms have been used with (partially) overlapping definitions. I here follow Bal (2009[1985]) and de Jong (2004) in much of the narratological definitions and methodologies. The *story* refers to the events that are being told. These events do not necessarily have to be temporally ordered, but they have been organized and interpreted in some way.[1] The *narration* is the

[1]The story is the result of the focalization of the *fabula*, which is the logically or chronologically connected sequence of events (Bal 2009[1985], de Jong 2004). The layer of the *fabula* will not be central in the later analysis and will therefore not be introduced here.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

telling of the events, and the *text* is the result of this telling event.

Stories told by literary means are always mediated by a narrator (Fludernik 1993: 28) and thus literary texts are entirely achieved by diegesis, i.e., they are the product of narration, of telling (de Jong 2004: 2–3). When talking about mimesis as a component of literary narrative texts, this additional layer of the narrator – which may be absent in other media – consequently complicates its definition. Plato and Aristotle analyzed speeches in a different way than the rest of the text. Plato views this as a distinction between mimesis and non-mimesis. He identifies *diegesis haple* (single narration), in which the poet speaks as himself, and *mimesis*, in which the poet impersonates a character. Aristotle, differently, seems to make a distinction between the poet speaking in first person as the poet (in for example the preface of a text) and the poet speaking as a narrator. The latter is considered to be mimetic by Aristotle, only the former is non-mimetic according to him. In Plato's division, both would fall under diegesis haple (de Jong 2004: 6-8, 38).

Plato has the more restricted application of the concept of mimesis, which I will follow here. However, since he did not distinguish between narrator and poet, a slight modification is in order. Therefore, I here assume that the narrator speaking as the narrator, i.e., as itself, is diegetic, whereas the narrator speaking as if a character is mimetic.

Both mimesis and diegesis can be illustrated by means of the example in (4). The first sentence of the example presents a diegetic segment. The external narrator – who is not a character in the story – tells us an event of the story. In contrast, character text – as presented in between quotation marks in (4) – is considered to be mimetic.

(4) Loki went to their underground forge. "Hello, sons of Ivaldi. I have asked around, and people here tell me that Brokk and Eitri, his brother, are the greatest dwarf craftsmen there are or have been," said Loki.
(Gaiman 2017: 54)

Technically, the quoted sentence is uttered by the same narrator as in the previous sentence. He imitates the voice of one of the characters, i.e., of Loki, and the events are told as if by him. As such, the answer to the question "who speaks?" is dual: Both the narrator and the character can be heard. Without going into much detail, this is possible due to the different layers a narration is build up off (Bal 2009[1985]) and the character's voice is embedded within that of the narrator. We will here, for simplicity sake, speak of a character's voice in such cases.

Cases of indirect speech are, following de Jong (2004) but against Genette (1980[1972]) taken to be diegetic, because it is the narrator speaking as the nar-

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

rator. The crucial distinction between indirect speech and other segments of diegesis lies in a difference in the embedded focalization (de Jong 2004: 255), viz. a difference in who sees, orders and interprets. In those cases, it is the perspective of one of the characters being filtered and voiced by the narrator without imitating a character.

I recognize that prefaces are different in some way. In the prose, they do not typically present the events of the story but comment on it and give context to the creation of the manuscript. Such instances in which the narrator speaks but does not narrate the story may also be found in other parts of the text than the preface. An example is given in (5).

(5) There. That is the story of the mead of poetry and how it was given to the world. It is a story filled with dishonor and deceit, with murder and trickery. But it is not quite the whole story. There is one more thing to tell you. The delicate among you should stop your ears, or read no further. (Gaiman 2017: 151)

Here, the events of the story do not come into play. Instead, the narrator evaluates the developments in the story that were just presented and comments on the structure of the narration. As such the "narrative level (...) is external to that of the world represented in the primary narrative" (Pier 1986: 211). In such cases, we speak of an extradiegetic segment.

In sum, the distinguishing factor between mimesis and diegesis haple – in the rest of the paper referred to as diegesis – is who speaks: Technically, if there is embedded character narration within the external narrator that we find in the Ponthus, we speak of mimesis. In the other cases, the narrator speaks in his own voice. These are mostly segments of diegesis but may also be cases of extradiegesis, in which not the story itself but something else is presented.

2.2 Narrative speed

The different styles of narration are associated with differences in narrative speed (e.g Genette 1980[1972], 1983[1980], Packard 2008). I will here first discuss the definition of narrative speed adopted in this paper and relate it to the notions of mimesis, diegesis, and extradiegesis.

The rate at which the story progresses in relation to the narration can be higher or lower, resulting in different velocities. This is what is here understood by narrative speed. Genette (1980[1972]: 95) distinguishes primarily between four relations between story time (ST) and narration time (NT):

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(i) Pause, in which NT continues, but ST halts (NT ∞ > ST);
(ii) Scene, in which NT and ST coincide (NT = ST);
(iii) Summary, in which NT goes quicker than ST (NT < ST); and
(iv) Ellipsis, in which ST is not represented in NT (NT < ∞ ST).

The theoretically possible stretch (NT > ST) is not canonically used, as slowing down of the narration is typically done by means of pauses (Genette 1980[1972]: 93-112).

Of course, this is a rather simplified representation of narrative speed (Packard 2008, Kukkonen 2020), but assuming that the base-line of isochrony is the dialogue, this categorical division suffices for the purposes of this paper. In fact, ellipsis will not be relevant, as this results in an absence of linguistic structure. Thus we are left with a three-part division: pause, scene, and summary.

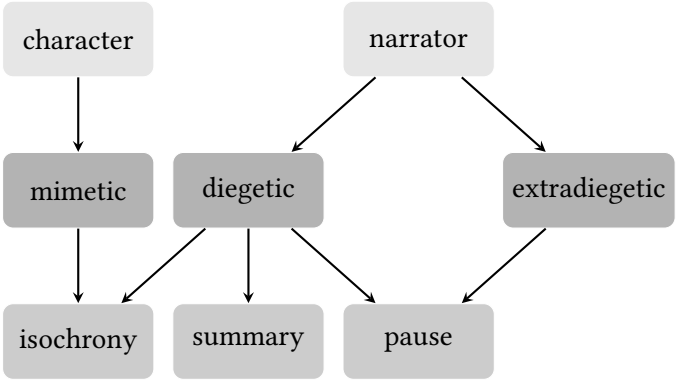


Figure 1: Narrative style and narrative speed in the Ponthus adaptations

The connection between the different narrative styles and the narrative speed is visualized in Figure 1. In the Ponthus adaptations that will be discussed, a text segment is either the product of a narration by an external but present narrator (i.e., homodiegetic), or by a character (and narrator). If it is a character speaking, i.e., in case of mimesis, narration and story time coincide, and we speak of isochrony. This is illustrated by the direct speech in (4).

Whenever the narrator is speaking, the situation becomes more complex, as what is being narrated can either be the story or something else than the story; the style can then be either diegetic or extradiegetic. Segments of extradiegesis pause the story, as was illustrated in (5), but segments of diegesis may constitute a pause or a summary. Example (6) illustrates this alternation nicely. The segment starts with a description, as state of affairs, which does not temporally progress

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

the story but still concerns it – a diegetic pause. With the introduction of the explicit temporal adverbial *over time*, the story temporally progresses, and a part of the story that must have taken centuries, if not millennia, is narrated in one sentence – a summary.

(6) Between Muspell and Niflheim was a void, an empty place of nothingness, without form. The rivers of the mist world flowed into the void, which was called Ginnungagap, the "yawning gap". **Over time** beyond measure, these posed rivers, in the region between fire and mist, slowly solidified into huge glaciers. (Gaiman 2017: 10)

Finally, diegetic segments may be isochronous, or very close to the temporal progression of the story in mimesis. Segments of indirect speech and free indirect speech may very closely mimic the words of a character, while the deictic origo and the voice do not lie with the character but the narrator. This is illustrated in (7), where the question closely imitates the presumed words of the man in the audience, yet the personal pronouns signal that this discourse segment is here spoken by the narrator.

(7) A man in the audience with a misshapen leg stood up and challenged her: was she seriously suggesting that the restoration of her legs was comparable to the loss of his wife? (Chiang 2002: 130)

In sum, mimesis is isochronous, diegesis is found with a variety of velocities, and extradiegesis pauses the story progression. The particular application of these categories to the current data will be discussed in §3.3.

3 Methodology

Prior to the analysis, this section concerns the methodology used in the present study. The texts that are used are presented in §3.1. The way in which the data is collected is explained in §3.2. §3.3 discusses the annotation of the data, and §3.4 mentions the statistical methods used.

3.1 The texts

The data used in the current study originate from a German and a Dutch reworking of the originally French prose text *Ponthus et la belle Sidonie*. The German rendering, specifically the Rhine Franconian Fassung B manuscript, was made in the second half of the 15th century, before 1474 (Bertelsmeier-Kierst 2019:

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

25). Schneider (1961: 28) suggests that the translation dates from somewhere between 1455 and 1470. The author of this text is unknown. The edition that is used here is a diplomatic transcription made within the DFG project "Wortstellung und Diskursstruktur in der Frühen Neuzeit" (Bloom et al. 2024). The transcribed manuscript is Pontus und Sidonia Tandareis und Flordibel – BSB Cgm 577. This is the same handwriting that underlies the modernized transcription by Schneider (1961). The Dutch version dates from 1564 and was printed in Antwerp by Niclaes vanden Wouwere and is signed by Philips de Lens. The used edition is the diplomatic transcription by Kuiper & Brüggeman (2010[1564]). Sporadically, I will also refer to the first printed edition of the French version, which dates from 1479 in Genève (Besamusca & Willaert 2019: 75) and has been made available online by Gallica, the online library of the Bibliothèque nationale de France. The text in these three editions is what I refer to when I mention the German, Dutch, or French (FR) Ponthus.

Very briefly, the tale is about Ponthus, a very virtuous son of a king. We follow his adventures and the war of the Christians against the heathen kingdoms. He falls in love with Sidoine, but he faces many obstacles before he finally can be with her. The story is highly moralistic and focuses the virtues a good Christian should have, which are embodied by Ponthus and Sidoine. For a full summary of the plot see Classen (2016: 549-551). Note that his paper concerns a different adaptation of the text than the ones discussed in the current paper. Although the names of the protagonists are quite different, it concerns the same main events.

Importantly, the narrative focuses on Ponthus but includes different, smaller strands focusing on others. These strands do not serve immediately to progress the main story, i.e., the adventures of Ponthus, but happen parallel to it and aid the development of other characters. The story is multi-stranded and shifts perspectives by explicit transitions like *Nv laten wy te spreken vanden .xiiij. kinderen ende comen tot den ridder diese versonden hadde metten schepe ter zee waert* [DU, A4ra] and *NV lassen wir die xiiij kynder vff dem mere farn vnd sprechen vorbaß von dem Ritter der sie ernert vnd dar gefertiget hette* [GE, 5ra].[2] This can be seen in a diachronic perspective, as the shift from a single to multiple plot-lines is likely related to the development from the oral narrative tradition to written prose (Fludernik 2003, Zeman 2023).

Though the German and Dutch Ponthus both portray the same story, they are of different lengths and vary in the speed in which the story progresses. In particular, the used German adaptation has a drastically increased amount of direct

[2]A similar pattern to 'now we leave X and come to Y' is often found in transitions and is attested in the French print as well, e.g., *Si laisseray icy a parler des quatorze enfans / et retourneray au cheualier qui les auoit mis dedansla nef.* [FR, 11l].

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

discourse in comparison to other versions (Phinney 1962: 959).[3] Additionally, narrative sequences are embellished as well but to a lesser degree than the dialogues (Schneider 1961: 19-26). Therefore, the text is quite long and the story progresses rather slowly at times. In the Dutch adaptation, the extradiegetic comments of the narrator are manifold. They pause the progression of the story, often in order to explicitly address issues of morality.

3.2 Data collection

The first data set contains all PACs that are followed by *so(o)* or TEMP in preverbal position in the German and Dutch Ponthus, as for example (1a) and (1b) for German and in (2a) and (2b). They were manually extracted from the text. This data set contains 327 observations, 276 from German and 51 of Dutch. The second data set is an expansion of the former and contains all PACs that are preposed to clauses that are not introduced by a complementizer. That is, I do not consider constructions as in (8a), where the PAC is preposed to the complement clause to *globen* 'promise', because this complement clause is introduced by the complementizer *das*. Such clauses have their finite verb at the end of the clause.

(8) a. *Da must er ir globen wan yme got gehulffe Das er sins*
then must he her promise if/when him God help.SBJV that he his
vatter konigrich gewonne ***das*** *er dan wider zu ir kommen* ***solt***
father's kingdom win.SBJV **that** he then again to her come **should**
'Then he had to promise her that he would come back to her when God would have helped him to win his father's kingdom.' [GE, 85rb]

b. *vnd beschiet sie wan sie qwemen zu der statt so* ***solten***
and instructed them if/when they come.SBJV to the city so **should**
sie guden kauff geben
they goods sell give
'And instructed them that they should present goods to sell when they would come to the city.' [GE, 2rb]

I include cases in which a dependent clause does not contain a complementizer, such as the complement clause to *beschiet* in (8b), where the verb is early. In cases where it was unclear whether the initial clause was an adverbial or an independent main clause after consideration of the wider context, the sentence was not included in the data set. An example is presented in (9).

[3]Though, since it is uncertain which French manuscript the translator has used, it is possible that the additions were made initially by a French scribe.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(9) *hebben wi goet betrouwen op hem hi sal ons wel helpen*
have we good faith on him he will us DISC help
'Let us have good faith in him. He will help us.' *Alternatively,* 'If we have good faith in him, he will help us.' [DU, A3vb]

In total, the texts contained 784 PACs, 427 in the German and 357 in the Dutch Ponthus. The more restricted data set, with only the correlative patterns, contains 327 clauses, 276 German and 51 Dutch ones.

3.3 Annotation

The full data set was annotated for a number of variables, which will be presented here.

The response variable is the element (besides the adverbial clause) in the prefield (LP), which could be empty (–), the element *so* (*so*), a subject (*subj*), a temporal adverb/*da* (*temp*), or another constituent (*other*) in the full data set. Note that an empty prefield does not necessarily mean that the PAC is integrated. This also depends on the structure assumed for the host. For example, the prefield of imperatives is typically empty. Hence, if the PAC combines with one, is not likely to be syntactically integrated. This is exemplified in (10a). Nevertheless, it is possible to have a preverbal element with imperatives (e.g., Ebert et al. 1993: 432; Müller 2015: 1450); seeing that there is an optional slot, the PAC in (10a) *could* be argued to be integrated. Furthermore, a prefield can also not be generally assumed for Early New High German independent main clauses, since V1 declaratives were widespread (Ebert et al. 1993: 431f. Demske 2018: 145). Thus, one cannot be certain either that the PAC in (10b) is syntactically integrated. For this reason, I have simply encoded the adjacency of the PAC and finite verb instead of integration.

(10) a. *Jst das ir sie zornig sehent* ***machent*** *sie zu freden mit gutten*
is that you her spiteful see make her to peace with good
dogenden
deeds
'If you see her spiteful, make her content through good deeds.' [GE, 127vb]

b. *Da sie zwo is lange wyle angedreben druckten sie ir*
then.CNJ they two it long while continued pressed they their
augen
eyes
'When the two of them had continued it for a long time, they dried their tears.' [GE, 59vb]

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

The situation is different in 16th century Dutch which did not have V1 main clauses, and the Ponthus only contains one case in which the PAC is directly adjacent to a finite verb in the imperative. This is exemplified in (11a). Note that this is represented with a similar construction in the German text (10a) and in the French print (11b). As such, the adjacency of the PAC and the finite verb may be due to transfer. Therefore, there is no direct evidence in the Dutch Ponthus of potentially unintegrated PACs when PAC and finite verb are adjacent.

(11) a. *Jst dat ghise te onureden siet* ***steltse*** *te vreden met duechdelike woorden*
is.it that you.her to unhappiness sees make.her to peace with virtuous words
'In case you see her unhappy, make her content through virtuous words.' [DU, n4ra]

b. *et se vous la voyez courroucee* ***appaisez*** *la par courtoysie*
and so you her see furious appease her through courtesy
'And if you see her furious, appease her through courtesy.' [FR, 119r]

In sum, the direct adjacency of the PAC and the finite verb of the host is thought to be reflective of the integration of the PAC into the host sentence in the Dutch Ponthus but not in the German text.

3.3.1 Who speaks?

The operationalization of mimesis and diegesis is encoded by annotating whether the utterance was spoken by a character (*char*) or by the narrator (*narr*). In the cases that that the narrator comments are extradiegetic, this was reflected in the label (*narr_ex*). Since this is infrequent, this falls under the narrator speaking for the statistical analysis.[4] (See §2.1.)

There are a couple of sentences in the data set for which the categorization was not as clear-cut. Especially the boundary between indirect speech, direct speech and extradiegesis is not always easily distinguishable. I have here chosen for the most likely interpretation given the context. In one case, there seems to be a shift from character to narrator. This sentence is presented below (12) and full glosses and a translation are given in the Appendix. The segment starts out with indirect speech. This is identifiable by third person pronoun *sie* with the plural

[4]In total, only six sentences were extradiegetic, five of which were found in the Dutch text with a subject following. Since they were few, these cases will not be further discussed separately in the paper.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

verb *wolten*, which indicates that *sie* refers to the speakers. Yet, the verb mood shifts from subjunctive to indicative in the last sentence. This suggests that the independent clause at the end of this segment is direct or free indirect speech (Demske 2019). The present tense of this final sentence (*ist*) is also reflected in the PAC by *kompt*, but the main clause has a tense shift to preterite (*wart*). This is indicative of at least a shift of the deictic origo from the story-now to the narration-now. This probably coincides with a shift in voice, from the characters' to the narrator's.

(12) *Sie sprachen da were vil abe zu sagen Dan siner schonheit weydelicheit vnd sins gutten gelaß were nit me in der werlt als sie wol gleuben wolten williche frauwe oder jungffrauwe yne zu bolen mochte haben da ist kein rede in Sie mocht wol berummen das sie den schonsten den besten den wol kommesten vnd den aller lustigesten frunt hette der da lebt vnd schon ist sin jugent* ***kompt er also inn das alter so wart sin gliche nie me ersehen*** *das ist vngezwiuelt die rechte warheit* [GE, 16rb]

This particular instance is removed from the further analysis.

In the vast majority of the cases, the context and the verbal morphology indicated clearly whether we hear the narrator's or a character's voice.

3.3.2 Narrative speed

The data was furthermore coded for narrative speed (NARR.SP). In dialogues, the story time and narration time are taken to coincide, and we speak of isochrony (*iso*). Segments of indirect speech where it is likely that the character's speech is presented as elaborate as if it were direct speech are annotated as isochronous as well. When narration time is slower than story time, the sentence is classified as a summary (*sum*), and pauses in the story time while the discourse is continued, both diegetically and extradiegetically, are annotated as *pause*. (See §2.2.)

In some cases, one sentence pertains to more than one level of narration. If the PAC and the host differ in this regard, the sentence is labeled *shift*, as in (13a). The PAC summarizes the events of the story, but the host sentence is an extradiegetic comment by the narrator, thus pausing the narration of the story.

(13) a. *Als sidonie dit hoorde men derf niet vragen of si ontstelt*
when Sidoine this heard one dares not ask if she distraught
wert van herten
was of heart
'When Sidoine heared this, one dares not ask if she was sick at heart.'

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

[DU, e1va]

b. *Doen pontus den sarazijn versleghen hadde ende dat hooft den*
then.CNJ Ponthus the Saracen defeated had and the head the
Eedelingen gegeuen hadde om dat sijt haren conick presenteren
noblemen given had to that they.it their king present
*souden **als ghy nu ghehoort hebt** sloech hi sijn paert met sporen*
should as you now heard have hit he his horse with spurs
'When Ponthus had defeated the Saracen and had given the head to the noblemen to present it to their king, **as you have now heard**, he spurred his horse on.' [DU, c3rb-c3va]

Whenever another embedded clause reflects a similar shift, this is ignored. For example, the *als*-clause in (13b), pauses the narrative and the narrator addresses his audience directly. The remainder of the utterance is a narrative summary, continuing the telling of the story. As the *als*-clause is a dependent clause within the preposed adverbial, the change is not annotated.

(14) *gebe er yme alle die werlt er neme ir nit vor sin dochter*
gives.SBJV he him al the world he take.SBJV he not for his daughter
'Even if he would give him the entire world, he wouldn't take it over his daughter.' [GE, 114vb]

As mentioned in §2.2, segments of indirect speech can be classified as (nearly) isochronous when the indirect speech is a close reflection of the words of a character, as for example in (14). Whenever the narrator summarizes what would have been a more elaborate monologue or dialogue, the segment is considered to be a summary.

3.3.3 Type of PAC

Finally, the type of PAC was coded. This is approximated by the element that introduces the PAC, conjunction or adverb, or the verb-initial position in the case of unintroduced PACs. This variable is encoded in order to see whether the above discussed narrative functions are reflected in or can be reduced to more local relations between PAC and host-sentence.

3.4 Statistical methods

The p-values and residuals resulting from Pearson's χ^2-test will be reported for simple frequency table. When the amount of data was insufficient, the p-values from Fisher's exact test will be reported instead.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

To rank the importance of the three annotated variables, "who speaks", narrative speed, and the type of PAC, I use random forests and variable importance measures. For illustrative purpose, I present conditional inference trees. The "party-package" and the functions CFOREST, VARIMP and CTREE are used to model these (Hothorn et al. 2006, Strobl et al. 2007, 2008). All modeling was done in R (R Core Team 2023).

4 Adverbial V3 with *so(o)* and TEMP

Why would *da* and *so* be associated with mimesis vs. diegesis? In Present-Day German, temporal correlatives (*dann, da*) are associated with the spoken domain, and resumption with *so* is restricted to the written, formal register (Catasso 2021, Axel-Tober 2023). Working with assumption that segments of direct discourse are more reflective of the spoken language, since the language of (fictional) persons is imitated, and that segments of diegesis are representative of more formal language (Oksefjell Ebeling & Ebeling 2020), one would expect *so* to be more dominant in diegetic segments of the narration and TEMP in mimetic parts. If this Present-Day German distinction has historical foundation, we would expect this to be the case for earlier stages of German as well.

Table 1: Who speaks and *so* & TEMP in the German Ponthus

	character		narrator		
	raw	(χ^2-residuals)	raw	(χ^2-residuals)	total
so	78	(7.69)	33	(-5.05)	111
TEMP	5	(-6.32)	159	(4.16)	164
total	83		192		275

It turns out that the opposite is true, as can be seen in Table 1. The data indicate that there is a distinction between *so* and *da* in the prefield, which is significant with $p < 0.001$. The residuals in Table 1 show that each cell has an effect, as the residual is outside of the range -1.96 to 1.96, meaning that the specific combination either occurs less or more than expected. Thus, it can be concluded that *so* occurs considerably more often than expected in character spoken segments and less often when the narrator speaks in the German Ponthus. The results indicate that the opposite is true for TEMP, which occurs more often than expected with the narrator's voice. What should be additionally noted is that all except one of the TEMPs is *da/do*, exemplified in (15).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(15) *Do* *er* *lange in dem walde geging sich* *zu bedencken in diesen sachen*
then.CNJ he long in the forest went REFL to ponder in these things
die *yme wieder faren* *weren* ***zu lest*** *fant* *er* *einen schonen*
that him again happened were **to latest** found he a beautiful
anslag
placard
'When he had went into the forest for a long time to ponder about these things that happened to him, finally he found a beautiful placard.'
[GE, 38va-38vb]

There is thus not much variation with regard to the types of temporal elements that can occur preverbally after PACs. In other words, it is not a productive schema but a lexically specific construction.

The findings for German thus point in the opposite direct than what was expected: Contrary to the hypothesis, *so* is associated with mimetic segments and *da* with non-mimetic ones. What about Dutch? First of all, Dutch did not have one adverb that is clearly recruited as a temporal resumptive. The following temporal adverbs are attested following an PAC in the Dutch Ponthus: *doen* 'then', *nemmermeer* 'nevermore', *rechtevoort* 'forthwidth', and *terstont* 'immediately'. Second, the overall frequency of TEMPs following a PAC is considerably lower in comparison to *so* than in German (3.08% vs. 38.41% of all PACs).

Table 2: The distribution of *so(o)* and TEMP wrt. voice

	so(o)	TEMP	total
German: character	93.98%	6.02%	83
German: narrator	17.19%	82.81%	192
Dutch: character	93.55%	6.45%	31
Dutch: narrator	55.00%	45.00%	20

As the frequency for the Dutch data is low, a Fisher's Exact test verifies that the voice distinction is significant (Dutch: $p = 0.002$). That an overwhelming amount of the PACs are followed by *so(o)* and not by TEMP in direct speech is also true for Dutch, as can be seen in Table 2. Where Dutch diverges from German is that it does not show a strong preference for TEMP in narrator text.

The results point in one of two directions: Either correlatives were not in the same way associated with different registers as in Present-Day German, or the

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

dialogues in the Ponthus are not reflective of spoken language (see for a similar point on direct speech in Old English Louviot 2016). Testing this is left for other studies. In any case, the data suggest that the two segments of discourse are distinctive so that the realization of the prefield shows different preferences. The unexpectedness of the results require an explanation. As mentioned earlier, direct discourse has been traditionally considered to be the temporal baseline, in which narrative and story time are taken to be isochronous. As such, when a character speaks, the story develops steadily but rather slowly. Differently, when the narrator speaks, the progression of the events in the story may be halted, go faster than the discourse, or there may be a shift between the PAC and the following clause. Due to this connection between 'who speaks' and narrative speed, it can be hypothesized that the observed difference in Table 2 is in fact an artifact of narrative speed. *so* would then be associated with isochrony or an overall slower pace of narration, whereas temporal adverbs might be associated with segments in which the evens of the story unfold comparatively rapidly in German, or with no particular narrative speed in Dutch.

If this is the case, the occurrence of *so* in segments of the story that are spoken by the narrator should be of a slower pace (i.e., pauses) than temporal adverbs, which should be found within narrative summaries. This seems indeed to be the case. (16a) illustrates one of the examples of *so* in the narrator voice and (16b) of *da* in a segment of direct speech.

(16) a. Context: No-one ever hears him vow more than "It is like that" in good faith.
vnd so er dure sweren solt so sprach er (...)
and so he dearly swear should SO sprach er (...)
'And whenever he had to vow dearly, he said (...).' [GE, 9va]

b. *Vnd do ich gewar wart das er erslagen was Da det ich*
and then.CNJ I aware became that he defeated was TEMP did I
myner jungffrauwen rock einen an
my lady's tunic one on
'And when I became aware that he had been defeated, I put on one of my lady's tunics.' [GE, 106rb]

The segment in (16a) is a rather straightforward case of description: The narrator describes Ponthus and all his positive traits. In such segments, the temporal progression of the story is halted.

In (16b), the queen is speaking to Ponthus and the sequence PAC + *da* is used. In this segment, which comes relatively late in the narration (106rb), both Ponthus

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

and the reader learn about what has happened to the queen when the city was invaded and she was separated from Ponthus, which happened at the beginning of the text (2vb). The segment has a retrospective function and summarizes the events from a different perspective, thereby progressing the story rapidly.

The question is now whether this observation can be statistically supported. Table 3 presents the proportions of the two patterns with the different velocities.

Table 3: *So* and TEMP and narrative speed in narrator speech

	so(o)	TEMP	total
German: pause	100%	0%	6
German: isochrony	94.74%	5.26%	19
German: summary	5.06%	94.94%	158
German: shift	11.11%	88.89%	9
Dutch: pause	0%	100%	1
Dutch: isochrony	100%	0%	2
Dutch: shift	100%	0%	2
Dutch: summary	46.67%	53.33%	15

For German, the prediction is indeed borne out. Albeit with a low frequency of pauses and shifts ($p < 0.001$), *so* occurs more than expected with a slower pace (isochrony & pause) than TEMP. Since the Dutch PACs with *so* or a temporal adverb only contains one instance of a pause and two of a shift and isochrony, no conclusion can be drawn.

These results indicate that, at least for the German data, narrative speed is a better explanation for the distinction between *so* and TEMP than voice: Parts of the narrative that are voiced by the narrator occur with *so* when they progress the story slowly (isochrony and pause).

Finally, the type of PAC is correlated with the element that resumes it. That is, certain PACs seem to be associated with *so*, whereas others have a tendency to be followed by *da*. PACs that are introduced by the typically temporal but also possibly causal conjunction *da* are also combined with *da* and not by *so*, see (17a). Differently, V1-clauses, which typically express conditionality, are almost always followed by *so*, as in (17b).

(17) a. *Da messe gesongen was Da ging man in den sale*
then.CNJ mass sung was TEMP went one in the hall
'When mass had been sung, then they went into the hall.' [GE, 51rb]

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

b. *wolt ir yne sehen so wil ich yne uch bringen*
wanted you him see so want I him you bring
'If you wanted to see him, I will bring him to you.' [GE, 11ra]

For Dutch, there are a few types of PACs that occur more than once in the smaller data set: *als, doen, nadien (dat), indien* or V1. Each of those occur *so*, that is, there is no clause type that is restricted to a correlative structure with TEMP.[5] *Als*-clauses occur more often than other clauses with a following temporal adverb (39.13%, n = 23), all other PACs have an even higher proportion of *so* as compared to temporal adverbs (88.89% for V1-clauses, and 100% for *nadien (dat), indien* and *doen*).

To evaluate which variable has the most impact on the realization of the correlative in German, a random forest model was fitted on the data. Due to their low frequency, the categories *pause* and *shift* of narrative speed are left out of the quantitative analysis. The random forest calculates 2000 trees (ntree = 2000) and selects two variable for each tree (mtry = 2). The random forest has an accuracy of 98.46%, and a C-value of 0.995 and Somers' D 0.990, which means that the model performs outstandingly.

Presented in Figure 2 are the results of the conditional variable importance measures for German. They show that TYPE.PAC, i.e., the type of PAC is the most important factor whether a temporal adverb or *so* is found in the prefield. Narrative speed, NARR.SP is relatively important as well, but WHO.SPEAKS is not identified as an important factor. In other words, this confirms what was noted earlier in this section, that narrative speed has more explanatory value than voice. Importantly, the effect of narrative speed does not disappear when accounting for the type of PAC, suggesting that the local discourse function and the wider narrative function of the construction are both contributing factors.

Let us take a look at the data for German by looking at a conditional inference tree, which is visualized in Figure 3. The first split in the data concerns the variable TYPE.S, with PACs introduced by *als*[6] 'when, while, as', *als bald* 'as soon', *da* 'then', *nyt lange* 'not long' and *so bald* 'as soon as' being more often temporally resumed than the other types of clauses, which occur more often with *so*. What each of these clauses have in common is that they prototypically express a temporal relation, in particular one of simultaneity or rapid succession. This is illustrated in (18)

[5]The conjunction *ten* occurs once with *nimmermeer* 'nevermore' but not with *so(o)*: *ten sy dat ghi op dese tijt uwen sin inden mijnen set nimmermeer en sal ic v lief hebben* 'if it is that you now your desires in mine put, I will never love you.' [DU, j1vb].

[6]It should be noted that *als* is, unexpectedly, rather infrequent in the Ponthus for a thus-far unknown reason.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

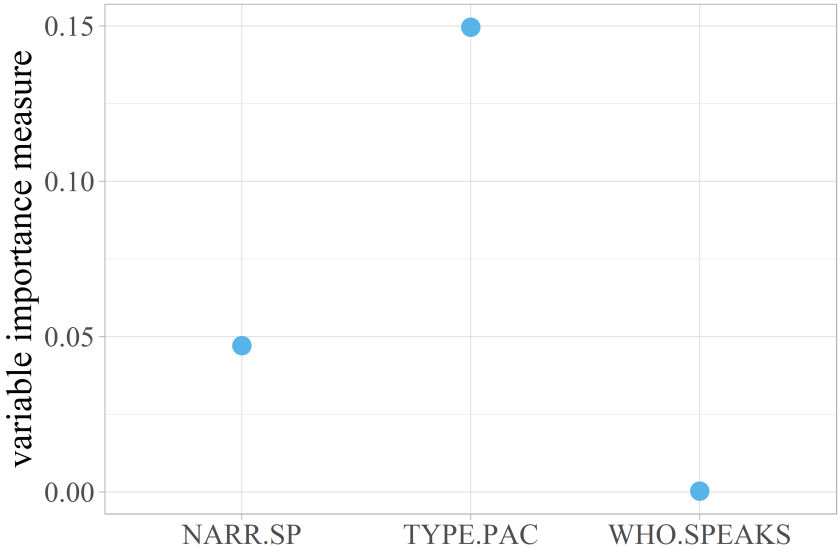


Figure 2: Conditional variable importance measures for German

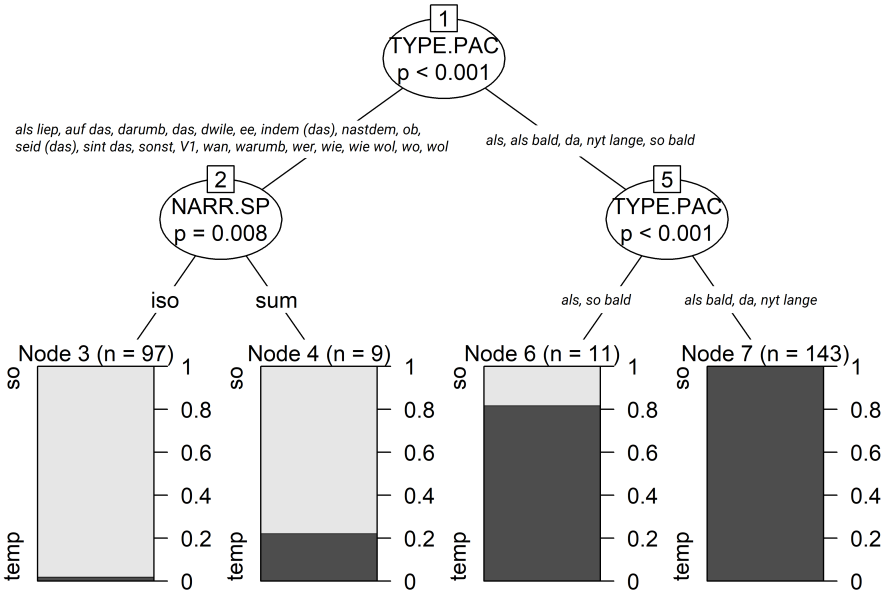


Figure 3: Conditional inference tree: TEMP versus *so* in the German Ponthus

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(18) *Vnd so balde das geschach da kanten sie yne von stont wol*
and as soon that happened TEMP knew they him of stand well
'And as soon has that happened, they recognized him immediately.'
[GE, 87va]

Within these temporal adverbial clauses of PACs, there is also a significant difference between those introduced by *als* and *so bald* versus those introduced by *als bald, da* and *nyt lange*. The former two occasionally occur with a following *so*, whereas this is not attested in the Ponthus for the latter group. Note that one should be rather careful with generalizing these findings: They are based on one text and thus we cannot conclude that *so* was categorically excluded from this group in other texts as well. Moreover, while the conditional inference tree nicely presents the optimal binning for the current data set, whether this is a robust result that can be extrapolated remains to be seen. Nevertheless, the marginal occurrence with *so* with conjunctions with a temporal meaning falls in line with what is known about these conjunctions in Dutch and in English.

Within the group of conjunctions that are associated with *so*, a significant distinction is made regarding narrative speed with isochrony having a significant higher proportion of *so* in comparison to temporal adverbs than summary. In other words, the types of PACs that are typically followed by *so* can be taken up by a temporal adverb instead when they progress the story rapidly.

A reviewer suggested that verbal tense might play a role in this relation between *da* and narrative speed. I acknowledge that this is very much probable, but the relation between the two is not as straightforward as one might suspect. The examples thus far have suggested that a pluperfect in the is typically combined with *da* and occurs in narrative summary. This is not exclusively the case, as the examples in (19) illustrate: In (19a) it combines with *so* and in (19b) with juxtaposition.

(19) a. *Aber dwile is* ***begonnen were*** ***so*** *wolt er is vß herten was*
but while it started was.SBJV SO wanted he it out harden what
da von qweme bose oder gut
there of came.SBJV evil or good
'But since he had started it, he wanted to tough it out, whatever would have come of it, evil or good.' [GE, 112va]

b. *Da sie nu* ***gescheyden waren Sydonie*** *begunde mit irn*
then.CNJ they now separated were Sidoine started with her

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

jungffern zu reden wie yne pontus gefiel
lady to talk how her Ponthus pleased
'When they had separated, Sidoine started to talk with her lady about how she liked Ponthus.' [GE, 16rb]

Yet, all adverbial clauses with a pluperfect verb tense that combine with *so* have subjunctive verbal mood, as in (19a). This is so for pluperfects with juxtaposition as well. In the few cases in which the verb is indicative (e.g., (19b)), the sentence presents a narrative summary. The indicative pluperfect indeed appears to be associated with *da* and segments of narrative summary. Due to the inconsistency in spelling, which results in a conflation of the third person singular present and past tense of *haben* 'have' in combination with the interaction of tense with mood, a more systematic analysis of the impact of tense goes beyond the scope of this paper and will be left for further research.

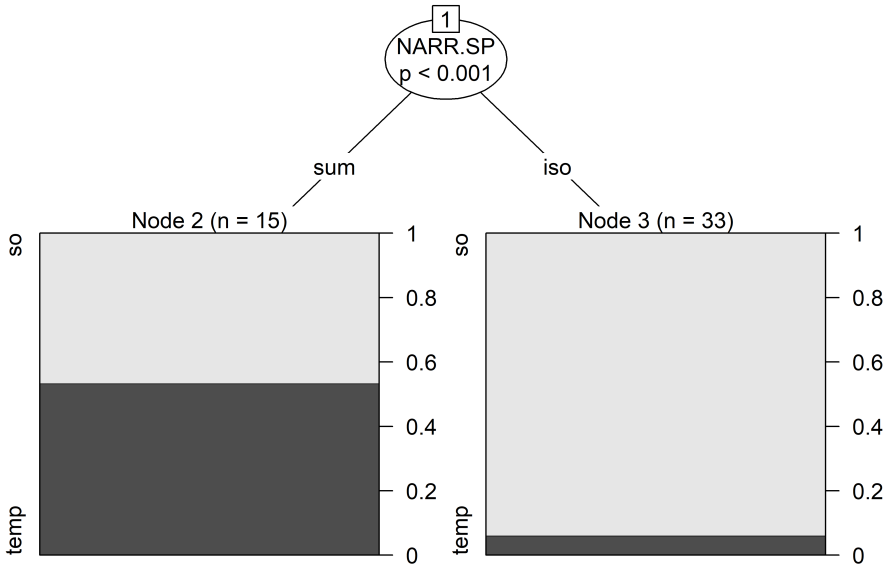


Figure 4: Conditional inference tree: TEMP versus *so* in the Dutch Ponthus

Turning now to Dutch, the conditional inference tree presented in Figure 4 shows, curiously, no identifiable distinction regarding the type of PAC. Like German, the model identifies a significant split in the data by narrative speed, with

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

isochrony showing a higher proportion of *so* than summaries, which have a relatively high amount of TEMP.

This section has illustrated that there is a systematic distinction between PACs that are followed by *so* versus those that are followed by TEMP: The former are associated with narrative segments in the character's voice and the latter with the voice of the narrator. This was argued to be a reflection not of narrative style but of velocity, with PAC + TEMP being indicative of a more rapid story progression. More specifically, while much of the variation between the two constructions in German can be explained by the type of PAC, narrative speed is shown to be an important factor as well. In Dutch, the type of is not significant, and only narrative speed comes up as a significant variable. It should be noted that this may be due to the low overall frequency of the two constructions, in particular of PACs with TEMP as well as the relatively low frequency of narrative summaries. Two things should now be explained: Seeing that the Dutch Ponthus is not significantly longer than the German one and presents the same story, it is safe to assume that the amount of narrative summary overall should not be less in the Dutch text than in the German. This raises the question whether this narrative structuring function is associated with other realizations of PACs, and whether this could simultaneously explain why PACs with TEMP are so rare in Dutch.

In addition, it must be checked whether the type of PAC has really no effect on the realization of the prefield in the Dutch Ponthus, whether this was an effect of low overall frequency of PAC + *so* and PAC + TEMP, or whether it plays a role after all but with different constructions. The following section will address these issues.

5 PACs in the German and Dutch Ponthus

The previous section identified considerable differences between the German and Dutch Ponthus. In particular, the Dutch data set contained few cases of narrative summary with TEMP in comparison to the German one and overall did not present as many correlatives. The question rises then whether the narrative summary is expressed by a different construction in Dutch. In addition, the noteworthy absence of the effect of type of PAC on the realization in Dutch requires an explanation. I will therefore address these issues here, making use of the larger data set containing all PACs.

To recap, following the PAC, the options are relatively limited. In German, the majority of the constructions belong to one of the following four types: 1) direct adjacency between PAC and finite verb (20a), 2) PAC + TEMP, predominantly *da*

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(20b), 3) PAC + *so* (20c), or 4) a subject intervenes between PAC and the finite verb (20d). On occasion, other elements – mainly adverbs and objects – stand in between PAC and the finite verb (20e).

(20) a. *Do is tag wart stont er vff*
then.CNJ it day became stood he on
'When the day came, he stood up.' [GE, 107ra]

b. *Da er das hortt* ***da*** *erschracke ere[sic!] sere*
then.CNJ he that heard TEMP shocked he much
'When he heard that, he was very frightened.' [GE, 111va]

c. *Aber dwile is begonnen were* ***so*** *wolt er is vß herten*
but while it started was.SBJV SO wanted he it out hard
'But because it had been started, he wanted to see it through.' [GE, 112va]

d. *lebet der lange* ***er*** *dott mir alle myn folk*
lives.SBJV he long he kills me all my folk
'If he lives longer, he will kill all my people.' [GE, 72vb]

e. *Ob ir yne dan sehent* ***was*** *wolten ir thun*
if you him then see what would you do
'If you see him then, what would you do?' [GE, 87va]

Table 4 presents the distribution of the types of elements that occur in between the PAC and the finite verb: *so* and TEMP are known from the previous section; the hyphen (-) stands for direct adjacency of the two; *subject* stands, obviously, for the subject; and *other* refers to all other elements that may occupy the position between the PAC and the finite verb, which predominantly are other adverbs (e.g., *darzu/daer bi* 'therewith') and object pronouns (e.g., *das/dat* 'that').

Table 4: Preverbal elements that co-occur with PACs

	-	other	*so*	subject	TEMP	total
German	31	18	112	102	164	427
Dutch	211	25	40	70	11	357
total	242	43	152	172	175	784

The languages are significantly different ($p < 0.001$). Residuals show that adjacency of PAC and finite verb has a higher than expected frequency in Dutch

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

and is less than expected in German (9.60 and -8.78 respectively). The reverse is true for *so* in the prefield (-3.51 vs. 3.21) and especially for TEMP (-7.69 and 7.04): These patterns are infrequent following Dutch PACs but frequent in German. It thus seems that where German uses PACs with TEMP or *so*, we find integration of the PAC in Dutch.

In the following section, this frequency difference is related to the two factors that were discussed in the previous section as well. In §5.1, narrative speed is considered and I will additionally take into account the type of PAC in §5.2.

5.1 Narrative speed

Figures 5 and 6 present the distribution of the different prefield fillings in relation to the narrative speed in the German and Dutch Ponthus respectively. What should be kept in mind while reading these figures is that the Dutch data contains less isochrony (29.69%) and more narrative summary (62.75%) in comparison to the German data (43.09% and 50.12% respectively). This is in line with the expectation, as both texts present the same story yet the German Ponthus is ca. 20000 words longer.

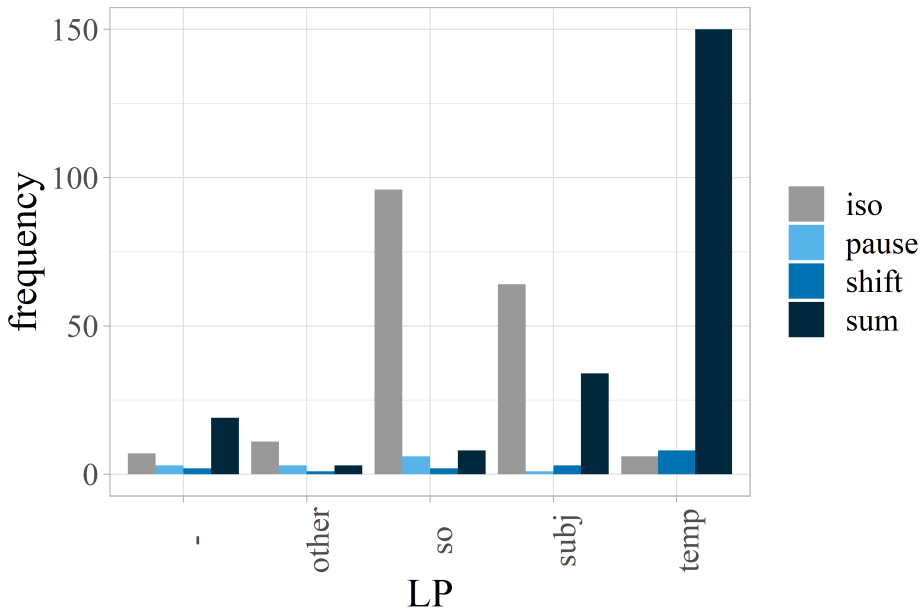


Figure 5: Narrative speed and the prefield in the German Ponthus

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

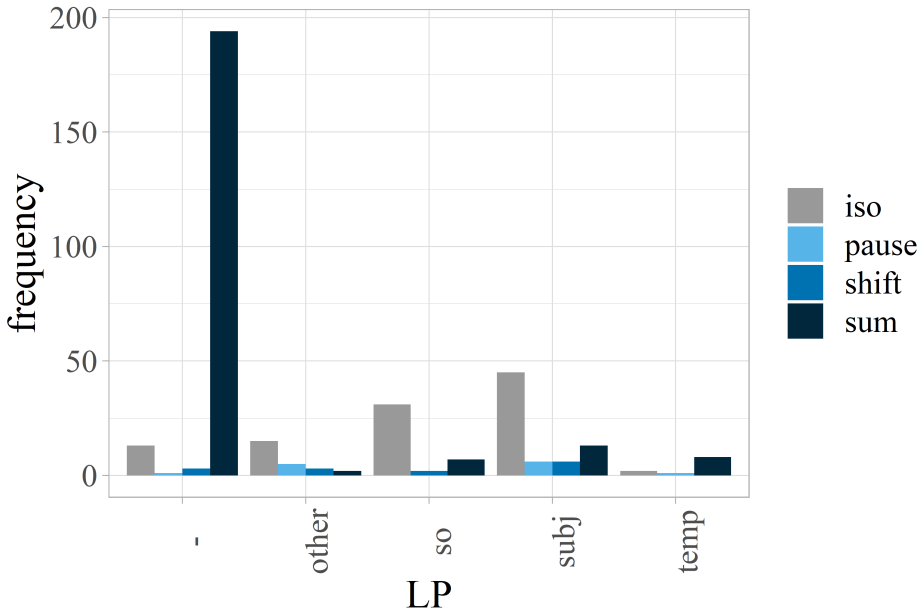


Figure 6: Narrative speed and the prefield in the Dutch Ponthus

Figure 5 shows that adjacency of PAC and finite verb is overall infrequent but occurs most frequently in narrative summaries. Narrative summaries are particularly frequent with TEMP, and isochrony with *so*. Subjects in the prefield occur mainly with isochrony (62.75%) though narrative summary is also frequent (33.33%). This suggests that PACs that occur with a preverbal subject are not strongly associated with a particular narrative speed. Other elements are overall infrequent and occur most often in isochronous segments (61.11%).

In the Dutch Ponthus, the situation is different. Figure 6 makes it clear that the function of narrative summary is dominant with the adjacency of the PAC and the finite verb (91.94%). This explains why we saw so little TEMP in the prefield in Dutch: The narrative function is expressed by integrating PAC. Moreover, the frequency of *so* is overall lower than in German; nevertheless, the strong preference for isochrony is likewise there (77.50%). A relatively high frequency of isochrony is found with subjects in the prefield (64.29%). In German we also saw this preference, but the construction occurred also relatively frequent with narrative summaries. In Dutch, subjects in the prefield are more rare with summaries (18.57% versus 33.33% in German). Other elements occur overall very infrequently, but they are attested with all velocities.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

In sum, the major difference between the two languages lies in narrative summaries and the construction that is associated with it: In the German Ponthus, this function is expressed by PACs with TEMP – specifically *da* – in the prefield; in the Dutch text, the PAC is directly adjacent to the finite verb in these contexts, viz. integrated in the host sentence.

5.2 Type of PAC

Let us now consider the types of PACs. In this section, I will first discuss the situation in the German Ponthus and subsequently that in the Dutch text. Thereafter, I will compare the two and discuss some implications.

Not all types of PACs show variation regarding the type of construction they occur in. Of course, those that only occur once in the data set exhibit no variation. These are *als liep*'as dear', *sint das* 'since', *sonst* 'otherwise if', *warumb* 'why', and *wol* 'although', which occur once with *so* ; *nyt lange* 'not long', which is followed TEMP; and *wie stark* 'how strong', which has a following subject. *Auf das* 'so that' (2) and *wie wol* 'although' (5) are furthermore only attested with *so* in the German Ponthus. The other types of PACs are attested in more than one of the constructions.

The most frequent types of PACs that occur ten times or more in the German Ponthus are *da* 'then', V1, *wan* 'if, when', *so bald* 'as soon as', *ob* 'if', *wo* 'where, when, if' and *ee* 'before'. *Da, so bald* and *ee* typically introduce temporal adverbial clauses, whereas *ob-, wo-* and V1-clauses mainly express conditionality. *Wan*-clauses are often ambiguous between the two readings. They all occur in more than one construction. Considering both the type of PAC and narrative speed, I will first illustrate by means of conditional inference trees the trends that are shown in the data. Then I will use random forests to predict the realization of the prefield. For this model, the input data only contains the most frequent types of PACs, which are *da,* V1, *wan, so bald, ob, ee,* and *wo* for German.

Figure 7 presents a conditional inference tree for the German Ponthus. The first split the model makes is between *da* and *so bald* on the one hand, and *ee, ob,* V1, *wan,* and *wo* on the other (node 1). In the latter group (node 7), *wo*-clauses are separated from the other, because they occur less often followed by SUBJ and more often by TEMP than the others. The group with *da* and *so bald* has a comparatively high proportion of TEMP and –. Within this group, the model identifies a separation in the data regarding narrative speed (node 2): Clauses that are introduced by *da* or *so bald* that encode isochronous narrative progression are found with *so* ; when they encode pauses, shifts and summaries, they occur more often with TEMP and –. Specifically, *da* and *so bald* clauses that encode

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

summaries may occur with a following subject and more often occur directly adjacent to the finite verb than the pauses and shifts (node 4). Still, TEMP is the preferred pattern for both.

In Dutch, the included types of PACs in the model were only *als,* V1, *doen* and *nadien (dat)*, because they were the only ones that occurred with a frequency of ten or more (and showed variation). The conditional inference tree can be found in Figure 8. The first split (node 1) in the data that the model identifies is between V1-clauses and those that are introduced by *als, doen* and *nadien (dat).* The V1-clauses have a relatively low proportion of integration, a high amount of SUBJ, and a slightly higher proportion of *so* following than the other clauses. Differently than what we have seen in Figure 4, narrative speed is relevant here for the relation of the prefield: In isochrony, *so* is more frequent than in pauses and shifts (node 2). Finally, summaries are realized differently than pauses and shifts as they have a higher proportion of integration and less *so*, SUBJ and TEMP.

The findings from these models is that narrative speed affects the realization of the prefield in both languages within the subset of temporal adverbial clauses. Where narrative summaries are encoded by PAC + TEMP in German, PACs occur directly adjacent to the finite verb in Dutch. To evaluate whether these findings are stable, I ran two random forest models; one for each language. 2000 trees are calculated (ntree = 2000), each selecting one variable (mtry = 1).

Table 5 presents the accuracy matrices resulting of the models. They present the observed distribution and the distribution as predicted by the model.

Overall, the German model erroneously predicts that there are no cases in which the PAC and the finite verb are directly adjacent (–), nor any cases of subjects in the prefield. It accurately predicts all instances of *si*, and it only erroneously classifies two cases of TEMP as *so*. The model thus overestimates the use of the correlative patterns. Interestingly, 64.29% of the PACs adjacent to the finite verb are classified as TEMP, whereas 65.93% of SUBJ is classified as *so*.

The Dutch model does not predict any TEMP and slightly over-predicts integration. It accurately predicts 94.61% of integration, which drops to 71.70% for juxtaposition with a subject in the prefield, and to a meager 48.39% for *so*. TEMP is mainly predicted to be integrated (90.00%). This is the case for the wrongly classified SUBJ as well (22.64% in total, 80.00% of wrong predictions). *so* is more often wrongly predicted to be SUBJ than integrated, but this difference is small (16 vs. 15 wrong predictions).

These matrices provide insight as to how accurately the random forests predict the element following the PAC based on two factors: narrative speed and the type of PAC. For German, the model reliably identifies *so* and TEMP, but it does not identify – and SUBJ. Differently, in Dutch TEMP is not predicted at all. This

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

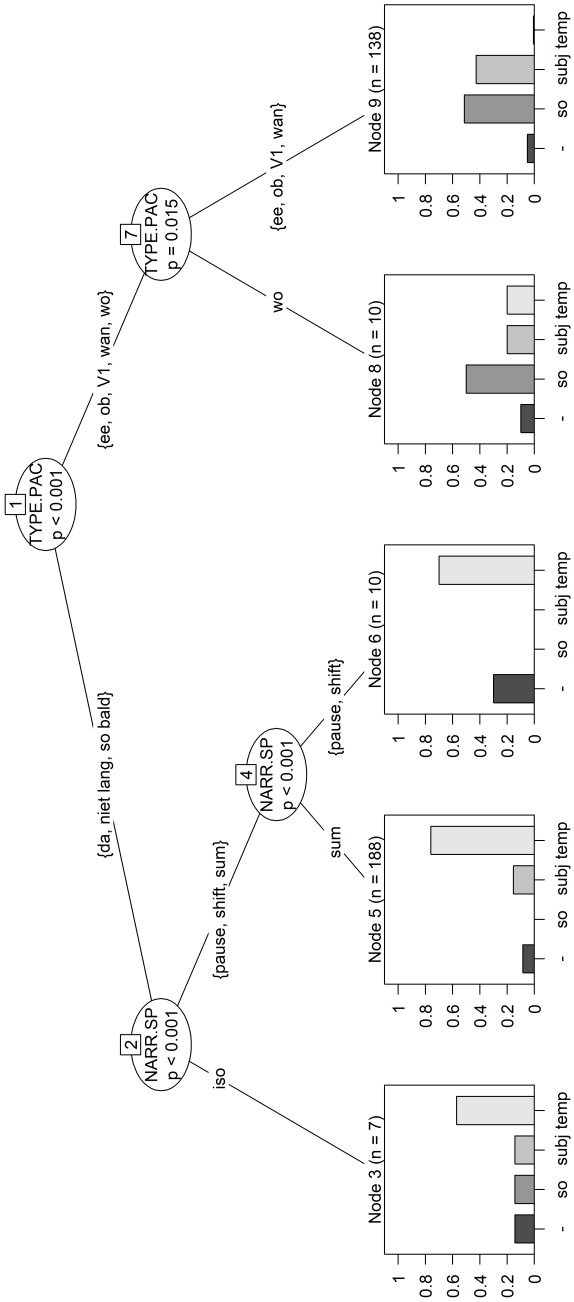


Figure 7: Conditional inference tree: Prefield filling in the German Pon-thus

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

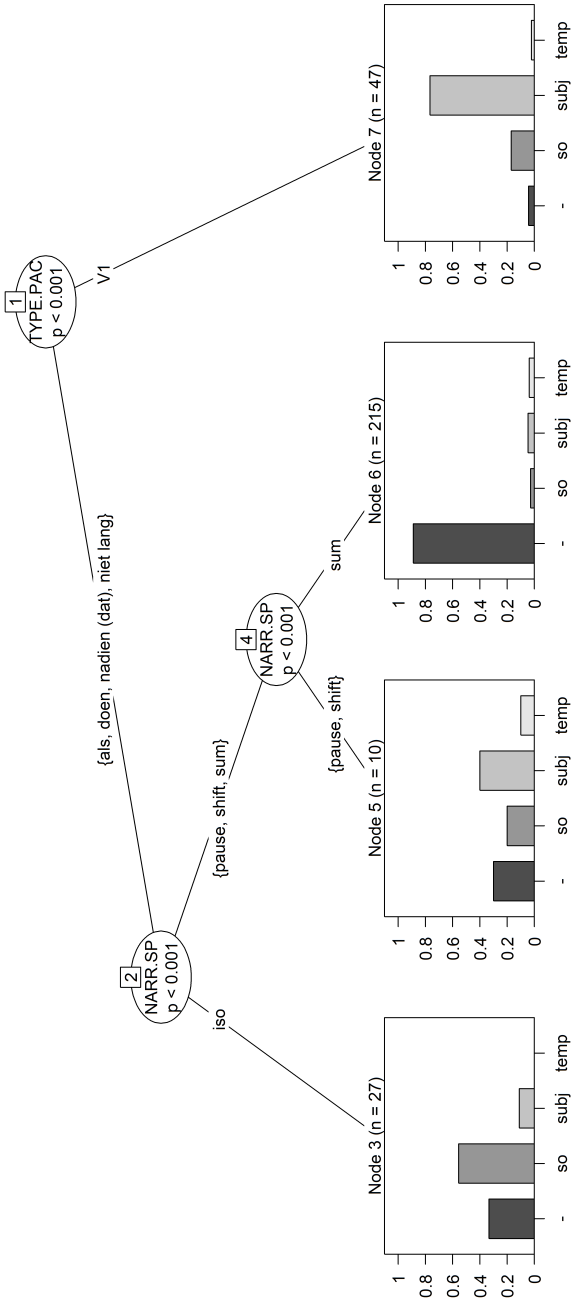


Figure 8: Conditional inference tree: Prefield filling in the Dutch Ponthus

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Table 5: Random forest accuracy matrices

(a) German

	–	*so*	subject	TEMP	observed
–	**0**	10	0	18	28
SO	0	77	0	0	77
subject	0	60	**0**	31	91
TEMP	0	2	0	**154**	156
predicted	0	150	0	202	

(b) Dutch

	–	*so*	subject	TEMP	observed
–	**193**	8	3	0	204
SO	7	**15**	9	0	31
subject	12	3	**38**	0	53
TEMP	9	0	1	**0**	10
predicted	221	26	51	0	

suggests that the factors that motivated the use of TEMP in German do not do so in Dutch. Instead, integration is predicted with a high accuracy, which is fully in line with the hypothesis that where the German Ponthus uses PAC + TEMP, Dutch prefers integrated PACs. Moreover, juxtaposition seems to be associated with particular settings in Dutch (the tree in Figure 8 suggest V1 and pauses/shifts if not V1) but not in German. This can reflect that juxtaposition has developed its own niche now that integration has become the default in Dutch but that German has not (yet) fully gone through this process.

5.3 Returning to Dutch: integration versus *so*

In the previous section, it was argued that where the German Ponthus uses *da*-resumption, Dutch tends to use the integration of PACs. This means that we need to revise the findings for Dutch in §4 for a complete picture, as it is not TEMP that contrasts with *so* but integration. For this reason, I present one final conditional inference tree, which parallels the German one presented in Figure 3. Here, all types of PACs are included.

Figure 9 illustrates that the two constructions contrast regarding narrative speed, with *so* being considerably more frequent with isochrony. With

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

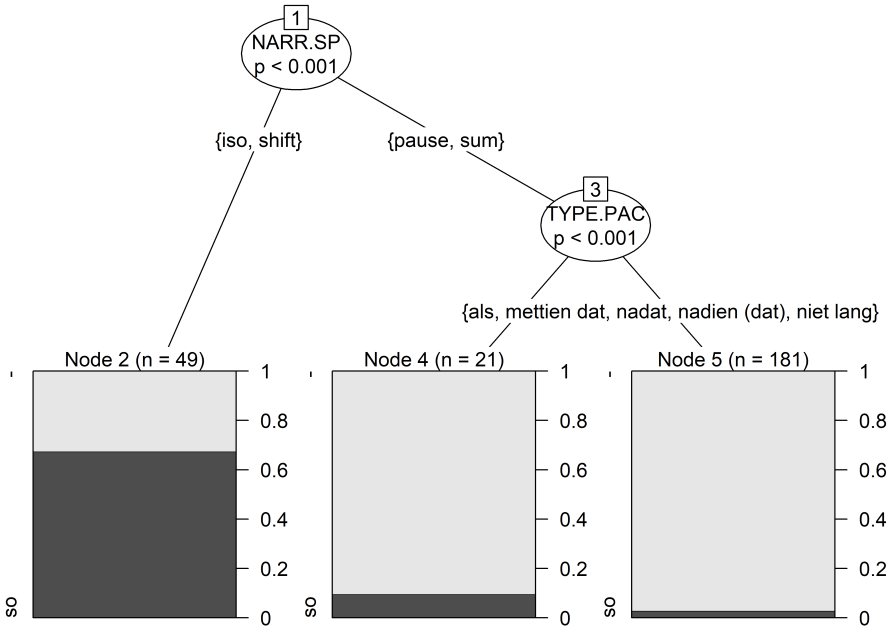


Figure 9: Conditional inference tree: *so* versus integration in the Dutch Ponthus

summaries, the preference for integration is strong.[7] This motivates the first split in the tree. Second, within the group of clauses that encode pauses or summaries, most of the types of adverbial clauses do not, or only rarely, occur with *so*. This preference for integration is significantly stronger with *als, mettien dat, nadat, nadien (dat)* and *niet lang* than with V1 and *doen*. This split reflects partly what we saw for German in §4: Both the type of PAC and the narrative speed are important factors. Yet, different constructions are associated with narrative summaries and isochrony. Moreover, the variation in the types of PACs was greater in German than in Dutch and PAC + *so* was overall more frequent.

6 Conclusion

This paper investigated how preposed adverbial clauses (PACs) are used to structure the narrative in the Early New High German Ponthus and compares the

[7]Since this data only contains one pause and five shifts, I do not consider this strong evidence.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

results to what is found in the Dutch adaptation.

The first part of the paper focuses the two correlative constructions, as PACs were very often followed by *da* or *so* in Early New High German. Their contrasting uses regarding narrative structure were evaluated. Although *so* is predominantly found in mimesis, viz. dialogic segments, and TEMP in diegetic segments, which are voiced by the narrator, it is argued that this is an artifact of a difference in narrative speed. Specifically, *so* is associated with isochrony, which slowly progress the story, and TEMP with narrative summaries, which progress it rapidly. This effect is observed even when including the type of PAC as a predictor variable. The Dutch Ponthus had, differently than German, a remarkably low amount of TEMP. The results were consequently not very convincing. The second part of the paper therefore considers PACs in all constructions in the two Ponthus-texts. That is, it includes subjects following the PAC, direct adjacency of the PAC and the finite verb, and 'other' elements that may intervene between the PAC and the finite verb of the host. The main result is that where German uses temporal resumption in narrative summaries, Dutch uses integration. Moreover, the sequence PAC + SUBJ is more easily predictable in Dutch than in German on the basis of the type of PAC and narrative speed, which suggests that it has developed its own niche in Dutch. Seeing that the Dutch Ponthus was published about a century later than the German manuscript, the differences between the languages may be a reflection of a larger diachronic development, but to draw any reliable conclusions other types of data will need to be considered.

It should be noted, however, that the current study investigates the use of the PACs in two adaptations of one and the same text. As such, the findings may be an artifact of different translation strategies or idiosyncrasies of the individual translators rather than a reflection of conventionalized patterns. Further research based on a broader base is needed to verify whether the relation between narrative speed and the realization of the prefield as well as the contrast between German and Dutch is represented in narrative prose or the languages more generally. In addition, the verbal tense and mood seemingly interact with both narrative speed and the type of adverbial clause. Further research in this interplay would likely be fruitful and provide deeper insight into the issue at hand. In any case, the paper forms a first basis to the interaction between the realization of the prefield following preposed adverbial clauses and narrative structure. It illustrates how the different constructions in which preposed adverbial clauses may occur can be used to temporally progress the story, and how the Dutch and German Ponthus use similar strategies but different realizations to do so.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Abbreviations

Glosses follow the Leipzig Glossing rules.

CNJ	conjunction	PAC	preposed adverbial clause
DISC	discourse marker	SBJV	subjunctive
DU	Dutch	SUBJ	subject
FR	French	sum	summary
GE	German	TEMP	temporal adverb
iso	isochrony	V1	verb-first
NEG	negation	V3	verb-third

Acknowledgements

This research was funded by the Deutsche Forschungsgemeinschaft (ProjectID 456973946, „Wortstellung und Diskursstruktur in der Frühen Neuzeit“).

Primary sources

German Pontus und Sidonia. (3rd quartile 15th c.). *Pontus und Sidonia. Tandareis und Flordibel - BSB Cgm 577.* München, Bayerische Staatsbibliothek. https://www.digitale-sammlungen.de/en/view/bsb00014877. Transcribed as part of the Romankorpus Früneuhochdeutsch.

Dutch Ponthus ende Sidonie. 1564. *Een schoone ende amoruese historie van Ponthus ende die schoone Sydonie.* Transcribed by Willem Kuiper & Paul Jacob Brüggeman. 2010. Diplomatic edition. Amsterdam: Bibliotheek van Middelnederlandse Letterkunde.

French Ponthus. 1479. *Le livre de Ponthus filz du roy de Galice et de la belle Sydoine fille du roy de Bretaigne.* Gallica, Bibliothèque nationale de France, département Réserve des livres rares, RES-Y2-567. https://gallica.bnf.fr/ark:/12148/btv1b8600082d/

References

Axel-Tober, Katrin. 2023. Adverbial resumption in German from a synchronic and diachronic perspective. In Karen De Clercq, Liliane Haegeman, Terje Lohndal & Christine Meklenborg (eds.), *Adverbial resumption in verb second languages*, 167–194. Oxford: Oxford University Press. DOI: 10.1093/oso/9780197651148.003.0007.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Bal, Mieke. 2009[1985]. *Narratology: Introduction to the theory of narrative*. 3rd edn. Toronto: University of Toronto Press.

Bertelsmeier-Kierst, Christa. 2019. Durchbruch zur Prosa und der Einfluss des Buchdrucks auf die deutschsprachige Erzählliteratur des 15. Jahrhunderts. In Bart Besamusca, Elisabeth de Bruijn & Frank Willaert (eds.), *Early printed narrative literature in Western Europe*, 17–48. Berlin: De Gruyter. DOI: 10.1515/9783110563016-002.

Besamusca, Bart & Frank Willaert. 2019. Continuities and discontinuities in the production and reception of Middle Dutch narrative literature. In Bart Besamusca, Elisabeth de Bruijn & Frank Willaert (eds.), *Early printed narrative literature in Western Europe*, 49–92. Berlin: De Gruyter. DOI: 10.1515/9783110563016.

Bloom, Barthe, Malika Reetz & Ulrike Demske. 2024. Romankorpus Frühneuhochdeutsch (Roko.UP): A collection of Early Modern German narrative prose. *Journal of Open Humanities Data* 10(17). 1–7. DOI: 10.5334/johd.188.

Catasso, Nicholas. 2021. Generalized and specialized adverbial resumption in Middle High German and beyond. *Journal of Historical Syntax* 5. 1–38. DOI: 10.18148/hs/2021.v5i1-13.40.

Chatman, Seymore. 1978. *Story and discourse*. Ithaca, NY: Cornell University Press.

Chiang, Ted. 2002. Hell is the absence of god. In Ted Chiang (ed.), *Stories of your life and others*, 123–150. New York: Tor Books.

Classen, Albrecht. 2016. The innovative perception of space (Europe) in late medieval German literature: The spatial turn in light of Eleonore of Austria's Pontus und Sidonia (ca. 1450–1460). *Neohelicon* 43(2). 543–557. DOI: 10.1007/s11059-016-0347-8.

Condoyannis, George E. 1944. Word order in colloquial German. *Monatshefte für Deutschen Unterricht* 36(7). 371–377.

Demske, Ulrike. 2018. Syntax and discourse structure: Verb-final main clauses in German. *Linguistische Berichte* 25. Special issue: Non-canonical verb positioning in main clauses, 135–159.

Demske, Ulrike. 2019. Zur Autonomie indirekter Redewiedergabe – Eine diachrone Perspektive. *Zeitschrift für Germanistische Linguistik* 47(1). 70–101. DOI: 10.1515/zgl-2019-0003.

Ebert, Robert, Oskar Reichmann, Hans-Joachim Solms & Klaus-Peter Wegera. 1993. *Frühneuhochdeutsche Grammatik*. Tübingen: Niemeyer.

Fludernik, Monika. 1993. *The fictions of language and the languages of fiction. The linguistic representation of speech and consciousness*. London: Routledge.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Fludernik, Monika. 2003. The diachronization of narratology. *Narrative* 11(3). 331–348. DOI: 10.1353/nar.2003.0014.

Gaiman, Neil. 2017. *Norse mythology*. New York: W. W. Norton & Company.

Genette, Gérard. 1983[1980]. *Narrative discourse revisited*. Ithaca: Cornell University Press.

Genette, Gérard. 1980[1972]. *Narrative discourse: An essay in method*. Ithaca: Cornell University Press.

Hothorn, Torsten, Kurt Hornik & Achim Zeileis. 2006. Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics* 15(3). 651–674. DOI: 10.1198/106186006X133933.

de Jong, Irene. 2004. *Narrators and focalizers: The presentation of the story in the Iliad*. 2nd edn. Bristol: Bristol Classical Press.

Kuiper, Willem & Paul Jabob Brüggeman. 2010[1564]. *Een schoone ende amoruese historie van Ponthus ende die schoone Sydonie*. Amsterdam: Bibliotheek van Middelnederlandse Letterkunde.

Kukkonen, Karin. 2020. The speed of plot: Narrative acceleration and deceleration. *Orbis Litterarum* 75(2). 73–85. DOI: 10.1111/oli.12251.

Louviot, Elise. 2016. *Direct speech in Beowulf and other Old English narrative poems*. Cambridge: Boydell & Brewer.

Meklenborg, Christine. 2020. Adverbial resumptive particles and verb second. In Rebecca Woods & Sam Wolfe (eds.), *Rethinking verb second*, 90–125. Oxford: Oxford University Press. DOI: 10.1093/oso/9780198844303.003.0005.

Müller, Stefan. 2015. German: A grammatical sketch. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – Theory and analysis*, vol. 3, 1447–1478. Berlin: De Gruyter. DOI: 10.1515/9783110363685-001.

Oksefjell Ebeling, Signe & Jarle Ebeling. 2020. Dialogue vs. narrative in fiction. *Languages in Contrast* 20. 288–313. DOI: 10.1075/lic.00019.oks.

Packard, Stephan. 2008. Two-dimensional narrative velocity. *Journal of Literary Semantics* 37(1). 55–67. DOI: 10.1515/jlse.2008.004.

Phinney, Mary Catherine. 1962. Review Pontus und Sidonia in der Verdeutschung eines Ungennanten aus dem 15 . Jahrhundert by Karin Schneider. *The Journal of English and Germanic Philology* 61(4). 958–960.

Pier, John. 1986. Diegesis. In Thomas Albert Sebeok (ed.), *Encyclopedic dictionary of semiotics*, vol. 1 A-M, 209–211. Berlin: De Gruyter.

Pittner, Karin. 1999. *Adverbiale im Deutschen. Untersuchungen zu ihrer Stellung und Interpretation*. Tübingen: Stauffenburg.

R Core Team. 2023. *R: A language and environment for statistical computing*. R version 4.2.3 (2023-03-15 ucrt). Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Schneider, Karin. 1961. *Pontus und Sidonia: In der Verdeutsschung eines Ungenannten aus dem 15. Jahrhundert*. Berlin: Erich Schmidt.

Strobl, Carolin, Anne-Laure Boulesteix, Thomas Kneib, Thomas Augustin & Achim Zeileis. 2008. Conditional variable importance for random forests. *BMC Bioinformatics* 9(307). DOI: 10.1186/1471-2105-9-307.

Strobl, Carolin, Anne-Laure Boulesteix, Achim Zeileis & Torsten Hothorn. 2007. Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics* 8(25). DOI: 10.1186/1471-2105-8-25.

Weinert, Regina. 2007. The structure and function of *wenn*-clauses and their role in problem-solving discourse. In Regina Weinert (ed.), *Spoken language pragmatics. Analysis of form-function relations*, 60–93. London: Continuum.

Zeman, Sonja. 2023. ‚Linksversetzungen‘ im Mittelhochdeutschen als Beispiel narrativer Syntax. *Jahrbuch für Germanistische Sprachgeschichte* 14(1). 68–84. DOI: 10.1515/jbgsg-2023-0007.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Appendix

(21) *Sie sprachen da were vil abe zu sagen Dan siner*
they said there was.SBJV much about to say because his
schonheit weydelicheit ~~were~~ vnd sins gutten gelaß were nit
beauty proficiency ~~was.SBJV~~ and his good behavior was.SBJV not
me in der werlt als sie wol gleuben wolten williche frauwe oder
more in the world as they DISC believe wanted which woman or
jungffrauwe yne zu bolen mochte haben da ist kein rede in Sie mocht
lady him to court might have there is no reason in she might
wol berummen das sie den schonsten den besten den
DISC state that she the most.beautiful the best the
wolkommesten vnd den aller lustigesten frunt hette der da
most.considerate and the of.all amusing friend have.SBJV who there
lebt vnd schon ist sin jugent ***kompt*** *er also inn das alter so wart*
lives and beautiful is his youth comes he thus in the age so became
sin gliche nie me ersehen das ist vngezwiuelt die rechte warheit
his equal never more seen that is undoubtedly the right truth
'They said that there was much to say about that, because his beauty, proficiency, and good manners were unequaled in this world. As they would believe whichever woman or lady might court him – there is no reason in that – she might proudly profess that she would have the most beautiful, the best, the most considerate, and the uttermost amusing friend in the world. And beautiful is his youth. **Comes** he of age, his equal was never to be seen again. That is undoubtedly the truthful truth.'
[GE, 16rb][8]

[8]The crossed-out *were* is written in the handwriting. This is an accidental doubling.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Chapter 4

Discourse structure and the reorganisation of the Icelandic aspectual system

Jordan Chark [a]

[a]Humboldt-Universität zu Berlin & Leibniz-Zentrum Allgemeine Sprachwissenschaft

Icelandic is commonly considered a conservative language. While this claim holds true for its morphology, especially in the broader Germanic perspective, one area that has undergone considerable changes since the 16th century is the mapping between syntax and discourse structure. In this paper, I will argue that otherwise puzzling facts from the history of the language can be better understood if one takes this shift in discourse structuring as fundamental. While most research to date has focused on a limited set of syntactic phenomena, I bring new data to bear on this question: the emergence of two aspectual markers, the *búinn*-perfect and the *vera að*-progressive. I trace the shift in discourse structure alongside the emergence of innovative aspectual markers in two corpora of historical Icelandic (**IcePaHC**, Wallenberg et al. (2011), and **Íslenskt textasafn**, Stofnun Árna Magnússonar í íslenskum fræðum (2023)). I show that what seem to be abrupt changes in the corpus record no longer seem so abrupt when considered in the larger context of this shift from a bounded to an unbounded system. Rather, the loss of topic-time anchors like *þá* 'then' paved the way for the emergence of new ways of construing events (Los 2012) during the 16th-18th centuries, a period of variability reflected in the corpus dataset by travelogues and memoirs that combine more archaic and innovative characteristics.

1 Introduction

Icelandic, despite its well-known morphosyntactic conservatism (cf. Friðriksson 2011; Rögnvaldsson & Helgadóttir 2011), has undergone considerable changes in

Jordan Chark. 2025. Discourse structure and the reorganisation of the Icelandic aspectual system. In Ulrike Demske & Barthe Bloom (eds.), *Discourse structure and narration: A diachronic view from Germanic*, 73–131. Berlin: Language Science Press. DOI: ??

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

the mapping between clause and information structure since the early modern period (1540 onwards).[1] To date, scholarly work in this area has focused on the syntactic reflexes of these changes (Rögnvaldsson 1995, 2002, Booth & Beck 2021). In this paper, I focus on a topic that has received little attention in the diachronic literature on Icelandic: the emergence of periphrastic aspectual constructions, the *búinn*-perfect and the *vera að*-progressive. I will argue that the trajectories of these periphrastic constructions can be understood within the context of this fundamental shift in mapping between form and discourse structure. The diachrony of Icelandic exhibits a shift from a predominantly 'bounded' system for expressing narrative to a predominantly 'unbounded' one (see Petré 2010, 2014; Los 2012; Bech 2014; Fanego 2024 on a similar shift in the history of English).

Drawing on corpora of historical Icelandic, I show that between 1540–1850 (what I define as the early modern period)[2] the perfect construction *búinn* and the dedicated progressive construction formed with *vera að* 'be to' went from once marginal constructions to become integral parts of the aspectual oppositions encoded in Icelandic grammar. Both are attested in earlier Icelandic, not having yet acquired their modern aspectual meanings. More broadly, I argue that these changes can be subsumed under a shift from a bounded to unbounded system.

I argue that this shift in boundedness is reflected in a change in grounding strategies in the construing of narrative sequences. This includes the marking of episodic boundaries in narratives and the way in which information-structural notions are tracked overtly in the syntax. In this paper, I focus on the early modern period of Icelandic, which I take to begin with the New Testament translation by Oddur Gottskálksson (completed in 1540) and to end with the publication of what may be considered the first modern Icelandic novel, *Piltur og stúlka*, published in 1850 (Bernharðsson 2017, 2018).

This paper is structured as follows. First, this introductory section continues with background information on the corpora used. Next, in Section 2, I discuss the grammaticalisation of the *búinn*-perfect and the dedicated *vera að*-progressive: Section 2.1 begins with the schema I assume for grammaticalisation in context, Section 2.2 discusses the emergence of the *búinn*-perfect, Section 2.3 discusses the emergence of the *vera að*-progressive. Section 3 continues with a discussion

[1]This period is demarcated by the publication of Oddur Gottskálksson's translation of the New Testament into Icelandic (Ottósson 1990).

[2]Note that terminology surrounding the periodisation of Icelandic is used inconsistently in the literature, with some authors including early 20th century (Svavarsdóttir et al. 2014), while others such as Brendel (2023) refer roughly to the period spanning the 16th to 19th centuries.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

of discourse traditions (historical registers): 3.1 introduces the term and 3.2 situates the texts used in this study within their respective discourse traditions. Section 4 then proceeds with a discussion of the shift from boundedness to unboundedness, which is the primary claim of this paper: 4.1 provides the motivation behind this approach and 4.2 discusses the literature in this area in more detail. I then argue that Old(er) Icelandic (prior to 1540) represents a predominantly bounded system 4.2.1), that 17th and 18th century Icelandic has an intermediate system (4.2.2) and that modern Icelandic (post 1850) ought to be considered unbounded (4.2.3. In Section 5, I delve into this decrease in bounded strategies: 5.1 discusses work on Old English in this regard, in order to set the stage for the Icelandic data in comparison. Section 5.2 then provides statistical corpus evidence to support my claim that Icelandic underwent a shift in boundedness during this period: this evidence includes a general increase in temporal subordinate clauses (5.2.1), that the rise of the subordinator *þegar* is negatively correlated with the decline of the topic-time adverbial *þá* (5.2.2), that the requirement for overt temporal adverbials in the left periphery becomes less strict (5.2.3) and finally that the adverbials associated with *búinn* and *PROG* (5.2.4), as well as their relative timing support this view as well (5.2.5). Section 6 concludes with a summary of my findings and a reflection on implications for future work on the diachrony of Icelandic as well as diachronic typology more generally.

Throughout the paper, I draw my data[3] primarily from two corpora:

1. **IcePaHC** (Wallenberg et al. 2011) ≈ 1 million words. Timeframe: 1150-2008, with diverse genre distribution. Parsed in Penn treebank format.[4]

2. **Íslenskt textasafn** (Stofnun Árna Magnússonar í íslenskum fræðum 2023). I utilise two components of this corpus: (i) the 16th-18th century subcorpus which consists of more substantial material than that available in IcePaHC for this time period ≈ 421.000 words, outlined in Table 1,[5] and (ii) I make use of *Ritmálssafn Orðabókar Háskólans*, the archive of written language for the dictionary of Icelandic compiled at the University of Iceland, which contains 2.5 million example sentences of 700.000 header words.[6]

[3]Data and scripts available via https://github.com/jordanchark/discourse_narration.

[4]Version 0.9. More information on IcePaHC can be found at https://github.com/antonkarl/icecorpus.

[5]The texts *Nikulás Klím*, *Nýja testamentið* and *Ævisaga síra Jóns* are present in IcePaHC in much smaller excerpts.

[6]This corpus can also be accessed at https://ritmalssafn.arnastofnun.is/

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Table 1: Íslenskt textasafn, 16-18th century subcorpus. Translator in parentheses. Count is total words present in documents provided by the corpus editors.

Title	Author	Year	Words
Einfalt matreiðsluvasakver	Marta María Stephensen	1800	16,592
Nikulás Klím	Holberg, Ludvig (Jón Ólafsson)	1745	23,332
Nýja testamentið	(Oddur Gottskálksson)	1540	189,692
Stuttur siðalærdómur	Campe, J.H (Guðlaugur Sveinsson)	1799	36,773
Uppkast til forsagna	Eggert Ólafsson	1767	35,242
Vídalínspostilla	Jón Þorkelsson Vídalín	1718-20	156,497
Ævisaga síra Jóns	Jón Steingrímsson	1791	107,926

2 Emergence of a new perfect and progressive

Modern Icelandic exhibits a tripartite system for aspectual distinctions: inchoative, progressive and completive. This system is comprised of periphrastic constructions which embed an infinitival phrase. By the end of the early modern period, this system is well established. Evidence for this comes from grammatical descriptions of this period. Jón Magnússon, in his grammatical description *Grammatica Islandica* (written 1737-1738) (Jónsson 1933) under the heading *De Verbis Auxiliaribus*, the headings are taken from this text, translated from the Latin:[7]

(1) These all distinguish the present

a. *Ég tek nú til að gjöra*
I take now to to do
'I now begin to do.'

b. *Ég er farinn að gjöra.*
I am gone to do
'I began doing.'

c. *Ég er að gjöra.*
I am to do
'I am doing.' (Jónsson 1933: 130)

[7]Spelling is adapted to modern orthography. The list is not complete as Magnússon's account also includes modal expressions, which are not relevant for the purposes of this paper.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(2) These signify the imperfect past

a. *Ég fór* *að gjöra.*
I went to do.
'I began doing.'

b. *Ég var* *að gjöra.*
I was to do.
'I was doing.' (Jónsson 1933: 130)

(3) These indicate the perfect

a. *Ég er* *búinn* *að gjöra.*
I am búinn to do.
'I have done.'

b. *Ég hef* *gjört.*
I have done.
'I have done.' (Jónsson 1933: 130)

(4) These indicate the pluperfect

a. *Ég hafði gjört.*
I had done
'I had done.'

b. *Ég var* *buinn* *að gjöra.*
I was BÚINN to do
'I had done.' (Jónsson 1933: 130)

As we see here, by the mid-18th century the primary tripartite oppositions are considered to be integral enough to the language for Jón to include all of their forms in the paradigm.[8]

Prior to this, the paradigm of aspectual distinctions made available by the grammar is arguably incomplete (cf. Árnason 1977). By this I mean that the distinctions made overtly are asymmetrical. First, prior to 1550, marking of inchoative aspect is already established, marked by an infinitival phrase embedded under the verb *taka* 'take'. However, a completive aspect and progressive aspect do not emerge until later. These will be the focus of this paper. We begin with the diachrony of the progressive construction with *vera að*.

[8]His description also makes it clear that Icelandic of this period had competing forms for the inchoative: one formed by *taka* 'take' along with an infinitive and the other formed by the participle *farinn* 'gone' in conjunction with an infinitival phrase.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

2.1 Incidental variation

In what follows, I argue for the role of what I term incidental variation. Incidental privileging contexts in diachrony have also been termed *beachheads* (Givón 1979, Aldai 2002). Such an approach is pursued by Petré (2016) in his account of the diachrony of the English progressive. I would like to suggest that changing frequencies of environments that themselves privilege usages in question were responsible for the seemingly rapid rise in frequency in the diachrony of Icelandic for both *búinn* and the progressive. Focalised progressive and anterior *búinn* serve similar text-structuring purposes – the prevalence of privileging contexts should correlate with dimensions of textual organisation.

It has long been recognised by scholars that grammaticalisation happens in context (Diewald 2002, Heine 2002). That is, ambiguities arise in particular contextual configurations, leading to the possibility of an alternation between two expressions at this juncture, which eventually can give rise to the cementing of new semantic meaning arising from contextually-dependent inferences. Schematically, this can be represented as follows (Heine et al. 1991):[9]

(5) Increased frequency → Context-dependent implicature becomes more salient ⟶ Ambiguity (Bridging contexts) → Reanalysis (Switch contexts) → Generalisation

In the following subsections, I argue for the importance of temporal subordinate contexts in the conventionalisation of the two aspectual constructions which are the focus of this paper: *búinn* and *PROG*. We begin by examining the emergence of *búinn* in 2.2, followed by *PROG* in 2.3.

2.2 Emergence of BÚINN

Modern Icelandic has two constructions that have a number of properties typically ascribed to perfects (Jónsson 1992, Thráinsson 2017): *búinn* 'finish' and *hafa* 'have'. All primary perfect readings are found with the *hafa* perfect (experiential, resultative, universal, cf. McCawley 1971, Comrie 1976) already in the earliest written Icelandic sources (Nygaard 1905, Pollak 1930). The past participle *búinn* is derived from the verb *búa* 'reside, prepare, adorn' and does not acquire an anterior, perfect -like meaning until the 17th century (Thráinsson 2017). Prior to the 17th century, *búinn* predominantly has a 'prepared' reading. The example

[9] In the system proposed by Heine et al. (1991), bridging contexts allow ambiguity between conservative and innovative meanings, whereas switch contexts only permit an innovative reading.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

below, for instance, cannot have anterior reference (i.e. receive a perfect -type reading).[10]

(6) *En er Egill var búinn og byr gaf þá siglir hann í haf.*
and when Egill was BÚINN and wind gave then sails he to sea
'And when Egill was prepared and there was wind, he sailed to sea.'
[**1250**.THETUBROT.NAR-SAG,.74][11]

My claim is that the later anterior reading is due to the underspecification of *búinn* (cf. Rosemeyer & Grossman, 2017 on a similar diachronic path in Spanish). Preparedness is a context-dependent notion which bleeds together with completeness (an example of 'type coercion'; Pustejovsky 1998). This is illustrated for historical Icelandic in (7) (14th century) where it is apparent that a net which is prepared for fishing implies a complete or finished net.

(7) *Ok er búið var netit, þá fara Æsir til árinnar ok kasta neti í forsinn*
and when BÚINN was the:net, then go Æsir to the:river and throw net into the:waterfall
'And when the net was finished, the Æsir went to the river and threw the net into the waterfall.'
(Sigurðardóttir et al. (ONP): Jónsson, 1931: 69, English from Sturluson 1995)[12]

In early stages of Icelandic *búinn* is often used to add emphasis to a temporal sequence, adding some punch to the narrative by highlighting the sequential relationship between two eventualities (for a diachronic trajectory with similar characteristics, see Rosemeyer & Grossman 2017). Clauses involving *búinn* are typically complement-less, though not exclusively so. In my investigation of the *Fornrit* corpus (Stofnun Árna Magnússonar í íslenskum fræðum 2023), out of 1591 tokens, 46 of them are infinitival.[13] These have a future-oriented prepared-type reading. By contrast, in my extended dataset ranging from the 16th to 19th

[10]The adjectival meaning of the participle outside of this construction also shifts from predominantly having a 'prepared' reading to meaning 'complete' after the grammaticalisation of the infinitival construction (Árnason 1977).

[11]Examples sourced form IcePaHC are given with metadata information in the format: Year.Text.Genre-Subgenre.TokenID

[12]ONP (Sigurðardóttir et al. 2024) is a digital dictionary of Old Norse prose and can be accessed at https://onp.ku.dk/onp/onp.php.

[13]The *Fornrit* corpus is part of the *Íslenskt textasafn* corpus, additional available for search at https://malheildir.arnastofnun.is/. The corpus consists of approximately 1.5 million words (Stofnun Árna Magnússonar í íslenskum fræðum 2023, Rögnvaldsson & Helgadóttir 2011).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

centuries, when this conventionalisation was ongoing, upwards of 80% of *búinn* tokens have infinitival complements.

(8) *En er Haraldur konungur var búinn að stíga á hest sinn þá*
and when Haraldur king was BÚINN to mount on horse his then
bað hann kalla til sín Áka búanda.
asked he call to him Áki farmer
'When King Harald was ready to (about to) mount his horse, he asked for Áki the farmer to be called to him.'
(Heimskringla, cited in *Fornrit* corpus: Stofnun Árna Magnússonar í íslenskum fræðum 2023)

In temporal subordinate contexts expressing a sequence between "event time" (E) and "reference time" (R) (E-R-S, as per Reichenbach 1947), it would typically be under-informative to overtly express the event description for which one is preparing, assuming that the following main clause asserts such an event. In other words, overtly expressing that an event that was an imminent occurrence (cf. Eckardt 2006: 91-127 on *be going to*) at some past interval did indeed become instantiated is under-informative (on Gricean assumptions; Holvoet 2014).

Below in (9), we see a prototypical early anterior usage of *búinn*, highlighting a relationship between two events, e_1 strictly preceding e_2.

(9) *Þegar ég var nú búinn að tala við Jón yngra úti kom ég inn*
when I was now BÚINN to speak with Jón younger outside came I in
aftur í kirkjuna
again into the:church
'When I was finished speaking/had spoken with John junior outside I came into the church again.' [**1659**.PISLARSAGA.BIO-AUT.56]

I assume that innovative speakers took advantage of the underspecification of *búinn* as well as analogy with related expressions (relating to the present) and began to combine *búinn* with infinitival complements. Hearers, in turn, were left to make sense of the utterance, interpreting pluperfect *búinn* constructions as expressing that the event description associated with this infinitival complement is informative relative to the discourse context (cf. Rosemeyer & Grossman 2017 for an analysis of a Spanish construction along similar lines). To be more precise, I am assuming a mechanism along the lines of Eckardt (2009)'s principle of Avoid Pragmatic Overload.[14]

[14] Eckardt (2009: 14) defines this principle as follows: "The utterance under a conservative inter-

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

One way of resolving informativity with regard to the discourse context here is to interpret the event description of the infinitival complement not as expressing a prospective, yet uninstantiated event but rather, highlighting the relationship between two instantiated eventualities. This results in the construction involving the participle *búinn* alongside the infinitive being reanalysed. The hearer can reason that what is relevant is the temporal interval associated with the result state of the event described by the infinitive, which brings about a new state of affairs (cf. Rosemeyer & Grossman 2017). In other words, asserting the existence of this state has immediate consequences for what can happen afterwards (cf. Wide 2002 on *búinn* in modern Icelandic discourse).

The question remains, however, of what constituted the system-internal motivation for the emergence of this marker, if Icelandic already had a canonical means of marking the perfect (i.e. with *hafa* 'have'). It seems to me that a key component of such a system-internal motivation is the considerable overlap between the *hafa*-perfect and the simple past that could be observed in Old Icelandic. In narrative texts, Old Icelandic is also known for considerable tense alternations between present and past for stylistic effect (Nygaard 1905, Heusler 1921). I take it that these facts led speakers of Icelandic to innovate a new marker for the purposes of highlighting the relationship between two events, i.e. the "perfect-in-the-past" reading (Comrie 1976). The canonical *hafa*-perfect, compatible with adverbs that specify Reichenbachian event time (i.e. "past-in-the-past" reading: Comrie 1976) and overlapping in use with the simple past, would have constituted a functional motivation for innovation.[15] This fact should be considered in light of the emerging system for overtly marking aspectual distinctions; Old Icelandic already used *taka* 'take' as an inchoative marker (Nordling 1928, Holm 1958). As I have discussed, at this stage we also see early focalised, aspectual usages of the progressive. I would like to propose that this constituted system-internal analogical motivation to complete this paradigm with a completive, an option that the chance reanalysis of the *búinn* construction made possible.

pretation will trigger presuppositions that the hearer can not easily accommodate or refute. The hearer has three options: S/he can (a) be uncharitable and refuse to interpret the utterance at all, or (b) face the pragmatic overload and attempt to reconceptualize the world such that the presupposition makes sense and is consistent, or (c) hypothesize a new meaning for parts of the utterance, notably the item that gave rise to the problematic presupposition. Option (c) will allow the hearer to Avoid the Pragmatic Overload."

[15]The canonical *hafa* perfect has an inferential flavour in older as well as modern stages of the language, which pertains to the "source of information" (Liberman 1990). This differentiates it from the simple past. However, it could also pave the way for a form to emerge that has a *current relevance* function (Comrie 1976: 52). A "past-in-the-past" reading can be seen as closer to a past perfective or remote past (Salkie 1989). That pluperfects can themselves incorporate a degree of current relevance is discussed in Wada (1995: 102-3) and Lee (2017: 72).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

These facts ought to, in turn, be considered in the context of inter-clausal relations being somewhat in flux during the transitional period of Early Modern Icelandic, which I draw attention to later on in this paper (Section 5.2). I take it that the tighter integration of adverbial clauses at this point in time had its part, too, in encouraging the explicit marking of temporal overlap, a hallmark of unbounded narrative.

While *búinn* is still young and/or not yet established, one finds a number of strategies in Icelandic for the expression of completive aspect. Constructions involving the past participle are often used to designate time in the place of subordinate clauses (Hauksson & Óskarsson 1994: 399). The past participle can also be found as part of a *dativus ablativus*, which is a feature of learned style in Icelandic of the medieval period (ibid.). The latter is illustrated below in (10).

(10) *Og strax honum uppgengnum, þremur stykkjum*
and immediately he.DAT up:go:PTCP.PST.DAT three.DAT pieces.DAT
affýrðum og kveðjum orðnum, afdró
off:fire.PTCP.PST.DAT and farewells:DAT become.PTCP.PST.DAT off:threw
hann sinn ysta kyrtil og gekk svo til káhyttu.
he his outermost cloak and went so to cabin
'When he was come on board and three pieces of cannon had been fired, and the greetings were ended, he threw off his outermost tunic and went down into the cabin.' (Ólafsson 1908: 111)[16]

These alternative strategies all reflect the situation at this point in the diachrony of the language in which no formal distinction is made between a stative, adjectival, reading of past participles and that of an eventive passive, as shown below in (11a). Kress (1982) argues that pressure to fill-in this paradigm was a factor in the grammaticalisation of *búinn*, as it makes overt this distinction, as shown in (11b).

(11) a. *Þeir voru ráðnir.*
they were hired
'They were hired.' (eventive or stative possible, cf. *Sie waren/wurden angestellt*) (Kress 1982)

b. *Það er búið að ráða þá.*
it is BÚINN to hire them
'They were/have been hired.' (eventive only) (Kress 1982)

[16]Translation from Ólafsson (2017: 203-4).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Crucial for the purposes of this paper is the role of context-dependent inferences for the rise of the modern, completive meaning of *búinn*. This modern usage arises in the language of those born shortly after 1550 (established by an investigation of writings by authors in and around this period by Árnason 1977). Early anterior readings of *búinn* are found predominantly in temporal subordinate contexts: these act as *switch* contexts (Heine 2002), as the older meaning is typically ruled out given the main clause that follows. The importance of past-marked, temporal subordinate contexts for the establishment of the construction is clear when one investigates the proportion of *búinn* tokens which appear in a subordinate clause expressing temporal sequentiality (which I term *SuB* in the plot below). This is shown in Tables 2 and 3 for the 17th and 18th century respectively.[17]

Table 2: Frequency of *búinn* perfect in subordinate clauses expressing temporal sequentiality (18th century, *Íslenskt textasafn*)

Source	Century	SuB	Count	Total	Proportion
(pislarsaga)	17	0	14	16	0.88
(pislarsaga)	17	1	2	16	0.12
(Rms-sitot-17)	17	0	59	70	0.84
(Rms-sitot-17)	17	1	11	70	0.16

This point is further illustrated in Figure 1, which displays the distribution in IcePaHC by clause type and century. We see that, while the innovative construction emerges in past-marked subordinate clauses, it subsequently gains ground in present-marked main clauses in the 19th century. Finally, Figure 2 illustrates the time-course of propagation for the *búinn* construction. Relative frequency is given in percentage of total words. It remains somewhat marginal until the early 19th century, never exceeding 0.5% relative frequency except for in one instance (*klim*), after which point many texts in IcePaHC either closely approach or exceed this threshold.

Below in Figure 3 and Figure 4, we see the proportion of *búinn* constructions (*búinn* followed by an infinitive) in the IcePaHC corpus relative to all perfect

[17] Here, 0 indicates the absence of the property, whereas 1 indicates its presence. More information on the texts can be found at https://corpus.arnastofnun.is/leit.pl?info=2. Rms-sitot designates examples that originate from *Ritmálssafn Orðabókar Háskólans*, the archive of written language for the dictionary of Icelandic compiled at the University of Iceland. This corpus can be queried independently at https://ritmalssafn.arnastofnun.is/.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Table 3: Frequency of *búinn* perfect in subordinate clauses expressing temporal sequentiality (18th century, *Íslenskt textasafn*)

Source	Century	SuB	Count	Total	Proportion
(ÆVISAGAN)	18	0	74	91	0.81
(ÆVISAGAN)	18	1	17	91	0.19
(campsid)	18	0	6	11	0.55
(campsid)	18	1	5	11	0.45
(eggbrud)	18	0	19	20	0.95
(eggbrud)	18	1	1	20	0.05
(klim)	18	0	13	28	0.46
(klim)	18	1	15	28	0.54
(martam)	18	0	11	18	0.61
(martam)	18	1	7	18	0.39
(Rms-sitot-18)	18	0	148	166	0.89
(Rms-sitot-18)	18	1	18	166	0.11

environments, i.e. in competition with *hafa*. The tokens are sorted by clause-type: IP-MAT designates matrix clauses and IP-SUB designates subordinate clauses. The importance of past-marked subordinate clauses early on in the diffusion of the construction is apparent from these plots, as well.

2.3 The progressive in Icelandic: Diachrony

The Icelandic progressive conforms to a cross-linguistically established trajectory, emerging from a locative (Bybee et al. 1994) e.g. *vera að kirkju* 'be at church', *vera að smíðum* 'be at building' (Benediktsson 2002). Bybee et al. (1994: 129) observe within their typological dataset a tendency for progressive markers to originate from source constructions that are locative expressions of some kind. The authors write that such locatives are "probably 'be in the place of verbing' or 'be at verbing' " (ibid. 130).[18] This function then seems to have undergone metaphorical or metonymic extension, whereby the spatial concept of location within an activity is extended to the temporal domain.

The Icelandic progressive construction combines the copula *vera* 'to be' accompanied by double *að*, which functions as a preposition and infinitival marker.

[18]That such a construction involves a copula is also cross-linguistically common, which is perhaps not surprising given the semantic import of copular predication.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

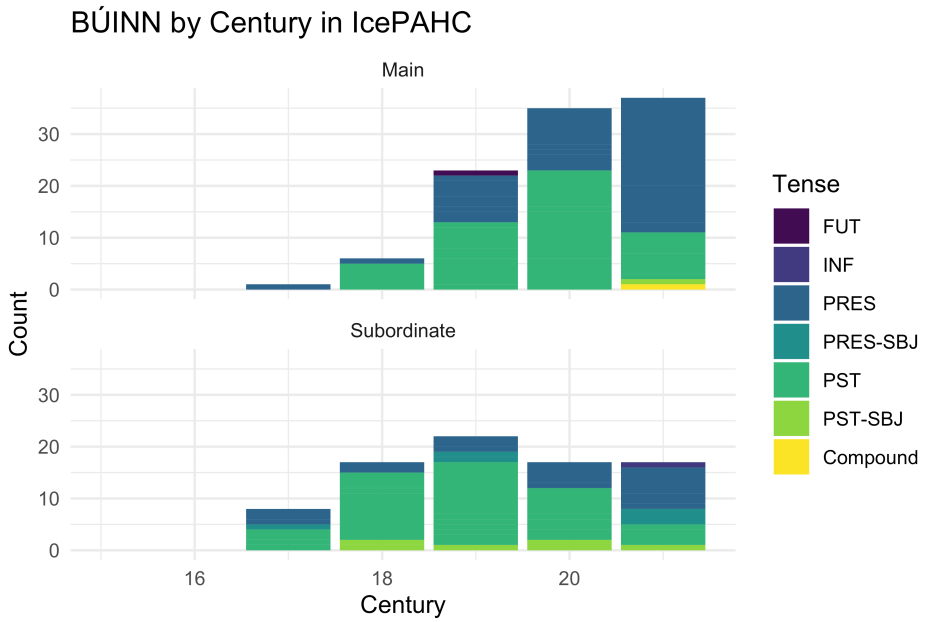


Figure 1: Clause-type/Tense by Century in IcePaHC

These fall together starting in the 14th century (ibid.).[19] Progressive readings provide an *internal perspective* of an event (Landman 1992). This corresponds to ongoingness, since the bounds of an event's inception or culmination are not taken into account.[20]

Bertinetto (2000) propose two fundamental types of readings for progressive constructions across Germanic and Romance: focalised and durative. The focalised type is characterised predominantly by the presence of a focalisation instant, which provides a vantage point (*As X occurred, I was Y-ing*): the event's ongoingness is viewed through a singular focal point. Durative progressives, by contrast, are evaluated relative to a larger temporal interval and are therefore often associated with adverbial phrases such as *during*, *until*, and *since* (Killie 2008: 76).

[19] An anonymous reviewer asks whether this is due to phonological or alternatively, phonetic fast-speech effects. As discussed by Benediktsson (2002), it is not entirely clear what the cause is. In any case, the doubled form is older (ibid.).

[20] An additional resource regarding the development of the progressive is Nordling (1928). Its status in modern Icelandic is discussed comprehensively in Jóhannsdóttir (2011).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

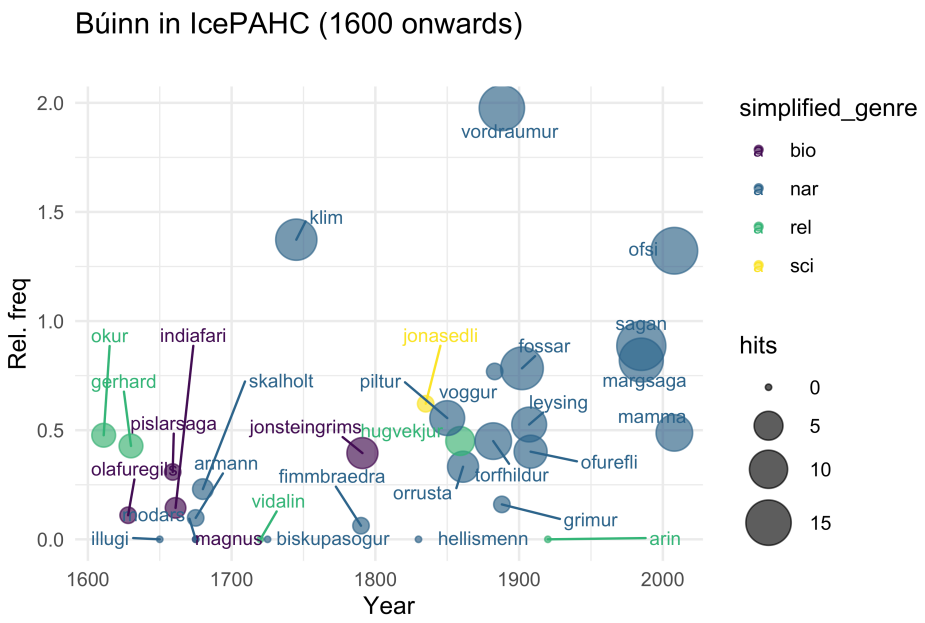


Figure 2: Time-course of relative frequency of *búinn*+Infinitive by text in IcePaHC, per 100 words. Anterior reanalysis occurs ca. 1600

An early example of an Icelandic progressive which is plausibly, though not definitively, focalised is shown below in (12). The wider context provides a singular focalisation point through which this event may be viewed, namely the man picking up a bundle of cloth. However, the focal point for this could also be the more specific time-frame given by 'when they were setting up the choir area' (*er þeir tjölduðu sönghúsið*).[21]

(12) *Menn voru að tjalda kirkju á einhverjum bæ.*
people were to decorate church at some farm
'People were decorating a church (with cloth) at some farm...' (But when they were setting up the choir area, a man picked up a bundle of cloth from the floor near the altar...) [**1210**.JARTEIN-REL.SAG,.2553]

Literature on the diachronic development of progressives cross-linguistically often assumes a progression from durative to focalised readings (Bertinetto 2000, Bertinetto et al. 2000). Looking at Germanic in particular, Killie (2014) (see also

[21]The verb *tjölduðu* is past tense. It refers to the practice of decorating the walls of a church with cloth (Kern 2019: 27).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

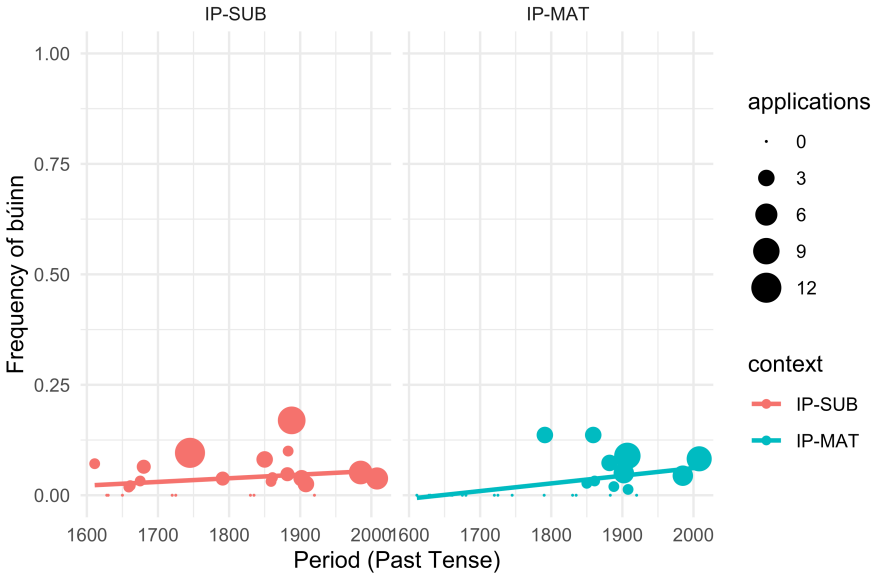


Figure 3: Pluperfect environments: *Búinn* and *hafa* in IcePaHC (past-marked)

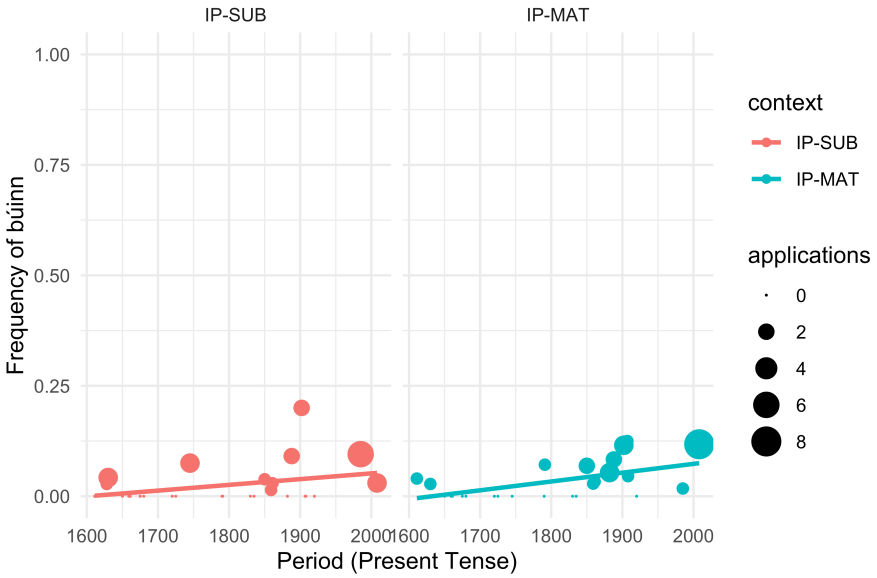


Figure 4: Perfect environments: *Búinn* and *hafa* in IcePaHC (present-marked)

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Killie 2008) provides a quantitative investigation the progressive in the diachrony of English through the Helsinki Corpus. The results in Killie (2014) demonstrate that, despite the fact that focalised progressives are more frequent than duratives in all time periods (Old, Middle and Early Modern English), one can observe an increase in the prevalence of focalised readings from one period to the next. This dataset is thus somewhat inconclusive: it is amenable to the proposal that duratives constitute an earlier stage, but this stage would have preceded the evidence in the corpus record. On the other hand, it remains possible that focalised and durative usages existed alongside one another from the earliest stages.[22]

As for the progression in the history of Icelandic, I carried out a preliminary corpus investigation focusing especially on the transitional stage of reanalysis. To do so, I investigated the portion of *Íslenskt textasafn* (16-18th century subcorpus) which includes the 16th century New Testament translation of Oddur Gottskálksson. This investigation was further supplemented by an assessment of the Old Icelandic examples discussed by Benediktsson (2002) in his article on the development of the periphrasis. We begin with the latter.

As we see below in Table 4, Old Icelandic differs from Old English in that durative examples are most common, with focalised usages coming in second. There are also a handful of indeterminate usages, which are somewhat ambiguous between durative and focalised. They do not clearly provide either a singular or multiple focal frame. Habitual usages are also represented in this corpus, as in Old English (cf. Killie 2008).

Table 4: Classification of progressive reading types in Old Icelandic examples listed by Benediktsson (2002)

Classification	Number of examples
Durative	15
Focal	10
Indeterminate (Indet)	8
Stative habitual	3

Moving now to the New Testament translation examples, I could find seven focalised usages, as opposed to one durative usage and two stative habitual us-

[22] A further complication is presented by the fact that in the Old English period include other usages, which appear to be non-aspectual. These include what Killie (2008) terms *narrative* progressives, which mark points of narrative climax for emphatic effect, as well as stative progressives, which denote "feelings, eternal truths, habits". Habitual progressives resemble the latter and are classified separately by Killie (2008).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

ages. In (13), a focalised example is provided. The progressive construction *en á meðan þeir voru að búa til* 'but while it was being prepared' sets up a focal frame for what follows.

(13) *Og er hann hungraði, vildi hann matar neyta. En á meðan*
and as he hungered wanted he food consume but when
þeir voru að búa til, leið yfir hann brjósthöfgi og sá
they were to prepare came over him breast:drowsiness and saw
himininn opinn og ofan fara að sér disk nokkurn mikinn svo
the:sky open and from.above go to himself disc some great so
sem línlak í fjórum hyrningum upp bundinn.
as linen:sheet in four corners up bound
'He became hungry and wanted something to eat; but while it was being prepared, he fell into a trance, and saw that the heavens were open, and that something like a great sail was descending, let down by its four corners towards the earth.'
(Acts 10; Nýja testamenti Odds, Íslenskt textasafn: Stofnun Árna Magnússonar í íslenskum fræðum 2023)[23]

Below in (14), a durative example is shown. As opposed to the focalised example above, there is no singular focal point through which the progressive-marked eventuality is viewed. Instead, the event in question (*þeir voru að tala um alla þá hluti sem við höfðu borið* 'they were talking about all those things that had taken place') is viewed as "relative to a larger interval" (Bertinetto 2000: 527). In this instance, the larger interval in question is the journey from Jerusalem to Emmaus.

(14) *Og sjá, að tveir af þeim gengu á þann sama dag til nokkurs*
and behold, that two of them went on that same day to some
kauptúns það er var frá Jerúsalem rúms sextigi skeiða, hvert
village DEM.N REL was from Jerusalem about sixty measures which
eð Emmahus var að nafni. Og þeir voru að tala um alla þá
COMP Emmahus was by name and they were to talk about all those
hluti sem við höfðu borið. Og það skeði þá þeir ræddust
things which had occurred and it happened when they conversed
við og spurðust á sín í millum að sjálfur Jesús
and asked themselves between that himself Jesus

[23]Translation from the Open English Bible (OEB Team 2022).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

> *nálægist þá og gekk jafnframt þeim.*
> approached them and went alongside them
> 'It happened that very day that two of the disciples were going to a village called Emmaus, which was about seven miles from Jerusalem, talking together, as they went, about all that had just taken place. While they were talking about these things and discussing them, Jesus himself came up and went on their way with them.'
> (Luke 24; Nýja testamenti Odds, Íslenskt textasafn: Stofnun Árna Magnússonar í íslenskum fræðum 2023)[24]

In conclusion, the corpus investigation at hand reveals the development of the progressive periphrasis in Icelandic does not neatly fit the path described by Bertinetto. Rather, it more closely resembles what Killie describes for the diachrony of English, in the sense that durative, focalised, as well as stative, non-aspectual usages are represented in the earliest stage of the corpus. As we move forward a couple centuries to the 1540 New Testament translation, we see that focalised usages predominate, in contrast to the earlier texts discussed by Benediktsson (2002). Further investigation is needed into the details of this trajectory, but we can nevertheless establish the fact that focalised usages became more entrenched in the language in the 16th century and beyond.

The IcePaHC corpus was queried for the progressive construction. Given the form of the construction, there is potential for false positives: combinations of the copula and infinitival marker show up in other environments (e.g. *Það er að segja* 'that is to say'). It was therefore necessary to manually sort out valid instances: I manually tagged tokens which could receive a focalised interpretation. The results are displayed below in Figure 5 (clause type and tense by century) and Figure 6 showing the time course of propagation.

As in the case of *búinn*, *PROG* remains relatively infrequent until the early 19th century. Focusing on our period of interest, what stands out is the higher frequency of subordinate instances in the 16th and 17th centuries. Though numerically the difference is not large, I would like to argue that these subordinate instances are especially important for the establishment of the focalised usage. This is due to the fact that temporal subordinate clauses provide the appropriate frame for focalisation. In summary, we see that despite the presence of modern-like focalised usages of the *PROG* construction as early as the 13th century, it does not become established in the language until much later.

[24]Translation from the Open English Bible (OEB Team 2022).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

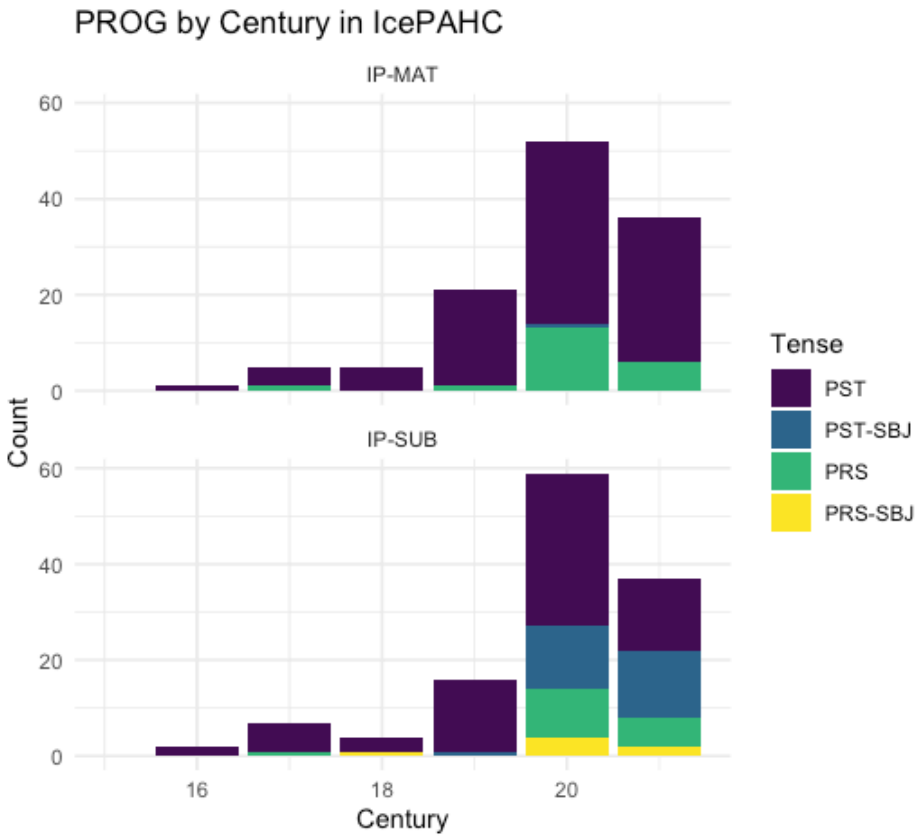


Figure 5: Clause-type/Tense by Century in IcePaHC

To better understand the status of the progressive in the early modern period, I also examined attested examples of the construction in the New Testament translation by Oddur Gottskálksson (1514–1556) in *Íslenskt textasafn*. Oddur, while a user of the focalised construction, does not use it in some instances where it would be obligatory in modern Icelandic. Helgason (1929: 135-136), for instance, wrote that Oddur uses the progressive "not quite as much as would be done today". Later authors represented in this corpus, such as Jón Ólafsson úr Grunnavík (*Nikulás Klím*) and Jón Steingrímsson (*Ævisaga Jóns prófasts Steingrímssonar*) in the 18th century both have a progressive in their grammar with a division of labour which resembles the modern usage. Future research is necessary in order to map out the diffusion of the progressive construction in the language, on the

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

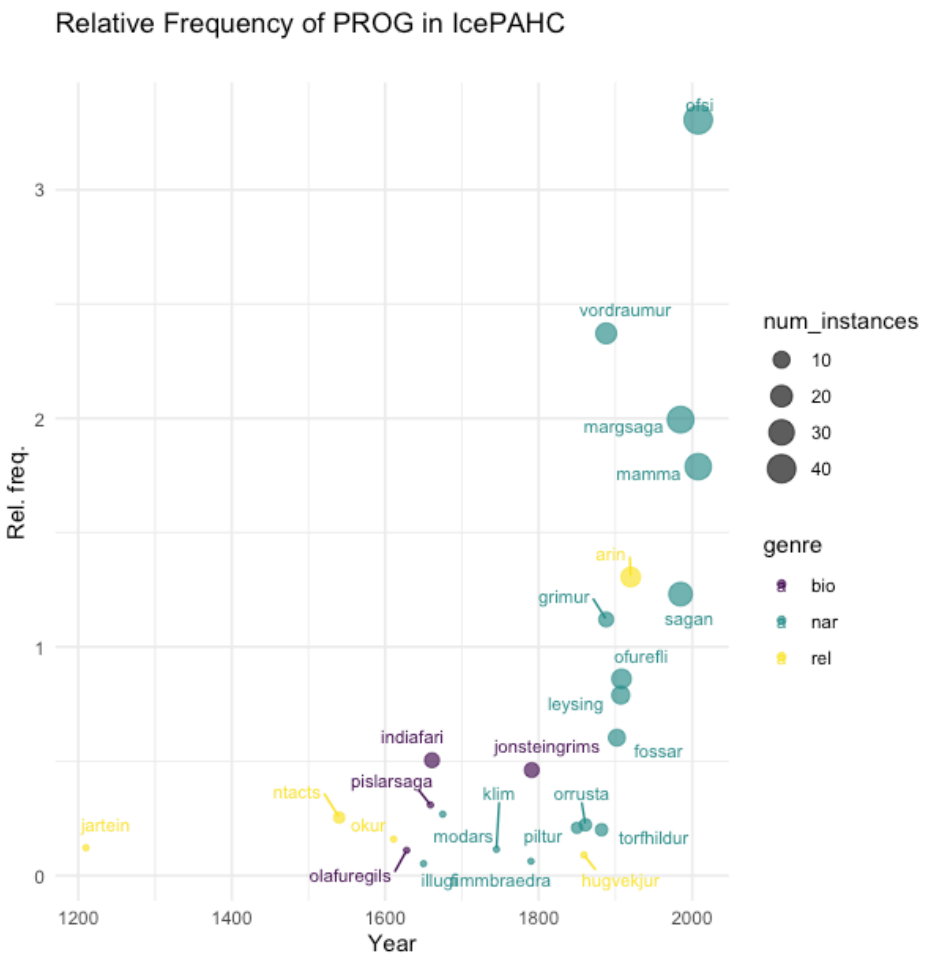


Figure 6: Time-course of *búinn*. Relative frequency over time in IcePaHC (by text).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

path towards obligatoriness in focalised instances. A classification of progressive instances in IcePaHC (manually tagged) broadly supports this view, as shown below in Figure 7.[25]

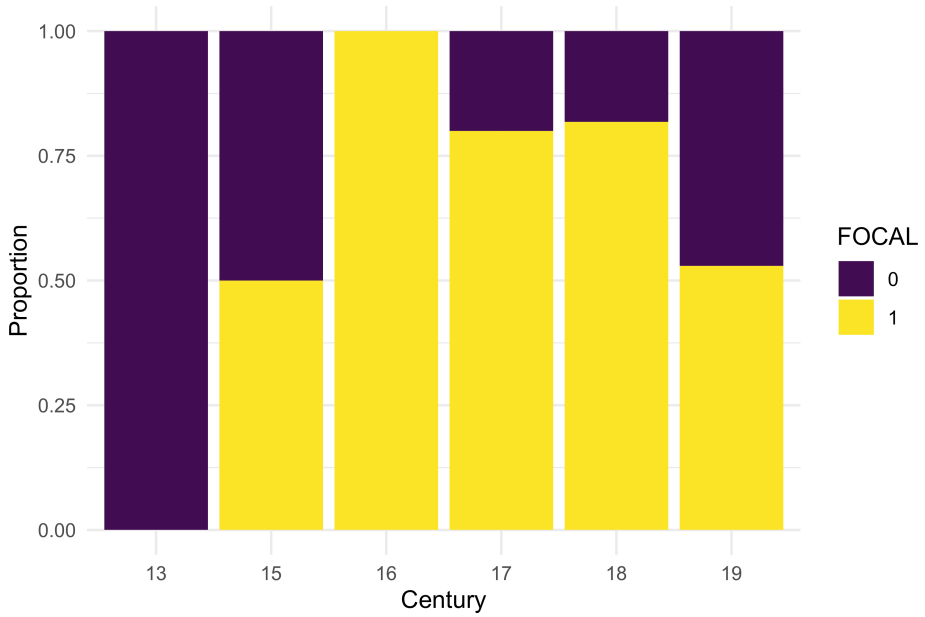


Figure 7: Progressives in IcePaHC prior to 1900 classified according to *focalisation* (n=68)

3 Tracing discourse traditions in the Icelandic context

In the following section, I argue that a shift in conventionalised strategies for structuring events and their overlap in narrative can be traced through the prominent discourse traditions of the early modern Icelandic period and that this shift is represented in our corpus evidence of the grammar of the language. Section 3.1 introduces the term *discourse tradition*. Section 3.2 discusses the outset for our investigation, the saga style narrative prominent in the corpus of Old Icelandic prior to 1540. Section 3.3 discusses the 17th and 18th century memoirs and travelogues which are an intermediate stage, I argue, with regard to the question of

[25]Note that only unambiguously focalised cases were counted, so the 13th century example discussed in Figure 12 is not classified as such.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

boundedness. Finally, Section 3.4 discusses the relationship between narrative traditions and orality.

3.1 On discourse traditions

Linguistic form has conventionalised associations with situations of language use, associations which are themselves always subject to change – these are what Koch (1987) calls *discourse traditions* (henceforth: DT).[26] As Kabatek (2022) spells out, changes can occur from within a certain discourse tradition and spread to others from there. The notion of DT as proposed by Kabatek (2022) is somewhat more expansive than the view espoused by Labov (1972: 123) conceiving of changes as either originating *from below* (common vernacular) or *from above* (prestige varieties). In the DT framework, language diachrony is modeled as consisting of developments parallel to one another temporally but differing in the degree to which they are conventionally tied to situations of language use.[27]

In what follows, I will advance the claim that discourse traditions surrounding Icelandic narratives, specifically how speakers construe events and their overlap in written language, underwent a shift between the 16th century and today.

A prominent discourse tradition of the Old Icelandic period is termed saga-style, as it is exemplified by style found in the *Íslendingasögur* ('Sagas of the Icelanders'). This style has a number of characteristics pertaining to the mapping between grammar and discourse structure: it is described by Hauksson & Óskarsson (1994) as being 'common'(*alþýðustíll*); 'learned' (*lærður*) characteristics are not prevalent.[28] The representation of events in saga style narrative is iconic: events are described in the order in which they occur. Relations between events are left largely underspecified. Thus, on the whole, texts of this sort are dominated by relatively brief matrix clauses, strung together paratactically (e.g. with the coordinating conjunctions *ok* 'and', *en* 'but', or non-coordinating, e.g. *þá* 'then', *nú* 'now', *síðan* 'since').

The way in which word order transformations map onto pragmatic functions resembles modern Icelandic, but there are some differences (Holm 1958, Hauksson & Óskarsson 1994). Two ways in which the older system differs from the modern one are with regard to the frequency of V1 and with regard to stylistic

[26]This discourse tradition model effectively adds a diachronic dimension to the notion of "register", which can be defined as "those aspects of socially recurring intra-individual variation that are influenced by situational and functional settings" (Lüdeling et al. 2022: 3).

[27]Kabatek (2022) assumes a continuum of situations from communicative distance to communicative immediacy (Koch & Oesterreicher 1985).

[28]'Common' refers to the language of the common people.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

fronting (movement of constituents to the left periphery which otherwise are base generated in the post-finite domain). Booth & Beck (2021) link both of these changes to the emerging status of a dedicated subject position. The authors (ibid.) also observe, in the IcePaHC corpus, a temporally concomitant decrease of non-subject initial V2 clauses – on the whole, it becomes increasingly dispreferred for subject topics to appear post-finitely: their corpus evidence (from IcePaHC) points to a period of considerable inter-textual variation around 1600, in which religious and biographical texts, as opposed to narratives, pattern in the opposite direction, which then gives way to a constant decrease in frequency.

For concreteness, I assume that both the progressive and the *búinn* perfect are first reanalysed prior to or at the early stages of the shift in boundedness. The circumstances that allowed for reanalysis (or supported the subsequent early entrenchment of the novel constructions) are independent of the shift in boundedness.

What is the relationship between changes in the language-system, on the one hand, and changes in norms surrounding narrative, on the other?[29] Language-specific properties have been shown to influence the way events are conceptualised and subsequently overtified linguistically: there exists experimental evidence for this in production/comprehension from monolinguals (von Stutterheim & Nüse 2003) as well as in acquisition (Carroll & Lambert 2003, Flecken et al. 2013). That is to say, if there exists a conventionally coded (grammaticalised option) distinction in a given language, it is more likely to be used than if only alternative, less conventionalised means, are available. Such work is couched within Slobin's (1994) 'thinking for speaking' hypothesis, which adopts the perspective that linguistic variation in terms of what is conventionally coded and that which is not serves as a kind of interface between the level of abstract conceptualisation and linguistic production (Levelt 1993). While "macro-planning" (ibid.) refers to the process where speakers determine what it is they want to say, "micro-planning" involves, among other things, the decisions a speaker makes in terms of intra- as well as inter-clausal relations, such as the use of referring expressions and whether to use coordinating or subordinating discourse relations (Asher & Vieu 2005).

There is, on the whole, considerable converging evidence regarding language-specific properties and their consequences for information packaging from cross-linguistic comparison in monolinguals (e.g. Starren 2017), acquisition (e.g. Schmiedtová et al. 2011), bilingualism (e.g. Daller et al. 2011) and in historical-typological research (Starren 2017). To summarise: the available evidence in the literature

[29]Thanks are due to an anonymous reviewer for prompting me to elaborate on this point.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

appears to support a view according to which language-specific properties, defined by which conceptual distinctions receive conventional encoding, results in a given language's propensity for bounded or unbounded construals.

As touched on above, I assume that certain discourse traditions can become conventionally associated with certain linguistic co-occurrence patterns (Kabatek 2022) (lexical and grammatical), which I assume occurs via a process of entrenchment (as conceived of by e.g. Schmid 2015). Saga style carries with it conventionalised associations to co-occurrence patterns of lexical, grammatical and structural characteristics.[30] Discussing the structure of the Icelandic family saga, Andersson (1967: 33) writes: "There is no such thing as a digression [...] All the episodes are linked in a sequence leading up to the climax or leading down from it. This is a fundamental rule and is the key to saga economy". The highly paratactic nature of saga-style is characterised by relatively minimal temporal backtracking or divergence from the main narrative storyline: "even with the most artless story-telling in the paractic style, however, there will always be greater or lesser degrees of syntactic divagation from the straight line of narrative events" (Amory 1980: 392). Such divergences, in narrative generally, serve evaluative functions (cf. Labov & Waletzky 1967: 26-35) and as argued by Amory (1980), saga writers used a number of stylistic devices to insert their evaluation into the narrative despite being "notoriously non-committal about their opinions" (Amory 1980: 394). The paratactic style is well-suited for such non-committal, minimally evaluated narrative.

"Saga style" refers here to what is termed "popular style" and typically contrasted with "learned style", the latter abounding with Latinate characteristics (Nygaard 1905, Kristjánsson 1985, Hauksson & Óskarsson 1994, Smári 1920). Latinate characteristics, as touched upon earlier on in this paper, include the appositive use of present participles and absolutive constructions (*dativus ablativus*), both of which represent devices for communicating temporal sequentiality and overlap, with these specific grammatical devices presumably being favoured due to translation influence (cf. Brendel 2023). This distinction between learned characteristics on the one hand and popular ones on the other thus provides an avenue for characterising the predominant discourse traditions in Icelandic prior to the 16th century. Already at this stage, we observe divergences between these traditions in the way in which events are structured. This seems to be shaped by the available grammatical means; translation influences led to the use of constructions which were presumably less frequent/conventionalised in the spoken

[30] As pointed out by an anonymous reviewer, saga style is typically conceived of as a written norm, which is partially what enabled the revival of some of its characteristics in modern Icelandic narrative style.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

vernacular but rendered more salient due to the influence of Latin, which in turn could impact the subsequent development of the language (within this particular discourse tradition and beyond, cf. Wagner et al. 2013 on scribes as agents of language change).

Our picture of the diachronic developments of a language are inevitably skewed by available corpus evidence. For Icelandic, the evidence furnished for different periods is shaped by dominant discourse traditions including the influence of conceptual orality/writtenness (cf. Koch & Oesterreicher 1985). The prevalence of aspectual expressions is, for instance, often tied to narratives in contrast with expository texts (Ragnarsdóttir et al. 2002). Overall, I think the corpus evidence has to be carefully considered in tandem with the socio-historical processes that shape discourse traditions.

The diachrony of Icelandic provides a powerful case study for the role of emerging norms and cultural processes in shaping discourse traditions. The modern literary standard begins to take shape in the 19th century with a concentrated effort to bring the language closer to the narrative style of the sagas (Bernharðsson 2017, 2018). Having now given some background on the norms surrounding written narratives for Icelandic, we now proceed to discuss narrative structure in more detail. This lays the groundwork for the discussion in Section 4, where I argue that Icelandic displays a shift towards unboundedness between the 16th and 19th centuries.

3.2 On narrative structure

Narratives can be segmented into episodes, which are a level of textual organisation situated between the smaller sentence-level and larger discourse-level (van Dijk 1976, Enkvist & Wårvik 1987, Fleischman 1990). These episodes are characterised by thematic or temporal coherence and episode transitions thus correspond to transitions of location, participants, perspective, etc. Episode transitions are commonly marked linguistically (Brinton 1996, Wårvik 1995): for instance by spatial or temporal adverbials or tense alternations.

Narratives in the saga style have a formulaic structure, often linked to their oral origins (Hauksson & Óskarsson 1994). Narrative episodes often begin, for instance, with formulaic introductions and scene setting, which then transitions into the description of a series of events (Clover 1974). Tense alternations are very common: past tense alternates with historical present, the latter giving the sense of the primary focus of the narrative episode coming more clearly into focus (Fleischman 1985).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

The way in which episode boundaries are marked overtly relates to the notion of grounding (Hopper 1979). The foreground is most basically speaking comprised of those events which comprise the main storyline – everything else is background. On the whole, one can distinguish between the linguistic expressions most associated with backgrounding and foregrounding, though these are non-deterministic correlations (Brinton 1996: 49). Adverbial clauses, our primary environment of interest, are frame-setters (Talmy 1978) – they signal continuity in a "temporal line" (Fleischman 1985).

The two expressions in question, *búinn* and *PROG* pattern with other imperfective elements in that they communicate temporal overlap and serve to reorient the reader to the narrative timeline (Fludernik 2012). They are backgrounded in the sense that they do not advance this timeline but rather have the function of providing "guideposts" for the reader (Chafe 1984). However, they differ from prototypically backgrounded elements in that they can provide new information (Reinhart 1984, Fleischman 1985) and are therefore well-suited to beginning a novel narrative episode. In the case of progressives, their focalised usage most clearly has this function, setting up a new temporal frame, which establishes a new topic time (in the sense of Klein 1994) as a vantage point through which subsequent events are viewed (see, e.g. the example in 23).[31] Likewise, the *búinn*-perfect often functions as a frame-setting strategy where it appears in temporal adverbial clauses, which typically serve to bring the reader up-to-date, setting the stage for what can subsequently occur (see, e.g. 21).

3.3 17th and 18th century memoirs and travelogues

Travelogues and memoirs of the 17th-18th cent. combine so-called *lærður stíll* '' and *alþýðustíll* 'common/popular style' (Hauksson & Óskarsson 1994).[32] These texts contain dialogues and indirect speech reminiscent of earlier styles, specifically saga writing. Direct event descriptions are often paratactic and iconic. Characteristics associated with the learned style are also present, however, including expository passages (usually interpreting events described) with high levels of subordination (ibid.). These travelogues and memoirs additionally exhibit influence from Low German and Danish due to language contact. This influence

[31]This is in accordance with the definition of frame-setters given in Krifka (2008: 270): "the frame-setter indicates that the information actually provided is restricted to the particular dimension specified." Focalised progressives (e.g. in *when*-clauses) restrict the following clause on the temporal dimension.

[32]Texts in IcePaHC with such characteristics include: *Nikulás Klím*, *Ævisaga Jóns Ólafssonar Indíafara*, *Reisubók séra Ólafs Egilssonar* and *Píslarsaga Jóns Magnússonar*.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

has been under-researched and is not well understood, though it is undeniably strongly present in texts of this genre at all linguistic levels (e.g. syntax and lexicon) (Ottósson 1990, Hauksson & Óskarsson 1994, Sigmundsson 1998).

It is precisely these texts, as far as our corpus dataset is concerned, that constitute the bridge between saga-style narratives and the emerging modern literary standard, as, for instance, reflected in Jón Thoroddsen's novels (Bernharðsson 2018). These texts are also the first in the corpus to exhibit frequencies of *PROG* and *búinn* which exceed 0.25% (overall relative frequency).

3.4 Adverbial clauses and orality

As we move forward, we will see that as the predominant narrative style changes, there is a consequent shift from a mainly iconic, strictly chronological representation of events, on the one hand, to other, new strategies for construing of event sequentiality as reflected in clause structure, namely overt grounding by means of (predominantly preposed) temporal adverbial clauses. Contrasting the two poles, hypotaxis is less dependent on the inference of causal and temporal relations, as these are now overtified. Saga style is quite sparse in these terms. Haiman (1985: 39-71) claims that there can be a kind of trade-off between linear order and other mechanisms which convey the way in which clauses relate to one another. This is roughly the line of analysis which I pursue here. In a predominantly unbounded system, the functional domain previously occupied by local anchoring adverbials (by means of tail-linking: Los 2012) gives way to new strategies: backgrounded, temporal clauses situate events within a global topic time.[33]

Wårvik (1995) and Brinton (1996) posit that an analogous situation in OE – from main clause constructions to subordinate *whan* clauses – constitutes such a shift.[34] Brinton (1996: 179) argues that an increase in subordination reflects the higher degree of "integration" reflected in (conceptually) written language, thus correlating with a degree of plannedness (cf. "thinking for writing" in the terms of Slobin 2003).

Corpus work on register variation broadly supports this general view but the picture is not straightforward: adverbial subordination is characteristic of written expository registers in PDE, however it also correlates with the 'interactive'

[33]The bounded/unbounded distinction as conceived of by Carroll & Lambert (2003) must be conceived of as a gradient of preferred ways in which to structure narrative. It is not absolute nor is it assumed that all situations of language use behave identically in terms of preference for boundedness.

[34]Middle English *whan* corresponds to PDE *when*.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

dimension of oral communication (Biber & Conrad 2005). Walkden (2021) investigates the ratio of hypotaxis to parataxis in IcePaHC. He finds that saga texts exhibit most parataxis on average. Religious and non-fiction texts exhibit the highest degree of hypotaxis, especially in the 17th-19th centuries.

4 From boundedness to unboundedness

4.1 Motivation behind approach

The main idea behind the approach pursued in this paper is as follows: A fundamental shift in mapping between clause and discourse/information structure occurred in the diachrony of Icelandic, with consequences for the the expression of aspect. To date, scholarly work in this area has focused on the syntactic reflexes of these changes (Holmberg 2000, Sigurðsson 2018, Hróarsdóttir 2000, Booth 2018, Booth & Beck 2021). In what follows I will show that the emergence of *búinn* and *PROG* was a consequence this broader shift in mapping to clause structure.

The syntactic reflexes of this shift are often discussed together in the literature, especially in reference to the relationship between the pre-finite position and subjecthood (Jónsson 1992, Barðdal & Eythórsson 2003, Booth & Beck 2021). These include a decrease in *Stylistic Fronting* (when, in clauses with a gap, non-subject XPs are fronted to the pre-finite position) (Holmberg 2000), a decrease in V1 declaratives (Sigurðsson 2018, Booth & Beck 2021), as well a decrease in argument drop, accompanied by an increase in the expletive element *það* (Rögnvaldsson 1995, 2002, Booth 2018). Booth (2021) argue that this shift can be understood at least partially in terms of the emergence of SpecIP as a dedicated subject position; the inflection point for this change is determined to be between 1500-1700 (based on IcePaHC).

What is novel in the present approach is its ability to explain otherwise puzzling facts, such as the time-course of propogation of these aspectual constructions. I argue that these can be better understood if one takes the shift in discourse structuring as fundamental. This shift led to fertile ground for emerging aspectual constructions to conventionalise.

How can the shift be characterised more broadly, moving beyond the syntactic phenomena more often discussed? My claim is that the shift in mapping between form and discourse structure in the history of Icelandic represents, more broadly, a shift from a predominantly "bounded" system for expressing narrative to a predominantly "unbounded" one (for related work on the diachrony of English, see Petré, 2010, 2014, Los, 2012, Fanego, 2024).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

4.2 (Un)Boundedness

What is the notion of "boundedness" that I use here? von Stutterheim et al. (2012, 2002) propose, on the basis of psycholinguistic experiments, an account of how the grammatical system can constrain how events are expressed in narratives: Bounded language invokes a series of completed events strung together one after the other in a sequence, while unbounded language views events in terms of their co-articulation within an overarching topic-time. Bounded construals thus broadly correspond to perfective/telic eventualities and unbounded ones to imperfective/atelic (Kupersmitt 2015).

Contrasting modern day English and German, the former seems to prefer unbounded construals, and the latter bounded ones, for constructing narrative sequences (von Stutterheim & Lambert 2005, Berman & Slobin 1994). There are other linguistic reflexes of boundedness which are discussed in this literature (von Stutterheim & Lambert 2005, Kupersmitt 2015): German, for instance, makes heavy use of topic-time adverbials and scene-setting on the left periphery which allows for anaphoric continuity. On the other hand, English makes little use of such topic-time adverbials, instead the determination of topic time is often context-dependent and there is more focus on overtly marking the co-articulation of events within the interval provided by topic time.

Older Germanic more closely resembles modern German than English. The linguistic reflexes present in these languages include heavy use of topic time adverbials and word order transformations for anaphoricity/discontinuity (Los 2012). The transition from bounded-unbounded system in the history of English is reflected in the loss of these topic-time adverbials, and the loss of V2 (ibid.).[35] An increase in the frequency of aspectual constructions could be observed: the progressive *BE -ende*, the ingressive aspectual verb *beginnan* (Petré 2010, Los 2000). Alongside these latter changes one also observes an increase in preposed *whan*-clauses (Brinton 1996: 169-172). The relevance of work on English for the Icelandic data is discussed later in this paper (Section 5.1.1). We turn now to the data from successive stages of Icelandic, demonstrating the language's shift from a bounded to unbounded system.

[35]Note, however, that there are alternative perspectives on the loss of V2, e.g. Haeberli (2002) who reduces the change to a consequence of a loss of inflectional morphology, or Speyer (2010) who proposes a primary role for prosodic factors. I thank an anonymous reviewer for making this point.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

4.2.1 Old Icelandic as a bounded system

In this section, I argue that Old(er) Icelandic (prior to 1540) is a language which exhibits a strong preference for bounded construals, in the sense of von Stutterheim et al. (2002). In order to discuss the ways in which this is reflected in the grammar, I begin by describing the basic characteristics of Old Icelandic grammar. Icelandic is at all stages a strictly V2 language. The question of what may be considered the basic or unmarked word order in Old(er) Icelandic is not entirely straightforward, as SVO and SOV co-exist for many centuries (Sigurðsson 1988, Rögnvaldsson 1994, 1995, 1996) – the latter only disappears entirely in the 19th century (Hróarsdóttir 2000). Icelandic at all stages makes heavy use of transformations for pragmatic purposes, including operations such as topicalisation, extraposition, and verb-initial orders (Faarlund 1990, Booth 2021). Old(er) Icelandic can in fact be considered to have Topic-Verb-Subject word order (Nielsen 2017).

Verb-first orders are used for discourse cohesion in order to express anaphoric continuity (Booth & Beck 2021), often termed "Narrative Inversion" (Platzack 1985). In narrative sequences, this results in continuity along the main line of narration.[36] There is additionally an argument to be made that the pragmatic effect of V1 is preserved even when the surface word exhibits V2; this is the case in the presence of the coordinating conjunction *ok*[37] (which may also resemble an adverbial, see Nielsen 2017) as well as the adverbial *þá*. For the purposes of this paper, I treat V1 clauses as parallel to those introduced with *ok* and *þá* as far as their relationship to the main narrative line is concerned: they are foregrounded and move the narrative timeline forward temporally (Enders 2001, Donhauser et al. 2006).[38]

Further evidence for the status of Old(er) Icelandic as having been predominantly bounded comes from Leiss (2000) who discusses V1 as covertly signaling aspectual value: perfectivity. Leiss builds on previous literature (e.g. Rieger 1968, Kossuth 1980) which develops the view that V1 is used to convey successive events, such as in battle descriptions found so commonly in the saga literature. Leiss (2000: 95) argues that certain syntactic contexts can coerce a perfective interpretation where the underlying verb is ambiguous (aspectually polysemous in her terms). These contexts include: V1, historical present, proximity to definite articles and the presence of perfective adverbials (cf. Richardson 1995).

[36] On the relationship between narration and topic anaphoricity in the diachrony of Icelandic, see Booth & Beck (2021: 20). Relatedly, Kinn et al. (2016) discuss the connection between V1 and null subjects.

[37] Old Icelandic *ok* corresponds to modern Icelandic *og*.

[38] For work on discourse adverbs with similar function in the diachrony of English, see van Kemenade & Los (2006). For Old High German, see Trips & Fuß (2009).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

The following passages are taken from *Finnboga saga ramma* (mid 14th century, present in IcePaHC) and serve to illustrate the core characteristics of saga style narrative. The passage below sets the scene for what is to come: a tail-linked topic-time adverbial *um sumarið* 'during the summer' provides a local anchor for what follows. The second sentence is V1, maintaining that the event directly succeeds the one in the first sentence. Here a perfective interpretation is most natural for the verb phrases in both sentences, though one is in historical present and the other is past-marked V1, in accordance with Leiss' claims.

(15) *Um sumarið býr Bergur skip sitt og ætluðu þau að sigla.*
during summer prepare Bergur ship his and intended they to sail
Lætur Finnbogi flytja varnað þeirra til skips.
lets Finnbogi move goods their to ship
'During the summer, Bergur prepared his ship and they intended to sail. Finnbogi had their goods moved to the ship.'
[**1350**, FINNBOGI.NAR-SAG, 1851-4]

The passage which follows is similar in structure; the adverbial phrase *og er Bergur er búinn* 'and when Bergur is ready' establishes a local topic-time. This is followed by a V2/V1 alternation in the next two sentences, indicating a shift in subject-topics. The concluding V1 sentence is most naturally construed perfectively: their riding has a concrete endpoint.

(16) *Og er Bergur er búinn ríða þeir Finnbogi og Þorkell með honum.*
and when Bergur is BÚINN ride they Finnbogi and Þorkell with him.
Hrafn hinn litli fór og með nogkura klyfjahesta. Ríða þeir þar
Hrafn the little went and with some pack:horses. Ride they there
til er þeir koma mjög svo vestur af hálsinum Hrútafjarðarhálsi.
until when they come very so west off the:ridge Hrútafjarðarháls.
'And when Bergur is ready, they, Finnbogi and Þorkell, ride with him. Hrafn the little went and [took] with him some pack horses. They ride there until they come very far west off the ridge of Hrútafjarðarháls.'
[**1350**, FINNBOGI.NAR-SAG, 1855-7]

A few passages later, following an exchange in direct speech, the storyline continues in a typical manner with the local topic-time anchor *þá* 'then' followed by a chain of perfective events marked in past tense, expressed paratactically with the coordinating conjunction *og* 'and'.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(17) *Þá hljópu fram tveir menn með vopnum og tóku hestana og*
then jumped out two men with weapons and took the:horses and
leiddu upp undir brekkuna.
led up under the:slope
'Then two men jumped out with weapons, took the horses, and led them up under the slope.' [**1350**, FINNBOGI.NAR-SAG,.1863]

In sum, we have seen that narrative in the saga style exhibits the characteristics of a bounded system: tail-linking (use of local topic-time adverbials at each discourse turn, cf. Los, 2012) is prominent; sequences of events are construed perfectively as though in a continuous chain of self-contained events which move the timeline forward.[39]

4.2.2 17th–18th century Icelandic as an intermediate system

In Section 3.3, I introduced the linguistic characteristics of 17th and 18th century memoirs and travelogues, which I argue constitute an intermediate stage in Icelandic's shift from a predominantly bounded to a predominantly unbounded system. To illustrate this, I draw on some passages from the memoir of Jón Ólafsson *Indíafari* (1661) (Ólafsson 1908) and the *Nikulás Klím* translation by Jón Ólafsson of Grunnavík (1745) (Holberg 1948).[40] We begin with an excerpt from Jón Ólafsson Indíafari: a scene-setting passage which establishes a global topic time (a Saturday in 1617 in Copenhagen). Notable in this passage is the use of extensive hypotaxis as well as the use of the progressive in the final sentence *og voru að spássera þar um grundirnar* 'and were strolling around the fields...', which does not move the global topic time forward.

(18) *Einn tíma bar svo við 1617 um sumarið, á einum*
one time happened so in 1617 during the:summer, on one
torgddegi sem var laugardag, að einn maður, Jens að nafni, sá eð
market:day which was Saturday, that one man, Jens by name, who rel
var nýgiftur einnri frómri dándiskvinnu og bjó í Litla
was newly:married a.DAT pious lady and lived in Little
Ferjustræti í Kaupinhafn, var út genginn fyrir borgina um
Ferry:street in Copenhagen, was out walked outside the.city through

[39] "Prominent" is a relative term; the decrease in tail-linking is discussed in quantitative terms in 5.2.3 below.

[40] A large excerpt of the Klím text is present in the *Textasafn* corpus, including pages 1-251.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Vesturport með einum frómum borgara og voru að spássera þar
West:gate with a distinguished citizen and were to stroll there
um grundirnar, sem bændurnir skyldu fram um fara, þeir
around the:fields, which the:farmers should forward around go, they
eð áttu heima uppá landsbyggðinni
who lived up:in the:countryside

'One market day, a Saturday, in the summer of 1617, it happened that a man, by name Jens, who had recently married an excellent lady and lived in Lille Færgestræde in Copenhagen, had left the town by the Western Gate with a worthy citizen. They were taking a walk on the plain crossed twice a week by the farmers who live in the neighbourhood, on their way to the market-place in the town.' (Ólafsson 1908: 38)[41]

The above passage continues as shown in (19), exhibiting both unbounded and bounded characteristics. The passage begins with an inchoative construction *upphóf að* 'began to', situating the event of Jens pelting the farmers with horse dung within the global topic time of their stroll around the fields. A sequence of bounded events begins once it is specified that a certain farmer is struck by Jens' insults: *En með því að hann mjög svo kenndi undan hans áköstum* 'Now as he was hurt by the pelting...'. This is the moment of inception for the farmer to react, establishing a new topic time at which the farmer jumps forth to attack Jens.

(19) *Þessi ungi borgari, Jens að nafni, upphóf eftir gamalli*
DEM.NOM.M young citizen, Jens by name, began according_to old
venju að grýta bændurna með hestaperum, þar til þessi drukkni
custom to mock the:farmers with horse:dung until this drunk
bóndi varð fyrir hans kasti, hrópi og spottyrðum. En með
farmer became for his throws shouts and insults but with
því að hann mjög svo kenndi undan hans áköstum og þar með
DEM.DAT.N that he very so felt from his attacks and with that
illa sveið hans hæðni, og var í þessu óaflátsamur og með engu
badly stung his scorn and was in this unyielding and with no
móti hirti um hins annars borgarans umtölur né afletjan,
way heeded of DEM:GEN other:GEN citizen:GEN words nor distraction
hljóp úr vagninum með ryðugan korða, sem bændur plaga
jumped out.of the:cart with rusty sword.M which farmers tend

[41]Translation from Ólafsson (2017: 95)

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

þar án bals liggja láta, og strax hann með honum í
there without sheath lie let and immediately he with he.DAT
gegn nísti.
against struck
'The young citizen, Jens, according to his time-honoured custom, set to pelting the farmers with horse-dung, and finally the drunken farmer became the object of his aim and taunts. Now as he was hurt by the pelting, and at the same time smarted sorely under the taunts, with which the other moreover continued to ply him, paying no heed to the advice and dissuasion of his fellow-townsman, the farmer leapt down from his cart with a rusty sword such as farmers use to have lying in their carts without a sheath; ran him through with it...' (Ólafsson 1908: 39)[42]

The narrative continues with a series of bounded events reminiscent of saga-style. A series of perfective events are construed in sequence, moving forward the narrative timeline: the farmer leapt, Jens fell dead, the news is carried into the city, Jens' body is collected and buried at St. Nicholas cemetery. The alternation between V1 and V2 in the passage in (20) resembles saga texts: an anaphoric subject topic is maintained across the V1 clauses and V2 is used to signal a shift in this topic (*Hljóp svo... Hinn féll...*). These events take place in a bounded sequence each moving forward the local topic time.

(20) *Hljóp svo á sinn vagn og burt keyrði. Hinn féll þar strax*
jumped then on his cart and away drove DEM.M fell there immediately
dauður niður til jarðar. Borgarinn hinn annar bar þessi tíðindi
dead down to ground the:citizen DEM.M other brought these news
í borgina. Var hann svo sóttur og inn fluttur og að morgni
into the:city was he then fetched and in brought and at morning
greftraður uppá Sankti Nikulás kirkjugarði.
buried up:on Saint Nicholas cemetery
'He leapt on to his cart again and drove away. Jens at once fell dead. The other citizen carried this news into the city, and the body was fetched and conveyed within, and buried the next morning in St Nicholas' churchyard.' (Ólafsson 1908: 39)[43]

Next, in (21) below, we turn to Nikulás Klím and examine a passage which describes the protagonist being medically investigated by a tree – the story is science-fiction and takes place in an underworld of sorts.

[42]Translation due to Ólafsson (2017: 95-96).
[43]Translation from Ólafsson (2017: 96)

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(21) *Meðan eg á þessa lund talaðe við sjálfann mig í einrúmi, geck*
while I in this manner spoke to self me in alone, walked
ein eik inn til mín, sem hafðe einn triangel (þríhyrníng) í
ART.INDF oak.F in to me, who had a triangle (tri:corner) in
hendinne. Eik þesse berađe minn handlegg og brjóst; sló hún
the.hand oak.F this touched my forearm and chest; struck she
rjett lagliga miðæðina með verkfære þessu, og þegar hún
precisely properly the:middle:vein with tool this, and when she
hafðe látið renna nockut blóð, so mikið sem henni sýndest þurfa, þá
had let flow some blood as much as her seemed to.need then
batt hún til arminn aptur með eigi minni lagvirkne. Þá hún var
tied she up the:arm again with not less skill when she was
þannin búin að útrjetta sitt erinde, og hafðe skoðat blóðit
thus BÚINN to accomplish her task and had examined the:blood
þegjande og forundrande, geck hún aptr sinn veg í burt.
silently and wondering went she back her way out

'While I mused on the strange things I had witnessed, a tree came into my cell, with an instrument resembling a lancet in his hand. It stripped one of my arms, and made a puncture in the median vein. When he had taken from me as much blood as he deemed sufficient, he bound up the wound with great dexterity. He then examined my blood with much attention, and departed silently, with an expression of wonder.'

(Holberg 1948: 29-30)[44,45]

From the perspective of boundedness, this passage begins with the establishment of a global topic time, namely that which occurs while the narrator is busy musing on what they had witnessed thus far. The entrance of the tree and its subsequent actions constitute events on the main time-line, situated within this global topic time. Temporal subordinate clauses beginning with *þegar* and *þá* are used to situate events within the topic time, without moving forward the global narrative timeline. This sequence of events is then summarised with a *búinn* construction, the material embedded under which is backgrounded in relation to the main timeline. Its function is to recapitulate the relation between this sequence of events and what comes after: having done his job, the tree is now in a position to leave. The passage above in (21) is followed by a sequence of expository

[44]Translation by John Gierlow sourced at Project Gutenberg (Holberg 1845). Page numbers are absent since reference is made to the digital edition.

[45]It should be mentioned that the noun *eik* 'oak' is feminine and thus glossed as such. In the subsequent discussion I refer to the tree as 'it', following English convention.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

passages detailing the circumstances which led to him being investigated in this strange new world, reflecting on on his understanding of the workings of this society. In the below passage, the *búinn* construction is used twice, again in order to summarise the sequence of events up until this point: the protagonist emphasises that had he not learned the subterranean language, his opinion would be quite different.

(22) *Þetta styrkte enn fremur meiníngu mína, sem eg hafðe*
DEM.NOM.N strengthened even further belief my which I had
feingið um fávitsku fólks þessa. Enn epter þat eg var
received about foolishness people.GEN DEM.GEN but after that I was
búinn at læra underjarðarmálit, og allt var búit at
BÚINN to learn the:underground:language and everything was BÚINN to
segja mier, þa snerest allt mitt forackt í forundran.
tell me, then turned all my contempt into wonder
'This circumstance by no means weakened the opinion which I had for some time entertained, that these people were shallow and foolish. But my judgment proved to be too hasty. When I was better enabled to judge of what passed about me, by acquaintance with the subterranean languages, my contempt was changed to admiration.'
(Holberg 1948: 30)[46]

In summary, travelogues and memoirs such as *Nikulás Klím* and *Æfisaga Jóns Indíafara* clearly exhibit both bounded and unbounded characteristics. The authors make use of unbounded strategies such as *búinn* and *PROG* alongside adopting stylistic characteristics favouring bounded construals from the earlier saga style.

4.2.3 Modern Icelandic as unbounded

I this subsection, I examine an excerpt from the novel *Piltur og stúlka*, written by Jón Thoroddsen and published in 1850 (Thoroddsen 1850). The novel was highly influential at its time of publication; Bernharðsson (2018) argues that it was an important ingredient in the establishment of a new linguistic standard in the 19th century. I therefore consider it to be at the dividing line between early modern and modern Icelandic.[47]

[46]Translation from Holberg (1845).

[47]The novel exhibits some archaicisms which hearkens back to older narrative styles, e.g. OV word order (Bernharðsson 2018). Nonetheless, it is in my view a representative example of narrative at this stage in the language.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(23) *Það var eitt kvöld nokkru eftir miðjan vetur, að þær Sigríður*
it was one evening sometime after mid winter that they Sigríður
og Guðrún sátu tvær einar í stofu. Veður var fagurt,
and Guðrún sat two alone in living_room. Weather was beautiful
sólin var að setjast, og kvöldroðanum kastaði á gluggana og
the:sun was to set and the:evening:glow cast onto windows and
inn um stofuna.
in through the:living_room
'One evening, not long after mid-winter, Sigrid and Gudrun were sitting alone together in the living-room. The weather was fine, the sun was just setting, and the evening-red fell upon the windows and the floor of the room.' (Thoroddsen 1850)[48]

In passage (23), we observe little in the way of tail-linking adverbials. Instead, the hallmarks of an unbounded narrative are present. A topic time is first established via the scene-setting construction at the inception of the paragraph, during which other events are ongoing. The narrative continues as follows:

(24) *Guðrún sat á stóli út við gluggann og var að sauma og*
Guðrún sat on chair out by the:window and was to sew and
þagði; sá, sem þá hefði séð hana og tekið eftir
was_quiet; DEM, which then have.SBJ seen her and notice
brosunum, sem voru að smáflögra um munnvikin á henni, og
the:smiles which were to flutter around mouth:corners on her and
séð hvernig spékopparnir á kinnunum á henni ýmist voru að
seen how the:dimples on the:cheeks on her either were to
myndast eða hverfa.
form or disappear
'Gudrun sat silently sewing in a chair by the window. Whoever could have seen her then and observed the smile which played (*was fluttering)* upon her lips, and the dimples in her cheeks, as they came and vanished (*were coming and vanishing*).' (Thoroddsen 1850)[49,50]

In example (24) above, progressive constructions are prominent *voru að smáflögra* 'were fluttering' and *voru að myndast eða hverfa* 'were forming or disappear-

[48]Translation due to Arthur Reeves (Thoroddsen & Reeves 1890: 164).
[49]Translation from (Thoroddsen & Reeves 1890: 164).
[50]My additions to highlight where the progressive is used; the translator did not always reflect it in the English.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

ing'. These describe unbounded events within the global topic time, as established by the expletive construction at the onset of the narrative episode (this one evening after mid-winter). The narrative continues with Sigríður singing lullabies to baby Sigrún, giving way to the following passages:

(25) *En er Sigrún litla var sofnuð, lagði hún hana hægt í*
but when Sigrún little was asleep laid she her slowly into
legubekkinn og lítinn kodda undir höfuðið og breiddi svuntuna sína
the:crib and little pillow under the:head and spread the:bib hers
ofan á hana, en settist sjálf út við gluggann allskammt
on onto her, and sit_down herself out by the:window all:close
frá Guðrúnu og horfði um hríð út.
from Guðrún and watched for.a while out
'And when little Sigrún was fast asleep, Sigríður laid her gently on the couch with a small pillow beneath her head, spread her apron over her, and then seating herself a short distance from Guðrún, looked out the window for a time.' Thoroddsen (1850)[51]

The passage above begins with a backgrounded adverbial clause resituating the following events on the main narrative timeline. The text continues:

(26) *Allt var kyrrt á strætum bæjarins; dálítill snjófölvi var yfir*
all was quiet on streets the.GEN:town a_little snow:paleness was over
jörðunni; veðrið var hreint og heiðríkt, og sólin var þegar
the:ground; the.weather was clear and fair and the:sun was already
sigin, og sló blóðrauðum geislum um allan vestursjóinn; jökulinn
set and cast blood.red rays about all the:west:sea; the:glacier
hillti upp; fiskibátarnir voru að koma að, sumir að lenda, en sumir
visible ; the:fishing:boats were to arrive some to land but some
voru komnir inn fyrir eyjarnar og skriðu fagurlega í
were arrived further in.front.of the:islands and glided beautifully in
logninu.
the:calm
'Everything was quiet on the town's streets; a little snow covered the earth; the weather was clear and fair, the sun had already set, casting blood-red rays across the entire western sea; the glacier stood visibly tall; the fishing boats were coming in, some approaching to anchor, and

[51]Translation due to Thoroddsen & Reeves (1890: 166).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

> others navigated the calm waters near the islands gliding beautifully in the calm winds' (Thoroddsen 1850)[52]

The above excerpt situates a number of atelically construed events within a global topic time: the sun's rays, the sight of the glacier, the arrival and anchoring of fishing boats. We again see the progressive used to describe such events: (the fishing boats) *voru að koma að, sumir voru að lenda, sumir voru komnir inn fyrir eyjarnar* 'were arriving, some anchoring, others had made it past the islands'.[53]

This section has served to illustrate the predominance of unbounded strategies found in the novel *Piltur og stúlka*, which I take to be representative of narrative structure more generally for Icelandic of this period. In the following section, I demonstrate that a decrease in bounded strategies can be observed in corpora for the time-period in question.

5 Decrease in bounded strategies

This section is structured as follows: in Section 5.1, I discuss the clear parallels between the shift in boundedness in the diachrony of English to the Icelandic case. Section 5.2 then draws on corpus evidence to substantiate similar claims for Icelandic.

5.1 Comparison to English

In this section, I draw on evidence regarding a shift in the frequency of temporal adverbials as a proxy for the shift in boundedness. This analytical move has precedent in the literature on Old English, a language which is not only closely related to Old Icelandic but also parallels it closely in this regard. In the following two subsections, I discuss the parallels with Old English in more detail, focusing first on the topic-time adverbials *þa* (OE) and *þá* (OI) in 5.1.1 and on the parallels between the rise of a dedicated progressive marker in both languages in 5.1.2.

5.1.1 Contrasting OE *þa* and OI *þá*

Just as Old Icelandic has the multi-functional adverbial item *þá*, so too did Old English have the item *þa*: both could be used as local discourse anchors, as sub-

[52]Translation due to Thoroddsen & Reeves (1890: 166)

[53]An anonymous reviewer points out that this sort of descriptive passage, outside of the main narrative, may represent one of the innovations of the emerging genre of *novel* (Fludernik 2003). The reviewer points out that in Dutch, a language with an emergent progressive construction, does not use it in passages such as this one.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

ordinating conjunctions or as resumptive items (Brinton 1996, Enkvist & Wårvik 1987).

In Old and Middle English, *þa* had a specific role in narrative episodes: the demarcation of incipient narrative episodes (Enkvist & Wårvik 1987, Wårvik 1995). Interestingly for our purposes, this functional domain was later overtaken by other expressions in the Middle English period – I contend that analyses of this development are highly relevant for understanding what happened in the transitional period that Early Modern Icelandic finds itself in.

As mentioned above in Section 4.2, a number of changes pertaining to the mapping between clause and information structure in historical Icelandic have been tied to the emergence of a dedicated subject position (Booth & Beck 2021). This is the case for historical English as well, as discussed by Fludernik (1995) who shows that the functional domain previously occupied by *þa* gives way to a number of competing expressions, among them *And whan NP+VP* constructions.

Wårvik (1995), too, investigates the role of *þa* and *when* in Middle English. As is the case in Old Icelandic, which has both *þá* and *þegar*, the two may co-occur. Wårvik shows that there is individual variation among others with regard to their preference for parataxis as opposed to temporal subordinate clauses: a decrease in the former corresponds to an increase in the latter.

Brinton (1996) looks specifically at the development of the discourse function of the inception of narrative episodes in Middle English in texts by Malory and Chaucer. She shows that, in the Middle English period, preposed *when*-clauses become the predominant means of doing so. Brinton (1996) also discusses the status of these clauses in terms of grounding: she shows that in earlier Chaucer as well as later Malory, events described in these clauses are more backgrounded. They express given information, which is off the main narrative timeline.

Finally, Enkvist & Wårvik (1987) discusses the status of *when*-clauses in Middle English. The authors conclude that their status as foregrounded or backgrounded depends on whether or not they can be interpreted as advancing the narrative timeline. This is the case when when is to be interpreted as 'after' rather than 'while' and the event it embeds is new, focused information that receives a perfective construal (Fleischman 1985: 876). In contrast, the embedded event can receive an imperfective construal which is backgrounded on account of containing given information, conveying temporal simultaneity.

In sum, Brinton (1996), Wårvik (1995) and Enkvist & Wårvik (1987) describe the linguistic system of Middle English as one with a tendency for increased backgrounding. Concretely, this is reflected in the gradual demise of *þa* as a foregrounding topic time adverbial. The functional domain occupied by *þa* in earlier

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

stages of the language – the marking of the inception of narrative episodes – becomes overtaken by backgrounded adverbial clauses with the subordinator *whan* by late Middle English. We will see later on, in Section 5.2, that Icelandic *þá* underwent a similar fate, and like shown in Wårvik (1995), its demise correlates with the rise of preposed clauses containing *þegar* 'when'.

5.1.2 The rise of the progressive

Petré (2016) discusses the rise of the dedicated progressive construction *be V-ing* in the history of English. His line of argumentation will broadly be adopted here and it can be summarised as follows: a change in frequency of associated environments ("co-texts") can be critical in the conventionalisation (semanticisation) of an emerging meaning. Petré argues that the progressive construction became semanticized in its focalised usage (where it is largely obligatory in PDE) as a result of a more general increase of backgrounded adverbial clauses. This increased frequency is, in turn, a consequence of a broader shift in grounding strategies. Past-marked, temporal subordinate contexts embedded under the subordinating conjunction *when* are of special importance to the conventionalisation of this focalised usage. These contexts allow a shift in meaning from the present-participle form merely expressing an ongoing state to a state which is ongoing at a particular focal point. Subordinate clauses are crucial here as they provide the time-frame at which this focalisation can occur – this then becomes part of the expressions conventional meaning that spreads to main clause contexts.

5.1.3 Summary: Decline of bounded strategies in the history of English

Petré (2010) and Los (2012) discuss the decline of bounded construals in the history of English. Both authors argue that Old English had a more transparent correspondence between syntactic transformations and information-structural functions. This argumentation is based upon a number of empirical observations:

1. Topic-time adverbials, used to re-establish the topic time with each discourse transition, lose prominence (see also Los et al. 2023). Clause-initial *þa* as a topic-time anchor decreases from 36.1% to 15.4% in early ME and 11.3% in late ME (Westergaard 2009: 93).

2. Environments which favour unbounded construals show a concomitant rise in frequency: dedicated progressive marking (Killie 2008, 2014), *ginnan*-class (*onginnan*, *beginnan*) with infinitival complements (Petré 2010: 468) begin as perfectivisers but acquire an inchoative usage late in the OE

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

period (Los 2000: 269-271), are subsequently subject to semantic bleaching and further conventionalisation into the Middle English period (Brinton 1988, Funke 1922, Mossé 1938).

The relation of the latter two environments to unboundedness has to do with their meaning: inchoative constructions focus on the onset of a new event, which has close pragmatic ties to the ongoingness of said event. Petré (2010) argues that *ginnan* expressions replaced the functional domain of *weorðan* 'become' in ME – the latter focusing instead on the state transition. This, Petre argues, reflects a greater flexibility in ME and afterwards with regard to framing events in an unbounded manner. Similarly, progressives directly focus event ongoingness and are thus fundamentally imperfective, they can be considered the "present counterpart" of inchoatives (Carroll et al. 2004).

5.2 Statistical corpus evidence

As outlined in the previous section, the trajectory for the diachrony of English, for which a shift from a bounded to unbounded system has been proposed, resembles Icelandic very strongly. Here, I propose that an independent case can be made for Icelandic on the basis of similar argumentation and empirical evidence. Concretely, I argue for the following on the basis of corpus evidence:

1. There is an overall trend towards a higher frequency of backgrounded temporal subordinate clauses, as discourse traditions shift.

2. There is a strong correlation between a decrease in foregrounded, iconic narration strategies (including verb-initial and þá-verb-subject structures) and those which rely more heavily on temporal backshifting and backgrounding.

3. The decrease in bounded strategies is further reflected in the tail-linking requirement: global establishment of topic-time no longer necessitates strict tail-linking. This can be seen in the decrease of adjunct-initial main clauses relative to other possible configurations.

4. Both *búinn* and *PROG* rise rapidly in frequency during the early modern period, going from marginal constructions to integral parts of the grammar. I argue that this rise in frequency should be tied to the increased preference for unbounded construals in the language. For both constructions, temporal subordinate contexts are of special importance, as argued in Section 1.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

This gives rise to two main predictions concerning both constructions: (a) they should show an association with unbounded adverbials and (b) their increased frequency should be correlated by-text.

In the following subsections, I show the empirical evidence that backs up the above claims. (1) is substantiated in 5.2.1, (2) is shown in 5.2.2, (3) is shown in 5.2.3 and the two claims in (4) are argued for with data in 5.2.4 and 5.2.5 respectively. Finally, 5.2.6 discusses an additional construction associated with the loss of boundedness, the decline of the present participle.

5.2.1 Increase in temporal subordinate clauses

Temporal subordinate clauses in IcePaHC were extracted via two separate queries. The first query targets prepositional phrases headed by *þegar* 'when' or *nær* 'when' which in turn immediately dominate an adverbial clause, itself immediately dominating a subordinate clause. The second query targeted temporal subordinate clauses introduced by *er* (an all-purpose relativiser), *þá er* (lit. 'at the time that'; typical way of forming a temporal subordinate clause in Old Icelandic), as well as *þá* or *nær* (both meaning 'when'). This is done by targeting what are coded as relative clauses in the corpus, which immediately dominate a subordinate clause, then embedding an adverbial phrase. The results were collated in order to yield a total relative frequency per 1000 clauses. The results were collated in order to yield a total relative frequency per 1000 clauses. These results are plotted in Figure 8. I looked at narrative texts specifically for the time-period from 1540 to 1900 and compared a statistical model including Year as a predictor to one without (an F-test using the `anova` function in `R`) which revealed that this increase is statistically significant ($p = 0.017$).

5.2.2 Tracing the availability of unbounded strategies: the temporal subordinator *þegar*

The temporal subordinator *þegar* becomes the predominant one in the language, taking over from *þá (er)* sometime between 1500-1700. Below in Figure 9, I show that the rise of the temporal subordinator *þegar* is negatively correlated with the decline of *þá* as a foregrounding topic-time adverbial, where it appears preverbally (cf. Wårvik (1995)'s finding for Middle English).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

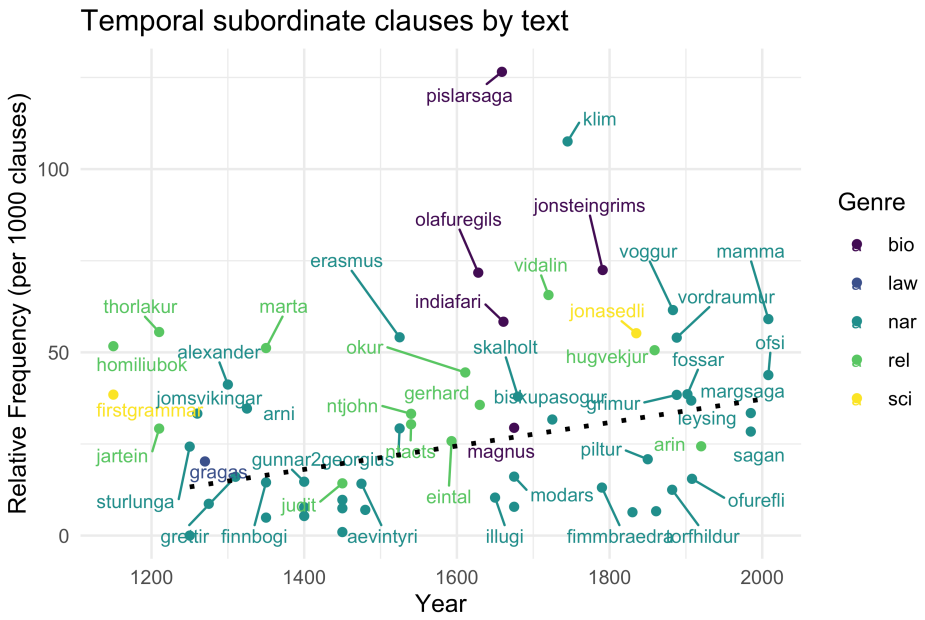


Figure 8: Rel.freq. of temporal subordinate clauses in IcePaHC

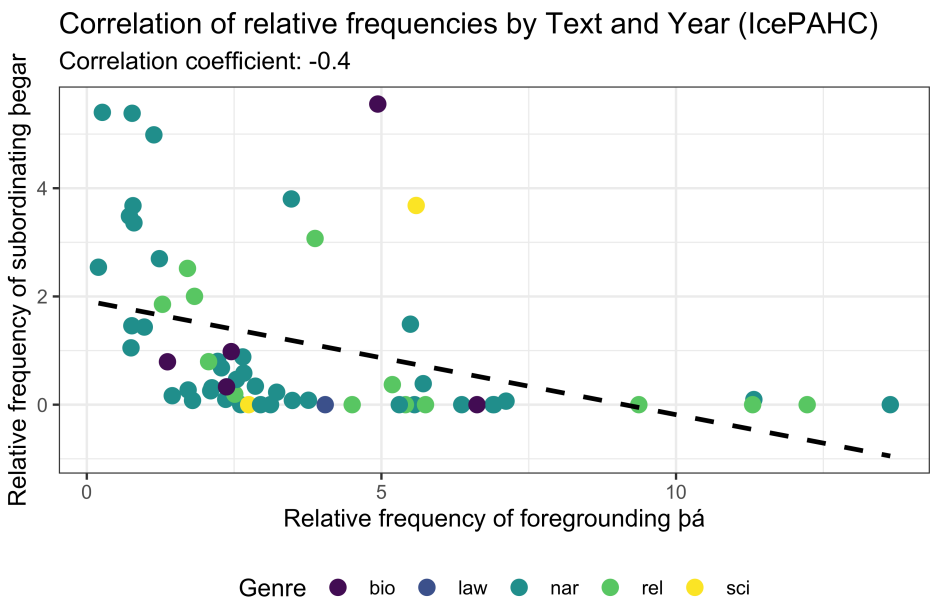


Figure 9: Correlation of foregrounding *þá* and subordinating *þegar* in IcePaHC. Pearson's coefficient: -0.4, $p < 0.002$.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

In the 16th to 19th centuries, written Icelandic prose style is under the influence of language contact via translation, e.g. due to both Latin and Low German in Oddur Gottskálksson's New Testament translation from 1540; this influence extends from the lexical to the syntactic (Helgason 1929). There are a couple areas where the language system is in flux during this period. Especially relevant for this paper is a shift in relativisation strategies (from *er* to *sem* with additional variation due to the interrogative strategy *hverr*) (Brendel 2023, Sapp 2019), which is intertwined with a change in subordinating strategies more generally.

Temporal subordinate clauses, which prior to the 16th century involved complementiser doubling, e.g. *þá er* 'then REL' and *þegar er* 'when REL', emerge both in new variants e.g. *þegar að* 'when COMPL' as well as without a complementiser or relativiser (Helgason 1929). In modern Icelandic written language, temporal subordinate clauses involve bare *þegar*. I take it that this change involves a tighter integration of the subordinate clause, as these changes are parallel to a decline in the sort of correlative constructions one finds in older texts to express the same sorts of temporal relations.

Kiparsky (1995) argues for the view that in Old Indo-European languages, what appear to be subordinate clauses should be analysed as instances of adjunction rather than true subordination. This explains, among other facts, why resumptive adverbials seem to show up in environments without there being, strictly speaking, an antecedent clause, such as clause initially. Such an environment is illustrated below in (27). On Kiparsky's analysis, this development is then effectively a transition from broadly paratactic to hypotactic systems.

(27) *Síðan varð hörð atlaga með þeim. Þa mællte Ólafr at hans menn*
then became hard attack with them. Then said Ólafur that his men
skyldu hlífa sér undir skjöldum, meðan þeir skyti spjótum ok
should protect themselves under shields, while they shot spears and
örum...
arrows
'Later they were attacked hard. Olaf said that his men should seek cover under the shields while the enemy shot spears and arrows at them...'
[*Ólafs saga helga* 24:44 cited in Meklenborg Nilsen 2020: 28]

It seems reasonable to me to tie these temporal resumptive uses to correlative relative clauses, such as the following cited in Wagener (2017: 229). What they both have in common is their exploitation of the left-periphery as an anaphoric position.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(28) *Þa ræðu er næst heyrða ec yður sægia um kaupmanna*
that speech REL next heard I you say about merchant:GEN.PL
iðrott þa var hon með glæggara froðleic fram flutt
profession then was she with more:intelligent wisdom performed
isvorum en í spurning
in_answers than in question
'The statement that I heard you make about the mercantile profession was uttered with more wisdom in the answer than in the question.'
[*Konungs skuggsjá* 38.18, cited in/translation due to Wagener]

Meklenborg Nilsen (2020: 29) discusses the difference between Older Indo-European languages and their modern counterparts in this respect: in the modern languages, only resumptives can occur between fronted adverbials and the finite verb, whereas in the older languages a number of other elements could appear in place of the resumptive. Clause-external adverbials in the older languages required that the resumptive be a full phrase in order to satisfy V2, but this is no longer the case in the modern languages, where resumptives are integrated and can fulfill V2 as they originate in the left-periphery. Meklenborg Nilsen (2020) takes this as evidence for the increased integration of the adverbial clause, which is accompanied by syntactic reanalysis of the resumptive.

Early Modern Icelandic thus represents an intermediate period where there is potential for ambiguity between hypotaxis and parataxis, embedding and adjunction. More conclusive evidence of this increased integration having played a role in Icelandic of the 16th to 19th centuries comes from a shift from double complementation strategies for temporal subordination, often involving extraposition (Sapp 2019, Wallenberg 2016), to the predominance of bare *þegar* for this purpose in the modern language.

My corpus investigation in IcePaHC shows that *þegar* wins out over *þá er/þá* by 1900, with the preceding period reflecting considerable genre variation. This is shown in Figure 10 below: we see that the overall proportion of *þegar* rarely exceeds 50% prior to 1600, while by 1850 alternative temporal subordination strategies have all but disappeared.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

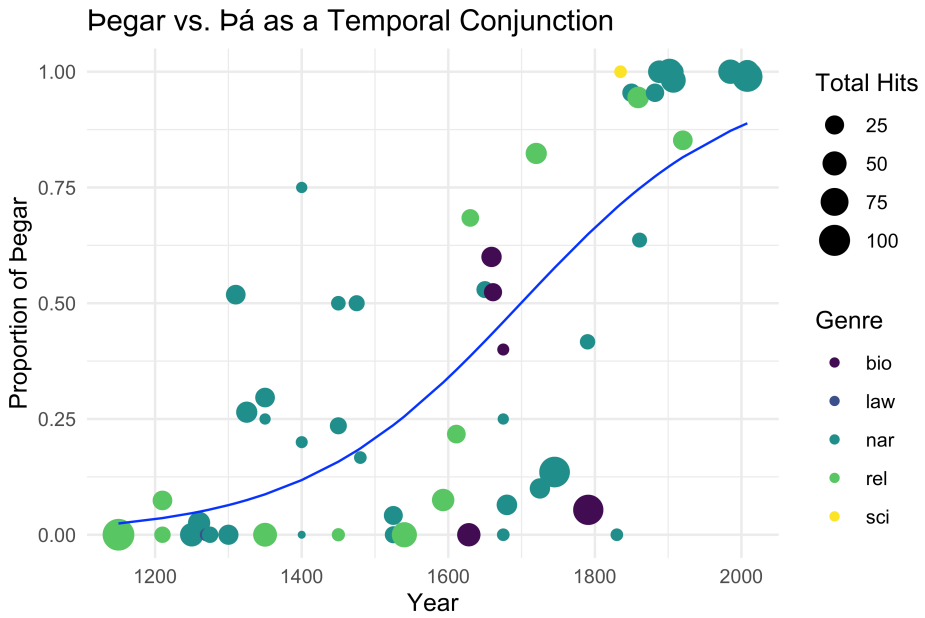


Figure 10: Subordinating *þegar* as opposed to subordinating *þá er/þá* in IcePaHC

The frequencies of the expressions relative to one another show that the conjunctionalisation is less advanced with *þá* 'then' than *síðan* 'since' and *þegar* 'when' (Bjerre 1935). This is tied to the fact that *þá* is a determinative temporal adverb in contrast with the others: "The demonstrative force in *þá* is then a hindrance for its development into a conjunction" (Bjerre 1935: 156). According to Bjerre, *er* has a tendency to disappear after all three, but this happens most clearly with *þegar* and least with *þá er*. The *er* in *þegar er* is perceived as pleonastic and thus gets ellided easily, whereas the relativiser function of *er* in *þá er* is supported by the determinative component of *þá* (Bjerre 1935). This shows that the conjunctionalisation there is less developed.

Why is *þá* poorly suited for unbounded construals, then? *Þá* is rare in temporal clauses following a determinative correlate. With few exceptions, *þá* er does not occur in temporal clauses after definite temporal expressions (Bjerre 1935). This is not solely do to the fact that *þá* in combination with a previous temporal expression would be pleonastic, since *þá er* does indeed occur in connection with temporal expressions in the main clause. However, this happens nearly exclusively in those cases where the connection between the temporal expression and the adverbial clause is relatively loose or the latter can be considered as some-

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

what freestanding (ibid). On the whole, it appears that *þegar* emerged as the best available candidate for an all-purpose subordinating conjunction.

5.2.3 Decrease in tail-linking

To investigate the role of discourse anchoring, I use adjunct-initial V2 clauses as a proxy – the vast majority of these involve temporal anchors (Booth & Beck 2021). Here I report on a regression analysis carried out to investigate the proportion of adjunct-initial clauses over time; I build on the discussion in Booth & Beck (2021: 35), but I make use of a different periodisation than discussed there. A binomial mixed-effects model was fitted in `R` (R Core Team 2023) using the `lme4` package (Bates et al. 2015) with a term for Year (centred and scaled) and Genre, as well as a random effect for Text. The model is specified as follows, including data from all genres for the time-period from 1540–1900:[54]

(29) `cbind(successes, failures) ~ scaled_Year + Genre + (1 | Text)`

The model specified above reveals that the overall decrease of non-subject initial V2 clauses (as a proportion of possible matrix clause configurations), while descriptively present, is not statistically significant ($\beta = -0.3640$, p = 0.1192). As observed by Booth & Beck (2021), biographies, religious and scientific texts pattern differently from narrative ones. However, for the given periodisation there is a decrease in all genres except for scientific texts. Of all genres, only narrative texts exhibit a statistically significant decrease for this period ($\beta = -0.4921$, $p = 0.00448$).

5.2.4 Frequency of associated adverbials

That the phenomena in question pattern with unbounded construals can be shown by examining the relative frequency of adverbials which pattern alongside them (cf. Petré 2010). Both phenomena collocate most commonly with no adverbial, which suggests that they pattern with unbounded construals. As far as overt adverbials go, *búinn* is predominantly associated with temporal subordinate clauses introduced either by *þegar* or *þá (er)*, both compatible with imperfective construals. Perhaps better evidence for the lack of boundedness is that the two constructions do not collocate with bounding adverbials (e.g. *after*, *at X time*, cf. Petré 2010: 474). *PROG* also collocates with *þegar* and *þá (er)*, however the most dominant adverbial is *meðan* 'while', clearly associated with ongoingness. In Tables

[54]This is a bit broader than the 1850 mark so as to not exclude otherwise valuable data.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

5, 6 and 7 below, *þá-s* designates subordinating *þá* (often, but not always accompanied by the relativiser *er*), while *þá-t* represents *þá* in its use as a topic-time anchor (clause-initial or clause-medial).

Table 5: Collocating adverbials for *búinn* IcePAHC until 1900 (n=90)

Adv	Percentage
none	69.10
þegar 'when'	11.16
þá-s 'when'	10.73
eftir 'after'	3.43
þá-t 'then'	2.15
fyrr 'before'	1.29
nær 'when'	0.86
er 'when'	0.43
sem 'as/when'	0.43
þangað til 'until'	0.43

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Table 6: Collocating adverbials for *búinn* in *Textasafn* 16th-18th cent. (n=215)

Adv	Percentage
none	36.67
þá-s 'when'	12.22
nú 'now'	11.11
þegar 'when'	11.11
eftir 'after'	7.78
er 'when'	3.33
nær 'when'	3.33
fyrr 'before'	2.22
svo 'then'	2.22
þá-t 'then'	2.22
allareiðu 'already'	1.11
eftir 'after'	1.11
fyrst 'first'	1.11
inn til þess 'until'	1.11
jafnskjótt og 'as soon as'	1.11
nú 'now'	1.11
oft 'often'	1.11

Table 7: Collocating adverbials for *PROG* in IcePAHC until 1920 (n=114)

Adv	Percentage
none	74.56
meðan 'while'	7.89
þegar 'when'	6.14
þá-s 'when'	3.51
þá-t 'then'	3.51
er 'when'	2.63
í þessu 'at this/that moment'	0.88
nú 'now'	0.88

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

5.2.5 Timing of changes

As displayed in Figure 11, the correlation between the relative frequency of *búinn* and *PROG* over time is moderate (0.61) and statistically significant. This supports the argument that the two changes were bolstered by the same underlying factor, the shift in boundedness. On the whole, this perspective provides insight into the time-course of propagation for both innovations, which is otherwise puzzling: why did the constructions stay relatively marginal before seeing a jump in frequency from the early 19th century onwards? My proposal is that the establishment of a predominantly unbounded system had to occur first. The chain of causality I would like to propose is as follows. The expressions in question could become available via chance reanalysis.[55] Their subsequent conventionalisation, leading to the well-developed aspectual system exhibited by modern Icelandic, was favoured by other ongoing structural developments in the language, i.e. tense indeterminacy, inter-clausal relations in flux. Regarding the latter, alongside the increased integration of adverbial subclauses, speakers had expressions available for explicitly marking event phase transitions.

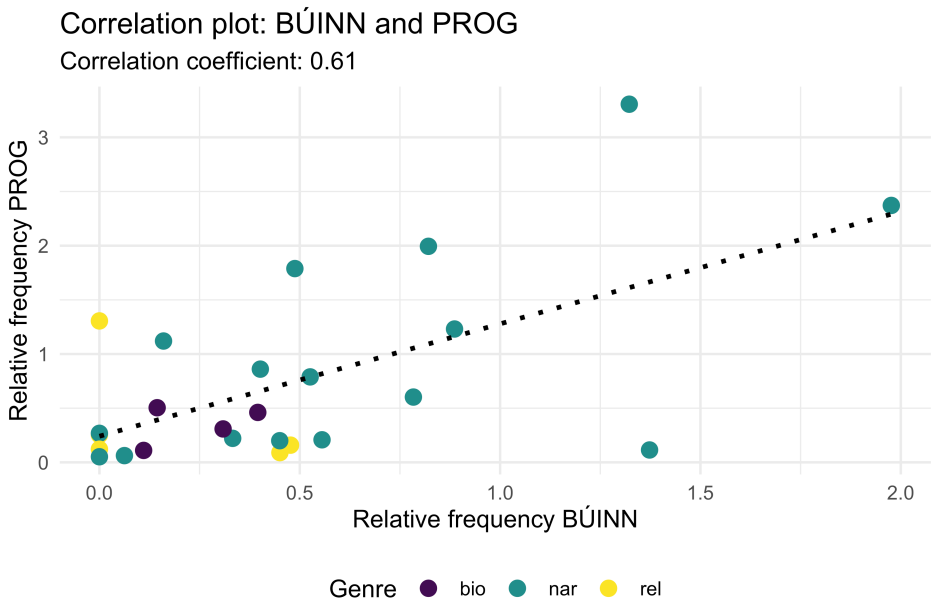


Figure 11: Correlation of rel.freq. of *búinn* and *PROG* by text/year. Pearson's coefficient: 0.586, $p < 0.003$

[55] I thank an anonymous reviewer for requesting I expand on this point.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

5.2.6 Additional construction associated with the loss of boundedness: present participle

Constructions involving the present participle often used to designate time in the place of subordinate clauses (Hauksson & Óskarsson 1994). This is accompanied by a decline in dative absolutive constructions. Both are considered features of the *learned style* of medieval Icelandic, but are less prominent in modern Icelandic narrative style (ibid.). The example below in (30) exemplifies both of these: *að morgni* 'at morning.DAT' is an absolutive, while *komandi* 'coming' is a present participle. The example is from *Indíafari*, which, as previously mentioned, exhibits features of bounded as well as unbounded narrative. The phrase as a whole *að morgni komandi* 'at (the) morning coming' (in other words, when morning arrives) is used in place of a subordinate clause indicating temporal sequentiality.

(30) *Þessi lofaði að morgni komandi þar upp á honum*
DEM.M promised at morning.DAT come.PTCP.PRS there up to him
andsvar að gefa.
answer to give
'The steward promised to give him an answer the following morning.'
(Ólafsson 1908: 262)[56]

The loss of prominence of the present participle in this function occurs alongside the increase of *búinn* especially. I would like to propose that this is due to the novel expression overtaking this functional domain across discourse traditions. Texts with a relative frequency of 2% or above are not represented in the corpus after 1900. This is shown below in Figure 12. The texts with highest relative frequency of the construction are predominantly translations with considerable Latinate influence, e.g. *Marta* and *Ectorssaga*.[57]

[56]Translation sourced from Ólafsson (2017: 83).

[57]See https://github.com/antonkarl/icecorpus/blob/master/info for more information on the source texts in the corpus.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

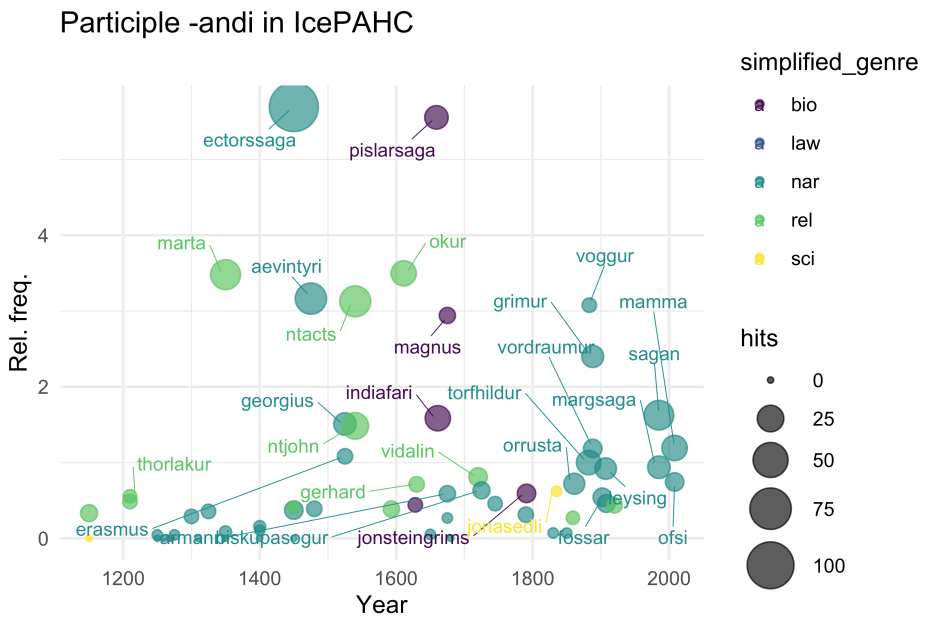


Figure 12: Present participle (*-andi*) in IcePaHC over the centuries

6 Conclusion

The seemingly abrupt rise in frequency of both *búinn* and *PROG*, starting in the 18th century, is difficult to explain without recourse to a shift in mapping between clause and discourse structure. In this article, I argue that this shift paved the way for these unbounded aspectual constructions to gain sufficient frequency in order to be propelled to conventionalisation.

To summarise, I have argued for the following:

1. Icelandic exhibits a shift in preference for predominantly bounded, to predominantly unbounded strategies for structuring narrative, during the early modern period (1540–1850).

2. This shift in boundedness is reflected in both a higher frequency of backgrounded temporal clauses as well as a concomitant decrease in foregrounding strategies, most notably clauses starting with the topic-time adverbial *þá* followed by the finite verb.

3. Two periphrastic aspectual constructions, *búinn* and *PROG* went from being relatively marginal to integral parts of the aspectual oppositions made

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

by Icelandic grammar. I show that an increased availability in unbounded construals favoured this change, as backgrounded temporal clauses are important for the conventionalisation of their meaning. This can moreover be seen in the time-course of the changes in question and the fact that they are highly correlated.

Alongside the aspectual constructions which are the focus of this paper, a number of other changes occurred in the history of Icelandic pertaining to the mapping between syntax/semantics and discourse structure, which I take to be compatible with this perspective (Booth 2021). Most prominently discussed among these changes in the literature on Icelandic diachrony is the rise of a dedicated subject position, which has furthermore been tied to the shift from optionally SOV to obligatorily SVO (Rögnvaldsson 1996, Hróarsdóttir 2000) and the decline of verb-first declaratives (Booth & Beck 2021). I leave it to future research to spell out the relationship between these changes in more detail.

Before concluding, a few words are in order about the implications of this paper for diachronic typology. The notion of boundedness is conceived of as a typological one in the literature (von Stutterheim & Lambert 2005, von Stutterheim & Nüse 2003) and there is as of yet little research which explicitly discusses the diachronic component, i.e. the switching of systems from predominantly bounded to unbounded or vice-versa.[58] My discussion in the paper can be related to the work of Bhat (1999), who proposes that languages can be classified as tense, aspect or mood-prominent. In my understanding of the literature on boundedness (Carroll & Lambert 2003, Carroll et al. 2004, von Stutterheim & Nüse 2003, von Stutterheim & Lambert 2005, von Stutterheim et al. 2002), bounded structures correspond, broadly speaking, to tense prominence and unbounded to aspect prominence. Bhat (1999: 179) discusses his typological findings in the context of Hopper (1979) on grounding, writing that "languages appear to select a verbal form that belongs to their most prominent verbal category for encoding the foregrounded material". In a tense prominent language like Kannada, this would be the simple (narrative) past, whereas in aspect prominent languages, like Russian, it is the perfective form. A shift in prominence would thus be expected to correspond to a change in preferred strategies for narrative grounding. This is also why I take it that boundedness and prominence are tightly related.

A number of diachronic studies relate their findings to Bhat's (1999) notion of prominence. The diachrony of Hebrew, for instance, is widely considered to have undergone a shift from aspect prominence to tense prominence between the Biblical and Tannaitic periods (Penner 2015: 74–81). An entirely unrelated language,

[58] I thank an anonymous reviewer for bringing up this point.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Basque, has been claimed to have undergone a shift from tense-prominence to aspect prominence, including the emergence of a dedicated progressive form (Martínez-Areta 2022). A further point of comparison is to the notions of satellite- and verb-framed languages (Talmy 1985); see Fanego (2024) for a recent discussion that ties a boundedness shift in historical English to that literature.

On the whole, I hope that this paper can provide a lasting contribution to this emerging empirical literature regarding the relationship between linguistic structure and the encoding of discourse structure.

Abbreviations

ACC	Accusative
ADV	Adverbial
ART	Article
BIO	Biographical texts (genre)
COMP	Complementiser
DAT	Dative
DEM	Demonstrative
DT	Discourse tradition
F	Feminine
FOCAL	Focalised (progressive reading)
GEN	Genitive
IcePaHC	Icelandic Parsed Historical Corpus
INDET	Indeterminate (reading type)
IP-MAT	Matrix clause (IcePaHC tagset)
IP-SUB	Subordinate clause (IcePaHC tagset)
LAW	Legal texts (genre)
M	Masculine
N	Neuter
NAR	Narrative (genre)
NOM	Nominative
OE	Old English
OI	Old Icelandic
PDE	Present-day English
PROG	Progressive (aspect)
PRS	Present
PST	Past
PTCP	Participle
REL	Religious texts (genre) OR Relativiser (glossing)
SCI	Scientific texts (genre)
SuB	Subordinate clause expressing temporal sequentiality

Acknowledgements

Thanks to Ulrike Demske and Barthe Bloom for their feedback and editorial work, without which this paper would not exist. For their thought-provoking comments on the talk that led to this paper, I am grateful for conversations with workshop participants: Malika Reetz, Sonja Zeman, Anna Cichosz and Bettelou

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Los. This paper was improved greatly thanks to comments from two anonymous reviewers; their contribution is hereby acknowledged. Finally, thanks are due to Uli Sauerland, Stephanie Solt and Artemis Alexiadou for their academic and personal support throughout the final phase of my PhD project, a part of which is reflected in this paper.

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – SFB 1412, 416591334.

References

Aldai, Gontzal. 2002. *The grammaticalization of present and past in Basque*. University of Southern California. (Doctoral dissertation).

Amory, Frederic. 1980. Narrative syntax in the typical saga scene. *The Journal of English and Germanic Philology* 79(3). 390–407. https://www.jstor.org/stable/27708685.

Andersson, Theodore M. 1967. *The Icelandic family saga: An analytic reading*. Harvard University Press. DOI: 10.4159/harvard.9780674729292.

Árnason, Mörður. 1977. Búinn er nú að stríða: athugasemdir um orðasambandið *vera búinn að* + nafnháttur (Has now fought. Remarks on the construction *vera búinn að* + Infinitive). Unpublished BA thesis. Reykjavík.

Asher, Nicholas & Laure Vieu. 2005. Subordinating and coordinating discourse relations. *Lingua* 115(4). 591–610. DOI: 10.1016/j.lingua.2003.09.017.

Barðdal, Jóhanna & Thórhallur Eythórsson. 2003. The change that never happened: The story of oblique subjects. *Journal of Linguistics* 39(3). 439–472. DOI: 10.1017/S002222670300207X.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1). 1–48. DOI: 10.18637/jss.v067.i01.

Bech, Kristin. 2014. Tracing the loss of boundedness in the history of English: The anaphoric status of initial prepositional phrases in Old English and Late Middle English. *Anglia* 132(3). 506–535. DOI: 10.1515/ang-2014-0056.

Benediktsson, Hreinn. 2002. Icelandic vera að + Infinitive: Age and Origin. In Guðrún Þórhallsdóttir, Höskuldur Þráinsson, Jón G. Friðjónsson & Kjartan Ottósson (eds.), *Hreinn benediktsson - linguistic studies, historical and comparative*, 449–469. Originally appeared in Gösta Holm, Nordiska studier i filologi och lingvistik, Lund, 1976 as Ísl. vera að + nafnh.: Aldur og uppruni. Reykjavík: Málvísindastofnun Háskóla Íslands (University of Iceland, Institute of Linguistics).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Berman, Ruth A. & Dan Isaac Slobin (eds.). 1994. *Relating events in narrative: A crosslinguistic developmental study*. New York: Psychology Press. DOI: 10.4324/9780203773512.

Bernharðsson, Haraldur. 2017. Jón Thoroddsen og málstöðlun nítjándu aldar (Jón Thoroddsen and language standardisation in the 19th century). *Orð og tunga* 19. DOI: 10.33112/ordogtunga.19.4.

Bernharðsson, Haraldur. 2018. Spreading the standard: The nineteenth-century standardization of Icelandic and the first Icelandic novel. *Journal of Historical Sociolinguistics* 4(2). 149–176. DOI: 10.1515/jhsl-2017-0014.

Bertinetto, Pier Marco. 2000. The progressive in Romance, as compared with English. In *Empirical approaches to Language Typology*, 559–604. Östen Dahl. DOI: 10.1515/9783110197099.4.559.

Bertinetto, Pier Marco, Karen H Ebert & Casper De Groot. 2000. The progressive in Europe. *Empirical Approaches to Language Typology* (6). 517–558. DOI: 10.1515/9783110197099.4.517.

Bhat, Darbhe N. Shankara. 1999. *The prominence of tense, aspect and mood* (Studies in Language Companion Series 49). Amsterdam: Benjamins. DOI: 10.1075/slcs.49.

Biber, Douglas & Susan Conrad. 2005. Register variation: A corpus approach. In *The Handbook of Discourse Analysis*, 175–196. John Wiley & Sons, Ltd. DOI: 10.1002/9780470753460.ch10.

Bjerre, Birger. 1935. *Nordiska konjunktionsbildningar med temporal innebörd i*, vol. 31:3 (Lunds universitets årsskrift: N. F. Avd. 1). Lund: C. W. K. Gleerup.

Booth, Hannah. 2018. *Expletives and Clause Structure: Syntactic Change in Icelandic*. Accessed: 06.01.2023. University of Manchester. (Doctoral dissertation).

Booth, Hannah. 2021. Revisiting the configurationality issue in Old Icelandic. *Glossa: A journal of general linguistics* 6(1). 130. DOI: 10.16995/glossa.5804.

Booth, Hannah & Christin Beck. 2021. Verb-second and verb-first in the history of Icelandic. *Journal of Historical Syntax* 5(28). 1–53. DOI: 10.18148/hs/2021.v5i28.112.

Brendel, Christian D. 2023. Translation-induced interrogative relativizers and stability in Icelandic. *Transactions of the Philological Society* 121(1). 1–32. DOI: 10.1111/1467-968x.12258.

Brinton, Laurel J. 1988. *The development of English aspectual systems: Aspectualizers and post-verbal particles*. Cambridge: Cambridge University Press.

Brinton, Laurel J. 1996. *Pragmatic markers in English: Grammaticalization and discourse functions*, vol. 19 (Topics in English Linguistics; 19). Berlin: De Gruyter. DOI: 10.1515/9783110907582.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

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

Carroll, Mary & Monique Lambert. 2003. Information structure in narratives and the role of grammaticised knowledge: A study of adult French and German learners of English. In Christine Dimroth & Marianne Starren (eds.), *Information structure and the dynamics of language acquisition*, 267–287. Amsterdam: Benjamins. DOI: 10.1075/sibil.26.13car.

Carroll, Mary, Christiane von Stutterheim & Ralf Nuese. 2004. The language and thought debate: A psycholinguistic approach. In Thomas Pechmann & Christopher Habel (eds.), *Multidisciplinary approaches to language production*, vol. 157 (Trends in Linguistics: Studies and Monographs), 183–218. Berlin: De Gruyter. DOI: 10.1515/9783110894028.183.

Chafe, Wallace. 1984. How people use adverbial clauses. In *Annual Meeting of the Berkeley Linguistics Society*, vol. 10, 437–449. DOI: 10.3765/bls.v10i0.1936.

Clover, Carol J. 1974. Scene in saga composition. *Arkiv för nordisk filologi* 89. 57–83. DOI: 10.7591/9781501740510.

Comrie, Bernard. 1976. *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge University Press.

Daller, Michael H., Jeanine Treffers-Daller & Reyhan Furman. 2011. Transfer of conceptualization patterns in bilinguals: The construal of motion events in Turkish and German. *Bilingualism: Language and Cognition* 14(1). 95–119. DOI: 10.1017/s1366728910000106.

Diewald, Gabriele. 2002. A model for relevant types of contexts in grammaticalization. In Ilse Wischer & Gabriele Diewald (eds.), *New reflections on grammaticalization*, vol. 49 (Typological Studies in Language), 103–120. Amsterdam: Benjamins. DOI: 10.1075/tsl.49.09die.

van Dijk, Teun A. 1976. Philosophy of action and theory of narrative. *Poetics* 5(4). 287–338. DOI: 10.1016/0304-422x(76)90014-0.

Donhauser, Karin, Lars Erik Zeige & Michael Solf. 2006. Informationsstruktur und Diskursrelationen im Vergleich. In Antje Hornscheidt, Kristina Kotcheva, Tomas Milosch & Michael Rießler (eds.), *Grenzgänger: Festschrift zum 65. Geburtstag von Jurij Kusmenko*, 1st edn., 73–90. Berlin: Nordeuropa-Institut. DOI: 10.18452/1849.

Eckardt, Regine. 2006. *Meaning change in grammaticalization: An enquiry into semantic reanalysis*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199262601.001.0001.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Eckardt, Regine. 2009. APO: Avoid pragmatic overload. In Maj-Britt Mosegaard Hansen & Jacqueline Visconti (eds.), *Current trends in diachronic semantics and pragmatics*, vol. 7 (Studies in Pragmatics), 21–41. Emerald Group Publishing Limited. DOI: 10.1163/9789004253216_003.

Enders, Irina. 2001. Verbets placering i fornisländska sagor [placement of the verb in Old Icelandic sagas]. In Jurij Kusmenko & Sven Lange (eds.), *Nordiska språk – insikter och utsikter*, 90. Föredrag från det 5:e studentkollokviet i nordisk språkvetenskap, Berlin, 6.–7. Mai 2000. Berlin: Nordeuropa-Institut. https://www.ni.hu-berlin.de/de/publ/publikationsreihen/kleine_schriften/pdf/insikterutsikter.pdf.

Enkvist, Nils Erik & Brita Wårvik. 1987. Old english *tha*, temporal chains and narrative structure. In Anna Giacalone Ramat, Onofrio Carruba & Giuliano Bernin (eds.), *Papers from the 7th International Conference on Historical Linguistics*, 221–237. Amsterdam: Benjamins. DOI: 10.1075/cilt.48.17enk.

Faarlund, Jan Terje. 1990. *Syntactic change: Toward a theory of historical syntax*, vol. 50 (Trends in Linguistics. Studies and Monographs [TiLSM]). Berlin: De Gruyter. DOI: 10.1515/9783110854947.

Fanego, Teresa. 2024. English motion and progressive constructions, and the typological drift from bounded to unbounded discourse construal. *Language Sciences* 101(598). DOI: 10.1016/j.langsci.2023.101598.

Flecken, Monique, Christiane von Stutterheim & Mary Carroll. 2013. Principles of information organization in L2 use: Complex patterns of conceptual transfer. *International Review of Applied Linguistics* 51(2). 229–242. DOI: 10.1515/iral-2013-0010.

Fleischman, Suzanne. 1985. Discourse functions of tense-aspect oppositions in narrative: Toward a theory of grounding. *Linguistics* 23(6). 851–882. DOI: 10.1515/ling.1985.23.6.851.

Fleischman, Suzanne. 1990. *Tense and narrativity: From medieval performance to modern fiction*. London: Routledge. DOI: 10.7560/780903.

Fludernik, Monika. 1995. Middle English *þo* and other Narrative Discourse Markers. In Andreas H. Jucker (ed.), *Historical pragmatics: Pragmatic developments in the history of English*, 359–392. Amsterdam: Benjamins. DOI: 10.1075/pbns.35.21flu.

Fludernik, Monika. 2003. The diachronization of narratology. *Narrative* 11(3). 331–348. DOI: 10.1353/nar.2003.0014.

Fludernik, Monika. 2012. Narratology and literary linguistics. In Robert I. Binnick (ed.), *The Oxford handbook of tense and aspect*, 75–101. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780195381979.013.0002.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Friðriksson, Finnur. 2011. Modern Icelandic: Stable or in a state of flux. *Scripta Islandica* 62. 19–50.

Funke, Otto. 1922. Die Fügung *ginnen* mit dem Infinitiv im Mittelenglischen. *Englische Studien* 56. 1–27.

Givón, Talmy. 1979. *On understanding grammar*. New York: Academic Press. DOI: 10.1016/C2013-0-10728-3.

Haeberli, Eric. 2002. Inflectional morphology and the loss of verb-second in English. *Syntactic Effects of Morphological Change*. 88–106. DOI: 10.1093/acprof:oso/9780199250691.003.0005.

Haiman, John. 1985. *Natural syntax: Iconicity and erosion*. Cambridge: Cambridge University Press.

Hauksson, Þorleifur & Þórir Óskarsson. 1994. *Íslensk stílfræði [Icelandic stylistics]*. Reykjavík: Styrktarsjóður Þórbergs Þórðarsonar og Margrétar Jónsdóttur, Háskóla Íslands, & Mál og menning.

Heine, Bernd. 2002. On the role of context in grammaticalization. In Ilse Wischer & Gabriele Diewald (eds.), *New reflections on grammaticalization*, vol. 49 (Typological Studies in Language), 83–101. Amsterdam: Benjamins. DOI: 10.1075/tsl.49.08hei.

Heine, Bernd, Ulrike Claudi & Friederike Hünnemeyer. 1991. *Grammaticalization: A conceptual framework*. Ilse Wischer & Gabriele Diewald (eds.). Chicago: University of Chicago Press. DOI: 10.1017/s0022226700000487.

Helgason, Jón. 1929. *Málið á Nýja testamenti Odds Gottskálkssonar [The language of Oddur Gottskálksson's New Testament]* (Safn Fræðafélagsins um Ísland og Íslendinga). Copenagen: Hið íslenzka fræðafélag.

Heusler, Andreas. 1921. *Altisländisches Elementarbuch* (Germanische Bibliothek / 1: Elementar- und Handbücher). Heidelberg: C. Winter.

Holberg, Louis. 1845. *Niels Klim's journey under the ground: Being a narrative of his wonderful descent to the subterranean lands; together with an account of the sensible animals and trees inhabiting the planet Nazar and the firmament*. Trans. by John Gierlow. Boston, MA: Saxton, Peirce & Co. https://www.gutenberg.org/files/27884/27884-h/27884-h.htm.

Holberg, Ludvig. 1948. *Nikulás Klím*. Jón Helgason (ed.). Trans. by Jón Ólafsson úr Grunnavík, vol. 3 (Íslensk rit síðari alda). Printed by S. L. Möller. Copenhagen: Hið íslenska fræðafélag í Kaupmannahöfn. DOI: 10.2307/jj.7193899.

Holm, Gösta. 1958. *Syntaxgeografiska studier över två nordiska verb [Syntactic geographical studies of two Scandinavian verbs]*, vol. 4 (Skrifter utgivna av Institutionen för nordiska språk vid Uppsala universitet). Uppsala & Wiesbaden: A.-B. Lundequistska Bokhandeln & Otto Harrassowitz.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Holmberg, Anders. 2000. Scandinavian stylistic fronting: How any category can become an expletive. *Linguistic Inquiry* 31(3). 445–483. DOI: 10.1162/002438900554406.

Holvoet, Axel. 2014. Phasal and proximative complementation: Lithuanian *baigti*. *Baltic Linguistics* 5. 81–122. DOI: 10.32798/bl.404.

Hopper, Paul J. 1979. Some observations on the typology of focus and aspect in narrative language. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"* 3(1). 37–64. DOI: 10.1075/sl.3.1.03hop.

Hróarsdóttir, Þorbjörg. 2000. *Word order change in Icelandic: From OV to VO*, vol. 35 (Linguistik Aktuell/Linguistics Today). Amsterdam: Benjamins. DOI: 10.1075/la.35.

Jóhannsdóttir, Kristín M. 2011. *Aspects of the progressive in English and Icelandic*. University of British Columbia. (Doctoral dissertation). DOI: 10.14288/1.0072196.

Jónsson, Finnur (ed.). 1931. *Edda Snorra Sturlusonar*. Copenhagen: Gyldendal.

Jónsson, Finnur. 1933. *Den islandske grammatiks historie til o. 1800 [The history of Icelandic grammar until 1800]*. Copenhagen: Levin & Munksgaard.

Jónsson, Jóhannes Gísli. 1992. The two perfects of Icelandic. *Íslenskt mál* 14. 129–145.

Kabatek, Johannes. 2022. Discourse traditions and the historicity of language: Discourse traditional knowledge and discourse universes. *Manual of Discourse Traditions in Romance* 30. 103–122. DOI: 10.1515/9783110668636-005.

van Kemenade, Ans & Bettelou Los. 2006. Discourse adverbs and clausal syntax in Old and Middle English. In Ans van Kemenade & Bettelou Los (eds.), *The handbook of the history of English*, 224–248. Oxford: Wiley Blackwell. DOI: 10.1002/9780470757048.ch10.

Kern, Jane Eleanor. 2019. *Oh, great mystery and wondrous sign: The miracles of St Þorlákr and the public experience of the divine in thirteenth and mid-fourteenth century Iceland*. Reykjavik, Iceland: University of Iceland. (MA thesis). https://skemman.is/bitstream/1946/34332/1/Jane%20Eleanor%20Kern%2c%20MIS%20MA%20Thesis%202019.pdf.

Killie, Kristin. 2008. From locative to durative to focalized? The English progressive and the 'PROG imperfective drift'. In Maurizio Gotti, Marina Dossena & Richard Dury (eds.), *English historical linguistics 2006, Volume 1: Syntax and morphology* (Selected Papers from the Fourteenth International Conference on English Historical Linguistics (ICEHL 14), Bergamo, 21-25 August 2006), 69–88. Amsterdam: Benjamins. DOI: 10.1075/cilt.295.07kil.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Killie, Kristin. 2014. The development of the English BE + V-*ende*/V-*ing* periphrasis: From emphatic to progressive marker? *English Language & Linguistics* 18(3). 361–386. DOI: 10.1017/s1360674314000148.

Kinn, Kari, Kristian A. Rusten & George Walkden. 2016. Null subjects in Early Icelandic. *Journal of Germanic Linguistics* 28(1). 31–78. DOI: 10.1017/S1470542715000136.

Kiparsky, Paul. 1995. Indo-European Origins of Germanic Syntax. In Adrian Battye & Ian Roberts (eds.), *Clause structure and language change*, 140–170. New York: Oxford University Press. DOI: 10.1093/oso/9780195086324.003.0006.

Klein, Wolfgang. 1994. *Time in language*. London: Routledge. DOI: 10.4324/9781315003801.

Koch, Peter. 1987. Distanz im Dictamen. Zur Schriftlichkeit und Pragmatik mittelalterlicher Brief- und Redemodelle in Italien. Unpublished habilitation thesis, Freiburg im Breisgau.

Koch, Peter & Wulf Oesterreicher. 1985. Sprache der Nähe—Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. *Romanistisches Jahrbuch* 36(1985). 15–43. DOI: 10.15496/publikation-20410.

Kossuth, Karen C. 1980. The linguistic basis of saga structure: Toward a syntax of narrative. *Arkiv för nordisk filologi* 95. 126–141.

Kress, Bruno. 1982. *Isländische grammatik*. 1st edn. Leipzig: Verlag Enzyklopädie.

Krifka, Manfred. 2008. Basic notions of information structure. *Acta Linguistica Hungarica* 55(3-4). 243–276. DOI: 10.1556/ALing.55.2008.3-4.2.

Kristjánsson, Jónas. 1985. The Court Style. In Régis Boyer (ed.), *Les Sagas de Chevaliers: Riddarasögur*, 431–440. Paris: Presses universitaires de France.

Kupersmitt, Judy R. 2015. Moving on and off the timeline in narratives of personal experience written in three languages: A form-function developmental study. *Linguistics* 53(6). 1303–1339. DOI: 10.1515/ling-2015-0033.

Labov, William. 1972. *Sociolinguistic patterns* (Conduct and Communication 4). University of Pennsylvania Press.

Labov, William & Joshua Waletzky. 1967. Narrative Analysis. In June Helm (ed.), *Essays on the verbal and visual arts*, 12–44. Seattle: University of Washington Press. DOI: 10.1075/jnlh.7.1-4.02nar.

Landman, Fred. 1992. The Progressive. *Natural Language Semantics* 1. 1–32. DOI: 10.1007/bf02342615.

Lee, EunHee. 2017. Pluperfect in discourse: When and why do we go back in time? *Journal of Pragmatics* 121. 76–90. DOI: 10.1016/j.pragma.2017.10.001.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Leiss, Elisabeth. 2000. *Artikel und Aspekt: Die grammatischen Muster von Definitheit*, vol. 55 (Studia Linguistica Germanica). Berlin: De Gruyter. DOI: 10.1515/9783110825961.

Levelt, Willem JM. 1993. *Speaking: From intention to articulation*. Cambridge, MA: The MIT press. DOI: 10.7551/mitpress/6393.001.0001.

Liberman, Anatoly. 1990. The Present Perfect in Old and Modern Icelandic. In Wolfgang Raible (ed.), *Erscheinungsformen kultureller Prozesse*, 263–280. Tübingen: Narr.

Los, Bettelou. 2000. *Onginnan/beginnan* with bare and *to*-infinitive in Ælfric. In Olga Fischer, Anette Rosenbach & Dieter Stein (eds.), *Pathways of change: Grammaticalization in English*, 251–274. Amsterdam: Benjamins.

Los, Bettelou. 2012. The loss of Verb Second and the switch from bounded to unbounded systems. In Anneli Meurman-Solin, María José López-Couso & Bettelou Los (eds.), *Information structure and syntactic change in the history of English*, 21–46. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199860210.003.0002.

Los, Bettelou, Gea Dreschler, Ans. van Kemenade, Erwin Komen & Stefano Coretta. 2023. The decline of local anchoring: A quantitative investigation. *English Language & Linguistics* 27(2). 345–372. DOI: 10.1017/s1360674323000047.

Lüdeling, Anke, Artemis Alexiadou, Aria Adli, Karin Donhauser, Malte Dreyer, Markus Egg, Anna Helene Feulner, Natalia Gagarina, Wolfgang Hock, Stefanie Jannedy, Frank Kammerzell, Pia Knoeferle, Thomas Krause, Manfred Krifka, Silvia Kutscher, Beate Lütke, Thomas McFadden, Roland Meyer, Christine Mooshammer, Stefan Müller, Katja Maquate, Muriel Norde, Uli Sauerland, Stephanie Solt, Luka Szucsich, Elisabeth Verhoeven, Richard Waltereit, Anne Wolfsgruber & Lars Erik Zeige. 2022. Register: Language users' knowledge of situational-functional variation. *Register Aspects of Language in Situation (REALIS)*. 1–58. DOI: 10.18452/24901.

Martínez-Areta, Mikel. 2022. Latin and Romance Influence on the Basque Verbal Morphosyntax. *Journal of Language Contact* 14(3). 497–523. DOI: 10.1163/19552629-14030001.

McCawley, James D. 1971. Tense and Time Reference in English. In Charles J. Fillmore & D. Terence Langendoen (eds.), *Studies in linguistic semantics*, 96–113. New York: Irvington.

Meklenborg Nilsen, Christine. 2020. Adverbial resumptive particles and verb second. In Rebecca Woods & Sam Wolfe (eds.), *Rethinking verb second*, 90–125. Oxford: Oxford University Press. DOI: 10.1093/oso/9780198844303.003.0005.

Mossé, Fernand. 1938. *Histoire de la forme périphrastique Être + participe présent en Germanique*, vol. I. Paris: Klincksieck.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Nielsen, Marit Aamodt. 2017. A remark on the Old Norse conjunction *ok*. *Scripta Neophilologica Posnaniensia* (17). 235–242. DOI: 10.14746/snp.2017.17.15.

Nordling, Arnold. 1928. Isländskt hann var at telja = hann talði ~ latinets legēbam o.s.v. In Johannes Brøndum-Nielsen (ed.), *Festskrift til Finnur Jónsson 29. maj 1928*, 399–406. Copenhagen: Levin & Munksgaards Forlag.

Nygaard, Marius. 1905. *Norrøn syntax*. Christiania: H. Aschehoug & Company (W. Nygaard).

OEB Team. 2022. *Open English Bible*. 2022.1 release. https://openenglishbible.org/.

Ólafsson, Jón. 1908. Æfisaga Jóns Ólafssonar Indíafara. In Sigfús Blöndal (ed.), *Hið íslenska bókmenntafélag*, vol. 1. Copenhagen: HINU ISLENSKA BÓKMENTAFJELAGI.

Ólafsson, Jón. 2017. *The life of the Icelander Jón Ólafsson, Traveller to India, Written by Himself and Completed about 1661 AD: With a Continuation, by Another Hand, up to His Death in 1679. Volume I*. Trans. by Bertha S. Phillpotts. Taylor & Francis. DOI: 10.4324/9781315556062.

Ottósson, Kjartan G. 1990. *Íslensk málhreinsun: Sögulegt yfirlit [Icelandic linguistic purism: A historical overview]*. Íslensk málnefnd.

Penner, Ken M. 2015. *The verbal system of the Dead Sea scrolls: Tense, aspect and modality in Qumran Hebrew texts*, vol. 64 (Studia Semitica Neerlandica). Leiden: Brill. DOI: 10.1163/9789004298446.

Petré, Peter. 2010. The functions of *weorðan* and its loss in the past tense in Old and Middle English. *English Language & Linguistics* 14(3). 457–484. DOI: 10.1017/s1360674310000158.

Petré, Peter. 2014. *Constructions and environments: Copular, passive, and related constructions in Old and Middle English*. Oxford University Press. DOI: 10.1093/acprof:oso/9780199373390.001.0001.

Petré, Peter. 2016. Grammaticalization by changing co-text frequencies, or why [BE V-*ing*] became the 'progressive'. *English Language & Linguistics* 20(1). 31–54. DOI: 10.1017/s1360674315000210.

Platzack, Christer. 1985. Narrative inversion in Old Icelandic. *Íslenskt mál* 7. 127–144. https://timarit.is/page/5401918.

Pollak, Hans. 1930. Über den Gebrauch der Tempora in Ares Islendingabóc. *Arkiv för nordisk filologi* 46. 171–187. https://journals.lub.lu.se/anf/issue/view/1802.

Pustejovsky, James. 1998. *The generative lexicon*. Cambridge, MA: The MIT Press. DOI: 10.7551/mitpress/3225.001.0001.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Ragnarsdóttir, Hrafnhildur, Melina Aparici, Dalia Cahana-Amitay, Janet G van Hell & Anne Viguié-Simon. 2002. Verbal structure and content in written discourse: Expository and narrative texts. *Written Language & Literacy* 5(1). 95–126. DOI: 10.1075/wll.5.1.05rag.

Reichenbach, Hans. 1947. *Elements of symbolic logic*. New York: The Free Press. DOI: 10.2307/2268978.

Reinhart, Tanya. 1984. Principles of gestalt perception in the temporal organization of narrative texts. *Linguistics* 22. 779–809. DOI: 10.1515/ling.1984.22.6.779.

Richardson, Peter. 1995. Tense, Structure, and Reception in Þorsteins þáttr stangarhǫggs. *Arkiv för nordisk filologi* 110. 40–55. DOI: 10.2210/pdb7rlk/pdb.

Rieger, Gerd Enno. 1968. Die Spitzenstellung des finiten Verbs als Stilmittel des isländischen Sagaerzählers. *Arkiv för nordisk filologi*. 81–139.

Rögnvaldsson, Eiríkur. 1994. Breytileg orðaröð í sagnlið. *Íslenskt mál og almenn málfræði* 16–17. 27–66. https://timarit.is/page/5409929.

Rögnvaldsson, Eiríkur. 1995. Old Icelandic: A non-configurational language. *NOWELE* 26(1). 3–29. DOI: 10.1075/nowele.26.01rog.

Rögnvaldsson, Eiríkur. 1996. Word Order Variation in the VP in Old Icelandic. *Working Papers in Scandinavian Syntax* 58. 55–86. https://eirikur.hi.is/wordord.pdf.

Rögnvaldsson, Eiríkur. 2002. ÞAÐ í fornu máli–og síðar. *Íslenskt mál* 24. 7–30. https://timarit.is/page/5413834#page/n8/mode/2up.

Rögnvaldsson, Eiríkur & Sigrún Helgadóttir. 2011. Morphosyntactic tagging of Old Icelandic texts and its use in studying syntactic variation and change. In *Language Technology for Cultural Heritage: Selected Papers from the LaTeCH Workshop Series*, 63–76. DOI: 10.1007/978-3-642-20227-8_4.

Rosemeyer, Malte & Eitan Grossman. 2017. The road to auxiliariness revisited: The grammaticalization of finish anteriors in Spanish. *Diachronica* 34(4). 516–558. DOI: 10.1075/dia.16024.ros.

Salkie, Raphael. 1989. Perfect and pluperfect: What is the relationship? *Journal of Linguistics* 25(1). 1–34. DOI: 10.1017/s0022226700012081.

Sapp, Christopher D. 2019. Arrested development: Case attraction as a transitional stage from Old Icelandic demonstrative to relative *sá*. *Language* 95(1). e1–e40. DOI: 10.1353/lan.2019.0007.

Schmid, Hans-Jörg. 2015. A blueprint of the Entrenchment-and-Conventionalization Model. *Yearbook of the German Cognitive Linguistics Association* 3(1). 3–26. DOI: 10.1515/gcla-2015-0002.

Schmiedtová, Barbara, Christiane von Stutterheim & Mary Carroll. 2011. Language specific patterns in event construal of advanced second language speak-

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

ers. In Aneta Pavlenko (ed.), *Thinking and speaking in two languages*, 66–107. Berlin: De Gruyter. DOI: 10.21832/9781847693389-005.

Sigmundsson, Svavar. 1998. Staða rannsókna á máli 18. aldar (The status of research on 18th century Icelandic). *Vefnir* 1.

Sigurðardóttir, Aldís, Alex Speed Kjeldsen, Bent Chr. Jacobsen, et al. 2024. *Dictionary of Old Norse Prose Online*. https://onp.ku.dk/.

Sigurðsson, Halldór Ármann. 1988. From OV to VO: Evidence from Old Icelandic. *Working Papers in Scandinavian Syntax* 34. 49–55. DOI: 10.1007/bf00134679.

Sigurðsson, Halldór Ármann. 2018. Icelandic declarative V1: A brief overview. *Working Papers in Scandinavian Syntax*. 49–55. DOI: 10.1007/bf00134679.

Slobin, Dan I. 1994. Discourse origins of the present perfect. In William Pagliuca (ed.), *Perspectives on grammaticalization*, vol. 109 (Current Issues in Linguistic Theory), 119–133. Amsterdam: Benjamins. DOI: 10.1075/cilt.109.04slo.

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*. Cambridge, MA: The MIT Press. DOI: 10.7551/mitpress/4117.003.0013.

Smári, Jakob Jóh. 1920. *Íslenzk setningafræði [Icelandic syntax]*. Reykjavík: Bókaverzlun Ársæls Árnasonar.

Speyer, Augustin. 2010. *Topicalization and stress clash avoidance in the history of English*, vol. 69 (Topics in English Linguistics). Berlin: De Gruyter. DOI: 10.1515/9783110220247.

Starren, Marianne. 2017. What comes second. Cross-linguistic analyses of information structure in Dutch between English and German. In Pieter de Haan & Bettelou Los (eds.), *Word order change in acquisition and language contact: Essays in honour of Ans van Kemenade*, 241–262. Amsterdam: Benjamins. DOI: 10.1075/la.243.11sta.

Stofnun Árna Magnússonar í íslenskum fræðum. 2023. *Íslenskt textasafn*. Accessed: 06.01.2023. http://corpus.arnastofnun.is/.

Sturluson, Snorri. 1995. *Edda*. Trans. by Anthony Faulkes (Everyman's Library). London: J. M. Dent. DOI: 10.17104/9783406692888-7.

von Stutterheim, Christiane, Ralf Nüse & Jorge Murcia-Serra. 2002. Cross-linguistic differences in the conceptualisation of events. In *Information structure in a cross-linguistic perspective*, 179–198. Leiden: Brill. DOI: 10.1163/9789004334250_012.

von Stutterheim, Christiane & Ralf Nüse. 2003. Processes of Conceptualization in Language Production: Language-Specific Perspectives and Event Construal. *Linguistics* 41(5). 851–881. DOI: 10.1515/ling.2003.028.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

von Stutterheim, Christiane & Monique Lambert. 2005. Cross-linguistic analysis of temporal perspectives in text production. In Henriëtte Hendriks (ed.), *The structure of learner varieties*, 203–230. Berlin: De Gruyter. DOI: 10.1515/9783110909593.203.

von Stutterheim, Christiane, Martin Andermann, Mary Carroll, Monique Flecken & Barbara Schmiedtová. 2012. How grammaticized concepts shape event conceptualization in language production: Insights from linguistic analysis, eye tracking data, and memory performance. *Linguistics* 50(4). 833–867. DOI: 10.1515/ling-2012-0026.

Svavarsdóttir, Ásta, Sigrún Helgadóttir & Guðrún Kvaran. 2014. Language resources for Early Modern Icelandic. In *Proceedings of Language Resources and Technologies for Processing and Linking Historical Documents and Archives – Deploying Linked Open Data in Cultural Heritage (LRT7HDA)*, 19–25.

Talmy, Leonard. 1978. Figure and ground in complex sentences. *Proceedings of the First Annual Meeting of the Berkeley Linguistics Society*. 419–440. DOI: 10.3765/bls.v1i0.2322.

Talmy, Leonard. 1985. Lexicalization Patterns: Semantic Structure in Lexical Forms. In Timothy Shopen (ed.), *Language typology and syntactic description III: Grammatical categories and the lexicon*, 225–282. Cambridge: Cambridge University Press.

Thoroddsen, Jón. 1850. *Piltur og stúlka*. Sæmundur Bjarnason, Benedikt Sæmundsson, Hafdís Sæmundsdóttir & Bjarni Sæmundsson (eds.). Accessed: 11.01.2023. https://www.snerpa.is/net/roman/pilturog.htm.

Thoroddsen, Jón Þórðarson & Arthur Middleton Reeves. 1890. *Lad and lass: A story of life in Iceland*. London: Sampson Low, Marston, Searle & Rivington. DOI: 10.1163/9789004342361_002.

Thráinsson, Höskuldur. 2017. Developing a new perfect: the rise of the Icelandic *vera búinn að*-perfect. *Acta Linguistica Hafniensia* 49(2). 118–142. DOI: 10.1080/03740463.2017.1357267.

Trips, Carola & Eric Fuß. 2009. The syntax and semantics of the temporal anaphor *"then"* in Old and Middle English. In Artemis Alexiadou, Jorge Hankamer, Thomas McFadden, Justin Nuger & Florian Schäfer (eds.), *Advances in comparative Germanic syntax*, 171–196. Amsterdam: Benjamins. DOI: 10.1075/la.141.08the.

Wada, Naoaki. 1995. On the English Pluperfect: Does it really represent the Prepreterite Tense? *English Linguistics* 12. 96–124. DOI: 10.9793/elsj1984.12.96.

Wagener, Terje. 2017. *The history of Nordic relative clauses*, vol. 304 (Trends in Linguistics. Studies and Monographs [TiLSM]). Berlin: De Gruyter. DOI: 10.1515/9783110496536.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Wagner, Esther-Miriam, Ben Outhwaite & Bettina Beinhoff. 2013. *Scribes as agents of language change* (Studies in Language Change). Berlin: De Gruyter. DOI: 10.1515/9781614510543.

Walkden, George. 2021. *Parataxis and Hypotaxis in the History of English*. Accessed: 04.01.2023. http://walkden.space/ICEHL21.pdf.

Wallenberg, Joel C. 2016. Extraposition is disappearing. *Language* 92(4). e237–e256. DOI: 10.1353/lan.2016.0079.

Wallenberg, Joel C., Anton Karl Ingason, Einar Freyr Sigurðsson & Eiríkur Rögnvaldsson. 2011. *Icelandic Parsed Historical corpus (IcePaHC)*. http://linguist.is/icelandic_treebank/Icelandic_Parsed_Historical_Corpus_(IcePaHC). Version 0.9. DOI: 10.21248/jlcl.26.2011.153.

Wårvik, Brita. 1995. The ambiguous adverbial/conjunctions *þa* and *þonne* in Middle English: A discourse-pragmatic study of then and when in Early English saints' lives. In Andreas H. Jucker (ed.), *Historical pragmatics: Pragmatic developments in the history of English*, 345–57. Amsterdam: Benjamins. DOI: 10.1075/pbns.35.20war.

Westergaard, Marit. 2009. Word Order in Old and Middle English: The role of information structure and first language acquisition. *Diachronica* 26(1). 65–102. DOI: 10.1075/dia.26.1.03wes.

Wide, Camilla. 2002. *Perfect in dialogue: Form and functional potential of the vera búinn að + inf. construction in contemporary Icelandic*. Helsinki: University of Helsinki. (Doctoral dissertation). https://rafhladan.is/bitstream/handle/10802/6228/cwide_diss.pdf?sequence=1.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Table 8: Classification of examples from Benediktsson (2002)

Source	Example	Classification
Njáls saga, ch. 11	hann [Þorvaldur] var at at hlaða skútuna	Indet
Biskupa sögur, Jarteinabók Þorláks biskups helga, AM 645 4° A	MeN voro at. at tialda kirkio...	Durative
Íslendingasaga of Sturla Þórðarson	Einarr la ísarvm. oc var helgí. prestr. Scelivngs son at græða hann.	Durative
Bósa saga, ch. 8	...ok véru þeir allan dag upp at brjóta....	Durative
Bjarnar saga Hítdælakappa, ch. 34	Þorfinnr...kvað hann [Ásgrim] vera at telja silfr	Focal
Icelandic Book of Homilies	fyr óvinom sinom. oc þeím [er] at voro at pína hann.	Durative
Book of Homilies	En epter dagmól var hon at ver-alldlego verke...	Stative habitual
Book of Miracles, 1st Miracle	...at þeim þore travt at aeínom dege mynde verþa.	Stative habitual
Book of Miracles, 33rd Miracle	þa var til at hogva biargit. Af oxanom. Oc veret at mikiN hlvt dags	Durative
Auðunar þætta vestfiryka Morkinskinna (Ed. 1932: 185)	Einn dag er aleið varit gec Sveinn konungr...	Focal
Haralds saga hins hárfagra, ch. 8	Þeir hofðu verit at þrjú sumur at gera haug einn	Durative
Eyrbyggja saga, ch. 37	...en þeir Arnkell voro þa at at gera annat hlassit	Focal
Jómsvíkinga saga, ch. 8	En ec þottvmz vppe æga vef en þat var lín vefr...	Indet
Þiðriks saga af Bern, ch. 273	...oc erv þeir at allan dagin at taka hæstin oc geta æigi tekit	Durative
Örvar Odds saga	váru þeir þá at at herklæðaz...	Indet
Hákonar saga Hákonarsonar, ch. 236	voro þa Varbelger at at taka af þa lag sem efter voro brvarinnar	Indet

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Table 9: Classification of examples from Benediktsson (2002) (continued)

Source	Example	Classification
Þorgils saga skarða, in Sturlunga saga	...ok váru menn at, at kasta steinum í stofuvegginn	Durative
Egils saga einhenda ok Ásmundar berserkjabana, ch. 5	Kerling var at at renna mjólk	Durative
Finnboga saga hins ramma, ch. 24	Finnbogi var epter oc hrafn hinn litle hía honom...	Durative
Stjórn, ch. 183	...þiat hann ottaðiz at hann mvndi æigi na at briota	Indet
Codex Wormianus, Prologue to four grammatical treatises	...þvi verðr spvrt hverr kvað þa er fra liðr enn æigi hversv lengi var at verit	Stative habitual
Fóstbræðra saga, ch. 8	hann var at byrgja kviadyrrnar	Focal
Hrafnkels saga Freysgoða, 3. Kap	en konur váru at mjólka	Durative
Grettis saga, ch. 17	Hafliði fór til þeira skipverja þar sem þeir váru at ausa, ok mælti	Focal
Reykdæla saga ok Víga-Skútu, ch. 27	hann var at ok smíðaði skot um skála Skútu-	Durative
Droplaugarsona saga, ch. 3	þeir foru brott (ok) þangat sem þeir uoru at ok gerdu hlauss(in)	Indet
Ragnars saga loðbrókar, ch. 16	Ok er hann var at at reka flottann...	Focal
Völsunga saga, ch. 8	Ok er þeir voru at tyrfa hauginn...	Focal
Morkinskinna, CCI 6	...oc taca nv oc havGva <s> is fvr scipom sinom oc er þeir voro at þa melti [maþr]...	Focal
Morkinskinna, CCI 6	...oc er þar ál lang vm baggann oc torsott at leysa oc er hann lengi at...	Indet

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Table 10: Temporal subordinate clauses by text (IcePaHC)

	text	year	t.T	total_rT	relative_frequency	Genre
1	ntacts	1540	1185	36	30.38	rel
2	ntjohn	1540	1685	56	33.23	rel
3	eintal	1593	1552	40	25.77	rel
4	okur	1611	629	28	44.52	rel
5	olafuregils	1628	906	65	71.74	bio
6	gerhard	1630	701	25	35.66	rel
7	illugi	1650	1929	20	10.37	nar
8	pislarsaga	1659	324	41	126.54	bio
9	indiafari	1661	1388	81	58.36	bio
10	armann	1675	1018	8	7.86	nar
11	magnus	1675	204	6	29.41	bio
12	modars	1675	373	6	16.09	nar
13	skalholt	1680	870	33	37.93	nar
14	vidalin	1720	1112	73	65.65	rel
15	biskupasogur	1725	1105	35	31.67	nar
16	klim	1745	874	94	107.55	nar
17	fimmbraedra	1790	1603	21	13.10	nar
18	jonsteingrims	1791	1518	110	72.46	bio
19	hellismenn	1830	1411	9	6.38	nar
20	jonasedli	1835	163	9	55.21	sci
21	piltur	1850	1440	30	20.83	nar
22	hugvekjur	1859	1107	56	50.59	rel
23	orrusta	1861	1804	12	6.65	nar
24	torfhildur	1882	2000	25	12.50	nar
25	voggur	1883	130	8	61.54	nar
26	grimur	1888	625	24	38.40	nar
27	vordraumur	1888	759	41	54.02	nar
28	fossar	1902	1659	64	38.58	nar
29	leysing	1907	1520	56	36.84	nar
30	ofurefli	1908	1743	27	15.49	nar
31	arin	1920	1149	28	24.37	rel
32	margsaga	1985	1705	57	33.43	nar
33	sagan	1985	2008	57	28.39	nar
34	mamma	2008	1845	109	59.08	nar
35	ofsi	2008	1210	53	43.80	nar

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

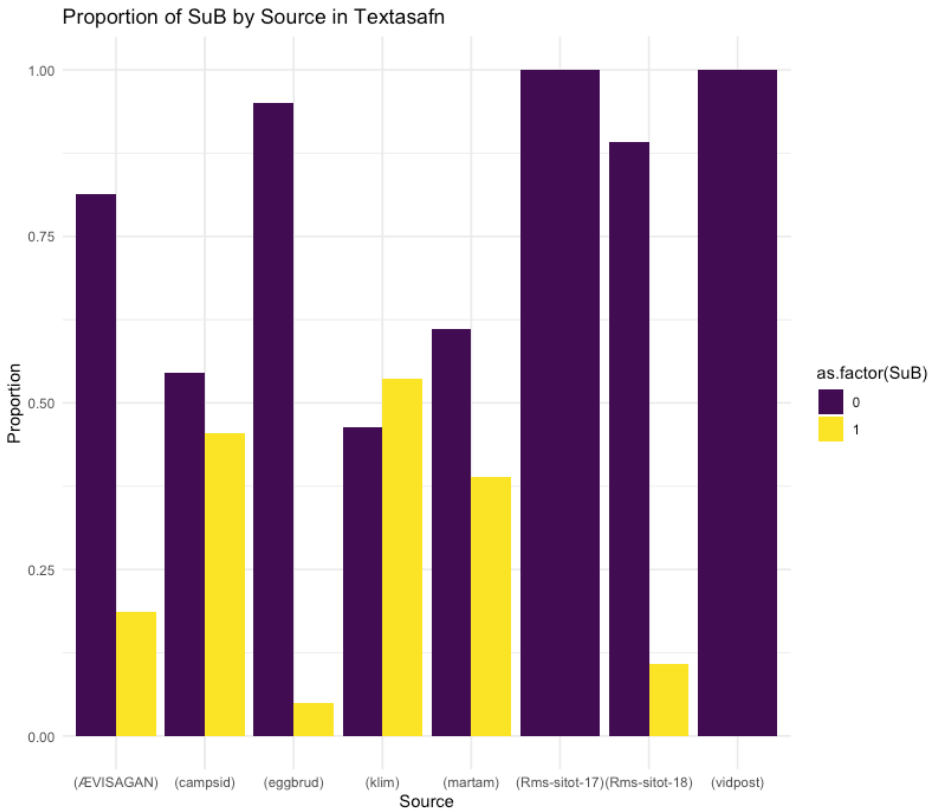


Figure 13: Prevalence of subordinate clauses expressing temporal sequentiality (*SuB*) in *Íslenskt textasafn* (16th-18th century subcorpus)

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Chapter 5

Narration and clause-internal adverbs: an information-structural watershed in Old Icelandic

Hannah Booth[a]

[a]Ghent University

Various discourse functions have been attributed to temporal/spatial adverbs in Early Germanic narrative texts, especially in Early West Germanic. By contrast, the role of such adverbs in Early North Germanic has been neglected, despite there being a rich attestation of saga narratives, ripe for studying such phenomena. In this paper, I investigate the behaviour of clause-internal *þá* (lit. 'then'), *þar* (lit. 'there') and *nú* (lit. 'now') in relation to discourse in Old Icelandic saga narratives. I show that these adverbs serve as an "information-structural watershed" (Krivonosov 1977, Grosz 2016), segmenting the clause into discourse domains, and show how this function can be modelled within Lexical Functional Grammar (Kaplan & Bresnan 1982, Bresnan et al. 2016). I also show that, diachronically, these adverbs are virtually in complementary distribution with the clause-initial expletive *það*, a later development which takes over some of the discourse function previously served by the clause-internal adverbs.

1 Introduction

Narratives as a genre bring their own specific stylistic characteristics which have been shown to interact with the expression of information structure in various ways (e.g. Carroll & Lambert 2003, Dimroth et al. 2010, Riester 2015). As one type of narrative, the Old Icelandic sagas have long been recognised by literary historians for their unique style and special rhetorical devices (e.g. Clover 1974, Ólason 2004, Phelpstead 2017), but so far their discourse-related properties have only

Hannah Booth. 2025. Narration and clause-internal adverbs: an information-structural watershed in Old Icelandic. In Ulrike Demske & Barthe Bloom (eds.), *Discourse structure and narration: A diachronic view from Germanic*, 133–172. Berlin: Language Science Press. DOI: ??

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

been scarcely investigated within theoretical linguistics (e.g. Hróarsdóttir 2009, Booth & Beck 2021). As such, their great potential to shed light on phenomena at the syntax-discourse interface remains largely untapped.

In this paper, I contribute to this area via an investigation of a small set of light adverbs (*þá*, lit. 'then'; *þar*, lit 'there'; *nú*, lit. 'now') which regularly appear clause-internally in Old Icelandic, especially in verb-initial clauses, e.g. (1).[1]

(1) a. *Sótti Ormur **þá** Sturlu að þessu máli og fóru*
pursued Ormur.NOM ÞÁ Sturla.ACC at this.DAT matter.DAT and went
***þá** margar orðsendingar millum þeirra Orms og*
ÞÁ many.NOM messages.NOM between they.GEN Ormur.GEN and
Sturlu...
Sturla.GEN
'Ormur pursued Sturla in this matter and there went many messages between them....' [IcePaHC: 1250, Sturlunga.396.269–270]

b. *Sátu þeir **þar** við er lengra fóru til. Var **þar** skipað*
sat they.NOM ÞAR PTCL REL further went to was ÞAR matched
saman þeim sem jafnsterkastir voru...
together they.DAT REL equally.strongest.NOM were
'Those who had come from further off would stay near by. They were matched together according to their strength...'
[IcePaHC: 1310, Grettir.162–163]

c. *Söðull var undir kviði niðri á hesti Grettis*
saddle.NOM was under belly.DAT lower.DAT on horse.DAT Grettir.GEN
*en í burt malurinn. Fer hann **nú** til leitar og fann*
and away knapsack.NOM went he.NOM NÚ to search.GEN and found
*eigi. Sér hann **nú** hvar maður gengur.*
NEG sees he.NOM NÚ where man.NOM goes
'Grettir's horse had got its saddle twisted round under its belly and the knapsack had gone, so he went to look for it, but couldn't find it. He sees a man walking by.' [IcePaHC: 1310, Grettir.237–240]

The characteristic presence of these elements in the Old Icelandic sagas has been previously noted (cf. Clover 1974: 71; Faarlund 1990: 71), but the precise discourse properties of such clauses have not been previously investigated. On the basis of corpus data from IcePaHC (Wallenberg et al. 2011) and MIcePaHC (Ingason

[1]Throughout the paper, I gloss the respective adverbs simply as ÞÁ, ÞAR, and NÚ to reflect the fact that, in clause-internal position, they often undergo some semantic weakening.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

2020), I claim that the clause-internal adverbs serve a specific discourse function, segmenting the clause into discourse domains, thus acting as an "information-structural watershed" (cf. Krivonosov 1977, Grosz 2016 on Modern German and van Kemenade & Los 2006 on Old English). I show how this specific function of clause-internal adverbs can be modelled within the parallel architecture of Lexical Functional Grammar (Kaplan & Bresnan 1982, Bresnan et al. 2016, Dalrymple et al. 2019), in which different types of linguistic information are represented at independent, interacting dimensions. I also examine the diachronic status of these clause-internal adverbs, showing that, across the history of Icelandic, they are virtually in complementary distribution with the clause-initial expletive *það*, which represents a later development (cf. Hróarsdóttir 1998, Rögnvaldsson 2002, Booth 2018, 2019, 2020). I relate this to a broader change whereby Icelandic gradually shifts its structural specification of discourse functions from the postfinite domain to the clause-initial position over time.

The paper proceeds as follows. In Section 2, I review previous work on adverbs and discourse in Early Germanic, especially in Early West Germanic where the bulk of previous research has been conducted, which serves as a backdrop to my investigation of Old Icelandic. Section 3 presents the corpus findings with respect to the frequency and distribution of clause-internal *þá*/*þar*/*nú*, and how this relates to the broader mapping correspondences between position and information structure in the postfinite domain. Section 4 outlines how the findings can be theoretically modelled within Lexical Functional Grammar and Section 5 presents some findings with respect to the diachrony of clause-internal adverbs in relation to broader changes in the grammar. Section 6 concludes the paper.

2 Adverbs, position and discourse in Early Germanic

Previous work on Early Germanic, in particular West Germanic (Old English, Old High German, Old Saxon), has highlighted the fact that certain temporal/spatial adverbs often serve particular discourse functions, especially in narratives (e.g. Foster 1975, Enkvist 1986, Enkvist & Wårvik 1987, van Kemenade & Los 2006, Trips & Fuß 2009, Wårvik 2013, Cichosz 2022). Crucially, some studies have highlighted the fact that the precise discourse functions of such adverbs vary by position (e.g. van Kemenade & Los 2006, Axel 2007, Donhauser & Petrova 2009, van Kemenade & Milićev 2011, Petrova 2011, van Kemenade & Links 2020, Catasso et al. 2021), with important differences between clause-initial adverbs and clause-internal adverbs, as I now discuss.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

2.1 Clause-initial adverbs

Various authors have claimed for Old English that *þa/þonne* (lit. 'then'), when occurring in the clause-initial position, marks a sequence of actions/events which represent the main line of the narrative ("foregrounding") (Foster 1975, Enkvist 1986, Enkvist & Wårvik 1987, Wårvik 2013, Cichosz 2022). An example of a continuous narrative whereby the main story line is carried forward by repeated use of clause-initial *þa* is shown in (2).

(2) Old English:
He sæde þæt he æt sumum cirre wolde fandian hu longe þæt land norþryhte læge, oþþe hwæðer ænig mon be norðan þæm westenne bude. ***Þa*** *for he norþryhte be þæm lande; let him ealne weg þæt weste land on ðæt steorbord & þa widsæ on ðæt bæcbord þrie dagas.* ***Þa*** *wæs he swa feor norþ swa þa hwælhuntan firrest faraþ.* ***Þa*** *for he þa giet norþryhte swa feor swa he meahte on þæm oþrum þrim dagum gesiglan.* ***Þa*** *beag þæt land þær eastryhte, oþþe seo sæ in on ðæt lond, he nysse wæðer buton he wisse ðæt he ðær bad westanwindes & hwon norþan & siglde ða east be lande swa swa he meahte on feower dagum gesiglan.* ***Þa*** *sceolde he ðær bidan ryhtnorþanwindes, for ðæm þæt land beag þær suþryhte, oþþe seo sæ in on ðæt land, he nysse hwæþer.* ***Þa*** *siglde he þonan suðryhte be lande swa swa he mehte on fif dagum gesiglan.* ***Ða*** *læg þær an micel ea up in on þæt land.*
'He said that at one occasion he wanted to find out how far that land extended northwards, or whether any man lived north of the wilderness. **Then** he travelled northwards along the coast; keeping all the way the waste land on the starboard and the open sea on the portside for three days. **Then** he was as far north as the whalehunters go furthest. **Then** he travelled still northwards as far as he could sail in another three days. **Then** the land turned east, or the sea into the land, he didn't know which, but he knew that he there waited for a wind from the west and somewhat from the north and sailed then east along the coast as far as he could sail in four days. **Then** he had to wait for a due north wind, because that land turned there directly to south, or the sea into the land, he didn't know which. **Then** he sailed from there southwards along the coast as far as he could sail in five days. **Then** there was a large river reaching up into the land.' [Or_1:1.14.5.226–235];
(Enkvist & Wårvik 1987: 234, as cited in Trips & Fuß 2009: 179–180)

Others have focussed on the discourse continuity/discourse-linking effect of clause-initial *þa/þonne* (e.g. Los 2000, van Kemenade & Los 2006, Trips & Fuß 2009).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Trips & Fuß (2009), for instance, argue that clause-initial *þa/þonne* function as discourse-anchoring temporal anaphors, which they relate to the fact that *þa/þonne* derive from former demonstratives (cf. Proto-Germanic **TO-*), and thus were always anaphoric/deictic in relation to something previously mentioned.

Similar observations have been made for certain clause-initial elements in Gothic (Klein 1994a) and Old Saxon (Linde 2009). Klein (1994a) claims that Gothic *þar-uh* (lit. 'there'+ *uh*) and *þan-uh* (lit. 'then' + *-uh*) function as "discourse articulators" which indicate narrative continuity, albeit associated with a change of grammatical subject. For Old Saxon, Linde (2009) observes that clause-initial *tho* (lit. 'then') and other adverbs such as *so* (lit. 'so') and *nu* (lit. 'now') have a similar function, signalling the continuation of the narration and guaranteeing progress in discourse.

2.2 Clause-internal adverbs

At the same time, a series of studies have shown that, when occurring clause-internally, such adverbs have a rather different function. Donhauser & Petrova (2009), for instance, examine the properties of clause-initial and clause-internal *tho* (lit. 'then') in Old High German and observe a clear difference: while clause-initial *tho* refers to a time span already established in the preceding section, e.g. (3a), clause-internal *tho* refers to a novel, indefinite time interval introduced as the topic time (cf. Klein 1994b) of a new episode, e.g. (3b) (see also discussion in Petrova 2011).

(3) Old High German:

a. ***tho*** *uuas man in hierusalem*
then was man in Jerusalem
'A man lived in Jerusalem at that time'
Latin: *homo erat in hierusalem* [T 37, 23]; (Petrova 2011: 219)

b. *uaas* ***thó*** *giuuortan in anderemo sambaztag*
was then become in another Sabbath
'It happened on another Sabbath'
Latin: *Factum est in alio sabbatum autem* [T 37, 23]; (Petrova 2011: 219)

Tying this in with the position of the finite verb, Petrova (2011: 224) argues that, in cases of clause-internal *tho*, a verb-initial (V1) structure is triggered by the novelty of the time interval referred to by *tho*, in line with the fact that V1 declaratives more generally are used to introduce a new situation (cf. Hinterhölzl & Petrova 2010).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

The special status of clause-internal *tho* and its close relation to V1 in Old High German is also observed by Axel (2007, 2009). Axel (2007: 156, 167) states that "most" Old High German V1 declaratives contain a clause-internal *tho* as "a very characteristic lexical feature". In such contexts, she observes, *tho* often has very weak semantics, as evidenced by the fact that it occurs with an additional temporal adverbial in the same clause, e.g. (4). Axel suggests that, in such contexts, *tho* serves as a narrative-emphatic particle, belonging to the broader, residual system of sentence particles exhibited in early Old High German texts, which is a vestige from an earlier period before the generalisation of verb-second (Axel 2007: 169–170, Axel 2009: 35–36).

(4) Old High German:
inti uuas ***tho*** *giheilit/ ira tohter* ***fon dero ziti/***
and was THO healed her daughter from that hour
'and her daughter was healed from that hour'
Latin: *& sanata est/filia illus ex illa hora/* [T 273,31]; (Axel 2007: 156)

The discourse function of clause-internal adverbs has also been highlighted for Old English in work by van Kemenade and colleagues (van Kemenade & Los 2006, van Kemenade et al. 2008, van Kemenade 2009, van Kemenade & Milićev 2011, van Kemenade & Links 2020). The central claim is that, in subordinate clauses, clause-internal *þa/þonne* ('then') functions as a "discourse partitioner", separating presupposed, given information from new, focussed information, cf. (5) (adapted from van Kemenade et al. 2008: 10).

(5) [$_{utterance}$ PRESUPPOSITION – *þa/þonne* – FOCUS]

The evidence for this claim comes from a series of quantitative corpus findings. For instance, van Kemenade & Los (2006: 232) observe that pronominal subjects in the postfinite domain occur overwhelmingly to the left to *þa/þonne*, e.g. (6); additionally, pronominal objects also regularly appear in this position, as in the example in (6).

(6) Old English:
forþæm he wenð þæt [he] [hi] ***þonne*** *ealle hæbbe*
because he knows that he them then all have
'because he know that he then had them all'
[coboeth,Bo: 24.56.16.1031]; (van Kemenade et al. 2008: 12)

By contrast, indefinite nominal subjects are preferred to the right of *þa/þonne* (van Kemenade & Los 2006: 237), e.g. (7); van Kemenade et al. (2008: 13) find

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

that indefinite NPs introducing new discourse entities do not occur to the left of *þa/þonne* in their data.

(7) Old English:
Gif ***ðonne*** *[hwelc mon] forbireð his synna*
if then any many forebears his sins
'If anyone then refrains from his sins'
[cocura,CP: 37.265.1.1719]; (van Kemenade et al. 2008: 13)

Definite nominal subjects can occur to the left or right of *þa/þonne* but, as van Kemenade et al. (2008: 13) observe, definite nominals occurring in the left position are +anaphoric. Via a statistical analysis, van Kemenade et al. (2008: 19) find that the most important trigger for a nominal to occur in the position to the left of *þa/þonne* is the presence of an antecedent in the previous discourse. The overall findings in these studies are related in van Kemenade & Los (2006: 233) to a broader claim that Old English is discourse-configurational (cf. Kiss 1995), i.e. is a language with designated structural positions for discourse functions such as topic and focus. This is reminiscent of work on modern German, where it has been argued that clause-internal discourse particles segment the postfinite domain into (aboutness) topic and focus fields (Krivonosov 1977, revisited by Grosz 2016), e.g. (8).

(8) Modern German:
Dann hat [Riko]$_{TOPIC}$ ***ja*** *[eine Frau]*$_{FOCUS}$ *geküsst*
then has Riko PTCL a woman kissed
'Then Riko has kissed a woman' (Grosz 2016: 338)

Elsewhere in Early West Germanic, Old High German/Old Saxon *tho* has also been claimed to serve an information-structural function, namely marking topicality in a specific verb-third construction of the form XP-*tho*-V (Axel 2007: 224–225, Catasso et al. 2021), e.g. (9).

(9) a. Old High German:
sie ***tho*** *antalengitun imo. neín*
they DO answered him.DAT no
'They said to him: 'No'.'
Latin *Responderunt ei: non* [T. 337]; (Catasso et al. 2021: 2))

b. Old Saxon:
Petrus thô gimahalde […]
Peter THO said
'Peter said […].' [Hel. XXXVIII, 3136]; (Catasso et al. 2021: 2))

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Catasso et al. (2021) observe that the clause-initial constituent in constructions like (9) is usually a pronominal shift topic, and they thus argue that *tho* is a topic marker in such contexts. Axel (2007: 225) similarly states that *tho* signals a change in topic in such contexts.

2.3 Summary

In sum, there is a good deal of work on Early West Germanic which has highlighted the special status of clause-internal adverbs in terms of their discourse functions. By comparison, the behaviour of clause-internal adverbs in Early North Germanic has generally been neglected, with only a few sparse references to their frequent occurrence in Old Norse/Icelandic texts. For instance, Clover (1974: 71) observes that there is "repeated interposition" of temporal/spatial adverbs in Icelandic translations of *chansons de geste* where the Old French original has none. Similarly, Faarlund (1990: 71), in relation to Old Norse texts more broadly, states that the adverb *þar* "often" occurs immediately after the finite verb, providing the examples in (10).

(10) Old Norse/Icelandic:

a. Konungs skuggsjá
Þá kemr ***þar*** *elldur af því jarni*
then comes ÞAR fire.NOM from that.DAT iron.DAT
'Then fire comes from that iron.'

b. Óláfs saga Tryggvasonar
Ok varð ***þar*** *mikit mannfall*
and was ÞAR great.NOM man.fall.NOM
'And there was a great loss of men'

c. Óláfs saga Tryggvasonar
Kemr ***þar*** *mikill fjǫldi manna*
comes ÞAR great.NOM multitude.NOM men.GEN
'There are many men coming' (Faarlund 1990: 71)

Nevertheless, the precise behaviour of clause-internal adverbs in Old Norse/ Icelandic, and the broader discourse-related properties of such clauses, has not been closely examined. In this paper, I address this gap in the existing literature on the morpho-syntax-discourse interface in Early Germanic.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

3 Old Icelandic *þá*, *þar* and *nú*

In this section, I examine the distribution and properties of three temporal/spatial adverbs in Old Icelandic saga narratives, namely *þá* (lit. 'then'), *þar* (lit. 'there') and *nú* (lit. 'now'). It is possible that there are other adverbs which fall into a class with these, but these three appear to be the most representative and most frequent clause-internally. The claims are made on the basis of a series of corpus investigations using IcePaHC (Wallenberg et al. 2011) and MIcePaHC (Ingason 2020). For IcePaHC, I restrict myself specifically to sagas dated to the period 1150–1450.[2] For MIcePaHC, I examine all texts dated to 1450 (all texts in the corpus represent saga narratives). Combined, this yields a total of 46 texts and 134,239 sentence tokens. I will refer to both corpora together as (M)IcePaHC.

3.1 Distribution of *þá*, *þar* and *nú*

All three adverbs occur both clause-initially, e.g. (11), and clause-internally, e.g. (12).

(11) a. ***Þá*** *gekk út kerling ein og hafði ullkamb í hendi.*
then went out woman.NOM one.NOM and had wool.comb.ACC in hand.DAT
'Then a certain woman went out and (she) had a wool comb in her hand.' [MIcePaHC: 1390, Graenlendingath.55–56]

b. ***Þar*** *sökk og niður lík Þorvalds…*
there sank also down corpse.NOM Þorvaldur.GEN
'There also the corpse of Þorvaldur sank…' [MIcePaHC: 1300, Njals.742]

c. ***Nú*** *komu þeir til Grænlands og eru með Eiríki rauða um veturinn.*
now came they.NOM to Greenland.GEN and are with Eiríkur.DAT red.DAT in winter.ACC.DEF
'Now they came to Greenland and (they) are with Eiríkur the Red for the winter.' [MIcePaHC: 1305, Eiriks.845–846]

(12) a. *Segir Ólafur* ***þá*** *ætlan sína…*
says Ólafur.NOM ÞÁ opinion.ACC his.ACC
'Ólafur says his opinion…' [MIcePaHC: 1275, Laxdaela.1179]

[2] From the IcePaHC saga texts I removed 1350.BandamennM and 1350.Finnbogi since these are duplicated in MIcePaHC.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

b. *Vésteinn gengur **þar** til húss...*
Vésteinn.NOM goes ÞAR to house.GEN
'Vésteinn goes to the house...' (MIcePaHC: 1400, Gisla.511]

c. *Fóru þeir **nú** sína leið.*
went they.NOM NÚ their.ACC way.ACC
'They went on their way.' [MIcePaHC: 1250, Egils.4036]

In the (M)IcePaHC data, clause-initial *þá/þar/nú* occurs at a frequency of 12.4% in matrix declarative V2 clauses, and 4.1% in subordinate declarative V2 clauses, see Table 1. Assuming that clause-initial *þá/þar/nú* serve a discourse-linking function and mark a sequence of events on the main line of the narrative as elsewhere in Early Germanic (see Section 2), it is not surprising that the phenomenon is more common in matrix clauses, since subordinate clauses typically present presupposed, backgrounded information (e.g. Hooper & Thompson 1973, Quirk et al. 1985, Matsuda 1998, Matić et al. 2014) and as such are less likely to express events on the main story line.

Table 1: Frequency of clause-initial *þá/þar/nú* in V2 matrix and subordinate declaratives in (M)IcePaHC saga narratives, 1210–1450

	matrix	subordinate
initial *þá/þar/nú*	10,276	1,738
all V2	82,687	42,256
% initial *þá/þar/nú*	**12.4%**	**4.1%**

Interestingly, if one isolates separately (i) subordinate clauses which are either non-conjunct clauses or initial conjunct clauses and (ii) subordinate clauses which are non-initial conjunct clauses, clause-initial *þá/þar/nú* are more frequent

Table 2: Frequency of clause-initial *þá/þar/nú* in V2 non-conjunct/first conjunct and non-initial conjunct subordinate declaratives in (M)IcePaHC saga narratives, 1210–1450

	non-/first conjunct	non-initial conjunct
initial *þá/þar/nú*	1,588	132
all V2	38,483	1,466
% initial *þá/þar/nú*	**4.1%**	**9.0%**

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

in non-initial conjunct clauses than their non-conjunct/first conjunct counterparts, see Table 2. Some examples of the relevant non-initial conjunct clauses with clause-initial *þá*/*þar*/*nú* are provided in (13).

(13) a. *En svo var djúpið mikið að það var*
and so was deep.water.NOM.DEF great.NOM that it.NOM was
*jafnskjótt er hann kom til grunna og [**þá** var þrotið*
immediately that he.NOM came to bottom.GEN and then was failed
örendi hans].
breath.NOM he.GEN
'And the deep water was so great that immediately he came to the bottom and then his breath failed.' [IcePaHC: 1210, Jartein.789]

b. *Svo er sagt að Háls Fjörleifarson gerði bú að*
so is said that Háls.NOM Fjörleifarson.NOM made home.ACC at
*Tjörnum í Ljósavatnsskarði og [**þar** gerðist vinátta*
Tjarnir.DAT in Ljósavatnsskarð.DAT and there became friendship.NOM
mikil í millum og Eysteins í Rauðaskriðu].
great.NOM between also Eysteinn.GEN in Rauðaskriða.DAT
'So it is said that Háls Fjörleifarson set up home at Tjarnir in Ljósavatnsskarð there became a great friendship between (him and) Eysteinn in Rauðaskriða.' [MIcePaHC: 1400, Reykdaela.74]

c. *Gerðist nú rómur mikill að því að eytt*
became now applause.NOM great.NOM because come.to.nothing
*væri vígsmálinu og [**nú** væri vörn*
was.SBJV manslaughter.case.DAT.DEF and now was.SBJV defence.NOM
framar en sókn].
more than prosecution.NOM
'There was now great applause because the manslaughter case had come to nothing and now there would be more deference than prosecution.' [MIcePaHC: 1300, Njals.9017]

This pattern, whereby a syntactic phenomenon which is overall more frequent in matrix than subordinate clauses is unusually frequent in subordinate clauses which are non-initial conjuncts, has been observed for another discourse-related phenomenon in Old Icelandic by Booth & Beck (2021). They find that narrative inversion, a special type of V1, which features a topical subject in the postfinite position, is robustly attested in subordinate clauses which are non-initial conjuncts, despite the fact that it is virtually unattested in subordinate clauses which

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

are either not coordinated, or first conjuncts. The examples which Booth & Beck (2021) provide are given here in (14).

(14) a. *"Eigi veit eg," segir Urðarköttur, "því að eg er*
NEG know I.NOM says Urðarköttur.NOM because I.NOM am
*ungur og [**kann** eg á fá skyn]"*
young.NOM and know I.NOM on few.ACC understanding.ACC
' "I do not know", says Urðarköttur, "because I am young and I do not have much knowledge" '
[IcePaHC: 1350, Finnbogi.631.351]; (as cited in Booth & Beck 2021: 23)

b. *Þú skalt... og mæla síðan þessum orðum við*
you.NOM shall and say then these.DAT words.DAT with
konunginn, að eg leiði hér eftir mér einn svein, [og
king.ACC.DEF that I.NOM lead here after I.DAT one.ACC boy.ACC and
***kalla** eg þar öngan mann...]*
call I.NOM ÞAR no.ACC man.ACC
'You should…and then say these words with the king that I lead here after me a certain boy and I call there no man…'
[IcePaHC: 1260, Jomsvikingar.1091]; (as cited in Booth & Beck 2021: 23)

Booth & Beck (2021: 25) offer an explanation for the occurrence of narrative inversion in subordinate non-initial conjuncts on the basis that narrative inversion is a marker of topic anaphoricity, such that it is only motivated in contexts where there is a directly preceding subordinate clause with its own topic as a potential antecedent. On the assumption that clause-initial *þá/þar/nú* serve a discourse-linking function and mark a sequence of events, a similar line of reasoning is a likely explanation for their higher frequency in subordinate non-initial conjuncts compared to their non-conjunct/first conjunct counterparts.

With respect to clause-internal *þá/þar/nú*, I examine their frequency across four environments, (i) V2 clauses with an overt subject, (ii) V2 clauses with an unexpressed subject, (iii) V1 clauses with an overt subject, and (iv) V1 clauses with an unexpressed subject.[3] The results for the two V2 environments are shown in Table 3, and for the two V1 environments in Table 4. In all four environments, clause-internal *þá/þar/nú* occur more frequently in matrix compared to subordinate clauses. Strikingly there are no examples of V2 subordinate clauses which

[3] I operationalise "clause-internal" in this context to mean either directly after the finite verb or one constituent later than the finite verb, and not the last constituent in the clause.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

have an unexpressed subject and also a clause-internal *þá/þar/nú*, see Table 3. Comparing V2 and V1 environments, clause-internal *þá/þar/nú* are a good deal more frequent in V1 clauses, and are particularly frequent in matrix V1 clauses with an overt subject (28.5%, see Table 4) and even more frequent in those with an unexpressed subject (34.4%, see Table 4). This is quantitative evidence which confirms the claim in Faarlund (1990: 71), that such adverbs "often" occur immediately after the finite verb in such contexts (see Section 2). More broadly, the relative commonness of the phenomenon in V1 declaratives specifically is in line with Axel's observations for Old High German, where she claims that the presence of clause-internal *tho* is a "very characteristic lexical feature" (Axel 2007: 167) (see discussion in Section 2).

Table 3: Frequency of clause-internal *þá/þar/nú* in V2 declaratives in (M)IcePaHC saga narratives, 1210–1450

	overt subject		unexpressed subject	
	matrix	subordinate	matrix	subordinate
internal *þá/þar/nú*	5,385	941	29	0
all V2	82,687	42,256	245	371
% internal *þá/þar/nú*	**6.5%**	**2.2%**	**11.8%**	**0.0%**

Table 4: Frequency of clause-internal *þá/þar/nú* in V1 declaratives in (M)IcePaHC saga narratives, 1210–1450

	overt subject		unexpressed subject	
	matrix	subordinate	matrix	subordinate
internal *þá/þar/nú*	4,311	38	145	12
all V1	15,135	827	422	341
% internal *þá/þar/nú*	**28.5%**	**4.6%**	**34.4%**	**3.5%**

Strikingly, there are cases where more than one of the adverbs (*þá/þar/nú*) occur clause-internally and adjacent to one another e.g. (15).

(15) a. *Og um daginn gerðist mæði mikil á þeim.*
and in day.ACC.DEF became exhaustion.NOM great.NOM to they.DAT
Þorgils var þó miklu hraustastur um
Þorgils.NOM was though much.DAT strongest.NOM in

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

allt. *Tekur* ***þá nú*** *að þyrsta* *mjög.*
everything.ACC begins ÞÁ NÚ to get.thirsty very
'And in the day a great exhaustion came upon them. Þorgils, though, was much the strongest. He begins to get very thirsty.'
[MIcePAHC: 1400, Floamanna.1313–1315]

b. *En er tíðindi þessi spyrjast fer Þorkell á*
but when news.NOM this.NOM is.reported goes Þorkell.NOM to
fund Þórarins og leitar þangað ráða og
meeting.ACC Þórarinn.GEN and seeks thither counsel.GEN and
meðferðar. *Hann segir: "Vera má* ***nú þá*** *að hún*
intercession.GEN he.NOM says be may NÚ ÞÁ that she.NOM
segi Ástríður að hann hafi eigi til engis
says.SBJV Ástríður.NOM that he.NOM had.SBJV NEG to nothing.GEN
risið á legginn."
risen to leg.ACC.DEF
'But when this news is reported, Þorkell goes to meet Þórarinn and seeks there counsel and intercession. He says: "It may be that she says, Ástríður, that he didn't get up for nothing."'
[MIcePaHC: 1350, Viga.538–539]

c. *Og þegar minnkar dirfð þeirra við þetta. Og nema*
and soon diminishes courage.NOM they.GEN with this.ACC and take
þeir að þjóna felmtinum. *Snerist* ***þar nú*** *skjótt*
they.NOM to serve fear.DAT.DEF turns ÞAR NÚ quickly
staðfestin í hræslu.
steadfastness.NOM.DEF in dread.DAT
'And soon their courage diminishes with this. And they take to serve their fear. The steadfastness quickly turns to dread.'
[IcePaHC: 1300, Alexander.901–903]

d. *Nú skal=tu þangað fara og reyna vini þína*
now shall=you.NOM thither go and try friends.ACC your.ACC
fyrir því að nú vættir mig að menn séu forkunnar
because now expect I.ACC that men.NOM are.SBJV remarkably
margir er lengi hefir dvalist. Mun ***nú þar*** *vera fjölmennt.*
many.NOM REL long have stayed may NÚ ÞAR be crowded.NOM
'Now you shall go there and try your friends because now I expect that there are remarkably many men who have stayed there for long. It may be crowded.' [MIcePaHC: 1300, Heidarviga.1166–1167]

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

e. *Þorgils var forstjóri fyrir búi í Torgum, þá er*
Þorgils.NOM was leader.NOM for farmstead.DAT at Torgar.DAT when
Þórolfur var eigi heima; hafði Þorgils ***þá þar*** *ráð*
Þórolfur.NOM was NEG home had Þorgils.NOM ÞÁ ÞAR authority.ACC
'Thorgils the leader of the farmstead at Torgar when Þórolfur was not at home; Þorgils had the authority there.' [MIcePaHC: 1250, Egils.515]

f. *Snorri góði kom þá að með flokk sinn. Var*
Snorri.NOM chief.NOM came ÞÁ PTCL with company.ACC his.ACC was
þar þá *Skafti í liði með honom…*
ÞAR ÞÁ Skafti.NOM in troop.DAT with he.DAT
'Chief Snorri came with his company. Also in the troop with him was Skafti…' [MIcePaHC: 1300, Njals.9384–9385]

The examples in (15) illustrate that all combinations and orders of two adverbs are attested in the (M)IcePaHC data (*þá nú, nú þá, þar nú, nú þar, þá þar, þar þá*). The examples which feature *þá nú* (lit. 'then now') and *nú þá* (lit. 'now then'), in (15a) and (15b) respectively, are particularly striking, as the literal meaning of the two adverbs presents a temporal contradiction. I take the attestation of sequences of clause-internal adverbs such as in (15a) and (15b) to indicate that such adverbs have grammaticalised to some extent and undergone some semantic weakening/bleaching. Next, I investigate the information-structural properties of clauses with clause-internal *þá/þar/nú*.

3.2 Mapping syntax to information structure in the postfinite domain

Any study of the syntax-information structure interface must first outline one's terminology and understanding of key information-structural concepts. In this paper, "topic" will be understood as roughly equivalent to "aboutness topic", i.e. the entity about which information is expressed (cf. "sentence topic", Reinhart 1981). In this context, the diagnostic tests provided by Götze et al. (2007: 165) can be used to identify the aboutness topic of an utterance, see (16).

(16) An NP X is the aboutness topic of a sentence S containing X if:

a. S would be a natural continuation to the announcement
Let me tell you something about X

b. S would be a good answer to the question
What about X?

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

c. S could be naturally transformed into the sentence
Concerning X, S'
where S' differs from S only insofar as X has been replaced by a suitable pronoun

Within the category of (aboutness) topic, I recognise different types of topic. One way of distinguishing between different subtypes of topic is to consider specifically the relation between a current topic and the topic of the immediately preceding utterance, i.e. the topic transition (see e.g. Daneš 1974 and "Centering Theory" in Grosz et al. 1995). In this paper, I recognise three subtypes of topic in this way, as defined in (17).

(17)
1. CONTINUING TOPIC: current topic is co-referential with the topic of immediately preceding utterance
   - *Helen arrived in Utrecht.* ***She*** *went directly to her favourite cafe in the city centre.*
2. SHIFT TOPIC: current topic is not co-referential with the topic of previous utterance
   - *Helen gave a book to Julie for her birthday.* ***The book*** *was a new collection of recipes by Mary Berry.*
3. SUBSECTIONAL TOPIC: current topic is an element of a previously introduced set of entities
   - *Helen and Julie were discussing their plans for the weekend.* ***Julie*** *wanted to go for a bike ride.*

Type 1 in (17), CONTINUING TOPIC, equates to what is elsewhere defined as "familiar topics" (e.g. Frascarelli & Hinterhölzl 2007). Type 3, SUBSECTIONAL TOPIC, is where the current topic constitutes a subset of a set of entities introduced in the previous discourse (cf. Hendriks 1996: 86); in this sense, it is roughly equatable with the notion of "contrastive topic", i.e. a topic which creates oppositional pairs with respect to other topics (cf. Frascarelli & Hinterhölzl 2007, Krifka 2007).

Another information-structural notion which will be relevant for this paper is INFORMATION FOCUS, as defined in (18) (cf. Reinhart 1981, Vallduví 1992, Götze et al. 2007).

(18) INFORMATION FOCUS: the part of the utterance which provides the new and missing information which is most relevant to the current discourse and serves to develop the discourse
- Q: *Where did you go in the Easter holidays?*
  A: *We went* ***to Italy****.*

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

To investigate the correspondence between syntax and information structure with respect to *þá/þar/nú*, I first extract matrix clauses with clause-internal *þá/þar/nú* and different types of nominative-marked subject in the postfinite domain (pronominal subject, proper name subject, nominal subject with definiteness marking, nominal subject without definiteness marking, quantified subject). I then examine the positional distribution of the different subject types with respect to *þá/þar/nú*.[4] An important detail to note is that, in Old Icelandic, the distinction between nominals with definiteness marking and nominals without definiteness marking does not map directly onto the distinction between semantically definite and semantically indefinite nominals, since definiteness marking was not fully obligatory with common nouns in Old Icelandic (Leiss 2000, 2007, Börjars et al. 2016). Thus, while all of the definite-marked nominals are expected to be semantically definite, the nominals without definite marking will not all be semantically indefinite, but may also include some semantically definite nominals. Since semantic definiteness is not annotated in the (M)IcePaHC data, it is not possible to extract semantically definite versus indefinite nominals and so I rely instead on the presence/absence of definiteness marking as a proxy, combined with manual examination of the examples in context.

The (M)IcePaHC data indicate that the five different types of subject show distinct and rather strong preferences in terms of their position with respect to clause-internal *þá/þar/nú*, see Table 5. Pronominal subjects, proper name subjects and definite-marked nominal subjects all favour the position to the left of *þá/þar/nú*, while nominal subjects without definiteness marking and quantified subjects favour the position to the right of the adverb. Next, I examine the relation between position and information structure for each of the five subject types.

3.2.1 Pronominal subjects

With respect to pronominal subjects, those that occur to the left of *þá/þar/nú* (the vast majority, see Table 5) are generally continuing topics, e.g. (19).

[4](M)IcePaHC also tags some non-nominative-marked arguments (typically experiencer arguments) as subjects, following the standard analysis of such arguments in Modern Icelandic as grammatical subjects (e.g. Zaenen et al. 1985). I exclude clauses annotated with such non-nominative subjects from this investigation of Old Icelandic, as it remains an ongoing debate to what extent non-nominative experiencer arguments already qualify as grammatical subjects at this early stage of the language (e.g. Barðdal & Eythórsson 2003, 2009, Barðdal et al. 2012, Schätzle et al. 2015, Schätzle 2018).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Table 5: Position of different types of (nominative-marked) subject relative to clause-internal *þá/þar/nú* in matrix declaratives in (M)IcePaHC saga narratives, 1210–1450

	V–SUBJ–*þá/þar/nú*		V–*þá/þar/nú*–SUBJ	
	n	%	*n*	%
pronominal subject	3,297	99.3%	22	0.7%
proper name subject	859	75.6%	277	24.4%
nominal subject$_{+\text{DEF}}$	78	94.0%	5	6.0 %
nominal subject$_{-\text{DEF}}$	275	21.9%	979	78.1%
quantified subject	10	2.5%	395	97.5%

(19) a. *Svo eignaðist Haraldur konungur Naumdælafylki og*
so acquired Haraldur.NOM king.NOM Naumdælir.county.ACC and
Hálogaland; setti [hann] ***þar*** *menn yfir ríki sitt.*
Hálogaland set he.NOM ÞAR men.ACC over kingdom.ACC his.ACC
'So King Haraldur acquired Naumdælir county and Hálogaland; he installed men over his kingdom.' [MIcePaHC: 1250, Egils.63]

b. *Maður heitir Ólafur og er Höskuldsson og er*
man.NOM is.called Ólafur.NOM and is Höskuldsson.NOM and is
[hann] ***nú*** *frægstur maður einnhver.*
he.NOM NÚ most.famous.NOM man.NOM anywhere
'There is a man called Ólafur, son of Höskuldur and he is the most famous man around.' [MIcePaHC: 1275, Laxdaela.1467–1468]

Of the 22 pronominal subjects which occur to the right of *þá/þar/nú*, 12 are instances of a dummy pronoun which is co-referential with a clausal argument later in the sentence, e.g. (20).[5]

(20) a. *Er* ***nú*** *[það]$_i$ ætlan þeirra bræðra [að venda til*
is NÚ it.NOM intention.NOM they.GEN brothers.GEN to turn to
hefnda við Helga Harðbeinsson því að hann veitti
avenge with Helgi.DAT Harðbeinsson.DAT because he.DAT dealt

[5]The dummy *það* is annotated in the (M)IcePaHC data as a subject, and so is included here, even though it is widely assumed that it does not qualify as a grammatical subject in Icelandic (e.g. Thráinsson 1979: 480–481; Platzack 1983; Maling 1988). Whatever the status of dummy *það* in Old Icelandic, I still include it here as an instructive example of an element which is neither a topic nor a focus.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Bolla banasár]$_i$.
Bolli.DAT death.blow.ACC
'It is the intention of the brothers to turn to attack in vengeance Helgi Harðbeinsson because he dealt Bolli his fatal blow.'
[MIcePaHC: 1275, Laxdaela.4401]

b. *Verður* ***nú*** *[það]$_i$ af ráðið [að brullaup skal vera að*
becomes NÚ it.NOM decided that wedding.NOM shall be at
Helgafelli að sex vikum sumars]$_i$.
Helgafell.DAT at six weeks.DAT summer.GEN
'It is decided that the wedding shall be held at Helgafell in the sixth week of summer.' [MIcePaHC: 1275, Laxdaela.4913]

At the same time, dummy pronouns which co-occur with a clausal argument can also occur to the left of *þá*/*þar*/*nú*, e.g. (21).

(21) *Fór [það]$_i$* ***þá*** *víða um héruð [hversu ólíkir þeir*
went it.NOM ÞÁ widely around district.ACC how unlike.NOM they.NOM
þóttu flestum mönnum vera í sínum framferðum]$_i$. Var [það]$_i$
seemed most.DAT men.DAT be in their.DAT procedures.DAT was it.NOM
þá *þegar vitra manna mál [að hvergi mundi vera*
ÞÁ soon wise.GEN men.GEN talk.NOM that nowhere would be
vænna til að leita en þar]$_i$...
more.likely.NOM to to search than there
'It went widely around the district how unlike most men they seemed to be in their procedures. It soon became the talk of wise men that nowhere would be more likely to search than there...'
[IcePaHC: 1210, Thorlakur.114–115]

The properties of *það* in contexts where it co-occurs with a clausal argument in Old Icelandic were examined by Booth (2018, 2019). Booth found that, unlike in certain present-day Germanic languages like German, where the parallel pronoun *es* is often assumed to mark the clausal argument as a topic (see Berman et al. 1998 and further references there), in early Icelandic the clausal argument in such contexts expresses discourse-new, non-topical information (Booth 2018: chapter 7, Booth 2019).[6] Thus the dummy pronoun in Old Icelandic does not merit

[6]In contrast to the traditional view on German *es*, Fuß & Hinterhölzl (2023: 44–45) point out that, while it can co-occur with a clausal argument which denotes a given situation, it can also occur with a clausal argument which introduces a new situation. This suggest that the relationship between such pronouns and information structure in Germanic is more complex than standardly assumed.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

a topical analysis.

Of the remaining 10 examples of pronominal subjects which occur to the right of *þá/þar/nú*, 9 involve misannotations and one involves Stylistic Fronting (cf. Maling 1990) of a nonfinite verb form to the prefield, shown here in (22).

(22) *Þorkell mælti: "Liðið er **nú** það. Gangi konur út úr*
Þorkell.NOM said passed is NÚ that.NOM goes.SBJV women.NOM out of
búðinni og viljum vér leita mannsins."
booth.DAT.DEF and wish we.NOM search man.GEN.DEF
'Þorkell said: "That is passed. Let the women go out of the booth and we want to search for the man."'
[MIcePaHC: 1350, Hallfredarmodruvallabok.10790–1081]

It remains an open debate as to whether Stylistic Fronting in Modern Icelandic is purely syntactic, without pragmatic effects (cf. Maling 1990, Holmberg 2000) or whether it contributes information-structural effects such as contrastive focus (cf. Hrafnbjargarson 2004) or backgrounding (cf. Egerland 2013). With respect to Old Icelandic, the picture is even more unclear, given that the phenomenon has not been studied in detail; in a preliminary survey, Booth (2018: chapter 4) showed that the syntactic constraints on Stylistic Fronting were much less restricted in early stages of Icelandic, and that Stylistic Fronting today may represent a remnant of a more general fronting process. As such, it is hard to say anything specific about this example with respect to Stylistic Fronting and the position of the pronominal subject. I simply note that it is another exception to the general preference for pronominal subjects to occur to the left of *þá/þar/nú*.

3.2.2 Proper name subjects

In the (M)IcePaHC data, proper name subjects show an overall preference for the position to the left of *þá/þar/nú*, but nevertheless occur to the right of *þá/þar/nú* in roughly a quarter of cases (see Table 5). Those proper name subjects which occur to the left of *þá/þar/nú* are generally either continuing topics, e.g. (23a), or shift topics, e.g. (23b).

(23) a. Preceding context: *Ólafur asked Gunnar to take care of himself and said that he had many ill-wishers "because you are now the most famous man in all the land". Gunnar thanked him for the gifts and counsel and rode home.*

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Situr [Gunnar] ***nú*** *heima nokkura hríð og er kyrrt.*
sits Gunnar.NOM NÚ home some.ACC time.ACC and is quiet.NOM
'Gunnar sits at home for some time and things are quiet.'
[MIcePaHC: 1300, Njals.14809]

b. Preceding context: *The King replies and says that he will never again accept hospitality from Þórólfur.*
Gekk [Þórólfur] ***þá*** *í brott og bjóst síðan til*
went Þórólfur.NOM ÞÁ away and prepared then to
heimferðar.
home.journey.GEN
'Þórólfur went away and prepared to travel home.'
[MIcePaHC: 1250, Egils.624–625]

By contrast, proper name subjects which occur to the right of *þá*/*þar*/*nú* show different information-structural properties and fall into two broad types. Firstly, many of the examples involve a proper name subject which is in information focus, e.g. (24).

(24) a. Preceding context: *Reginbald and Hercules go into battle against one another.*
Samson hans merkismaður bar djarflega fram
Samson.NOM he.GEN standardbearer.GEN carried boldly forth
merkið og hjó á báðar hendur og stökk
standard.ACC.DEF and struck at both.ACC hands.ACC and sprang
allt undan honum en í mót honum kom ***nú***
everything.NOM under he.DAT and towards he.DAT came NÚ
[Balabán] með merki Erkiles…
Balabán.NOM with standard.ACC Hercules.GEN
'Samson, his standardbearer, carried the standard boldly forth and struck at both hands and everything sprang from under him and towards him came Balabán with Hercules' standard…'
[IcePaHC: 1450, Vilhjalmur.81.1629–1632]

b. Preceding context: *And after that Gunnar, Karl, Svarthöfði and Ögmundur set off for home. There were two Eastmen besides Gunnar, and Karl was with the sixth man.*
Og er þeir komu ofan á hólana fyrir sunnan
and when they.NOM came up on hills.ACC.DEF before south
Brimnessá og til dælar þeirrar er ofan er og suður er
Brimnessá.ACC and to valley.GEN that.GEN REL up is and south is

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

frá ánni þá spretta þar upp fyrir þeim þrír
from river.DAT.DEF then spring there up before they.DEF three.NOM
*tigir manna og var **þar** [Ljótólfur].*
tens.NOM men.GEN and was ÞAR Ljótólfur.NOM
'And when they came to the top of the hills to south of Brimnessá and to the valley that is above and south of the river, there sprang up before them thirty of man and there was Ljótólfur.'
[MIcePaHC: 1450, Svarfdaela.3915–3916]

c. Preceding context: *Vémundur, Þorkell, Þórodður and Hánefur engage in battle with Hrói and his entourage.*
*Og nú kemur **þar** [Áskell] við þrjá tigu manna*
and now comes ÞAR Áskell.NOM with three.ACC tens.ACC men.GEN
og vildi sætta þá.
and wished reconcile they.ACC
'And now comes Áskell with thirty men and wanted to reconcile them.' [MIcePaHC: 1400, Reykdaela.2478–2479]

Secondly, proper name subjects which occur to the right of *þá/þar/nú* can be subsectional topics. This is illustrated in the continuous passage in (25), which relays action concerning Hallur and Þórólfur, and three proper name subjects which are subsectional topics occur to the right of *þá/þar/nú.*

(25) Preceding context: *One evening Hallur and Þórólfur came ashore and were to divide their catch. Hallur wanted to both divide and choose, because he thought himself to be the better man. Þórólfur didn't want to lose his share and did not hold back in his words. They exchanged some words and each insisted on their way.*
*Þrífur **þá** [Hallur] upp höggjárn er lá hjá honum og*
grasps ÞÁ Hallur.NOM up hewing.iron.ACC REL lay by he.DAT and
vill færa í höfuð Þórólfi. Nú hlaupa menn í milli
wishes strike to head.ACC Þórólfur.DAT now run men.NOM in between
þeirra og stöðva Hall en hann var hinn
they.GEN and restrain Hallur.ACC and he.NOM was DEF.NOM
óðasti og gat þó engu á leið komið að
most.stubborn.NOM and could though nothing.DAT on way.ACC come at
*því sinni og ekki varð fengi þeirra skipt. Réðst **nú***
this.DAT time.DAT and NEG became catch.NOM they.GEN divided went NÚ
[Þórólfur] á brott um kveldið en Hallur tók einn
Þórólfur.NOM away in evening.ACC.DEF but Hallur.NOM took one.ACC

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

upp fang *það* *er* *þeir* *áttu báðir* *því að* *þá*
up catch.ACC THAT.ACC REL they.NOM had both.NOM because then
kenndi *að ríkismunar.* *Fær* ***nú*** *[Hallur]* *sér*
recognised to power.difference.GEN gets NÚ Hallur.NOM himself.DAT
mann *í* *stað* *Þórólfs* *á* *skipið.*
man.ACC in place.ACC Þórólfur.GEN on ship.ACC.DEF.

'Hallur (not Þórólfur) picks up the hewing iron which lay by him and wishes to strike at Þórólfur's head. Now men run between them and restrain Hallur and he was the most stubborn man but could not get his way this time and their catch was not divided. Þórólfur (not Hallur) left in the evening but Hallur himself took up the catch which they had both had because he was recognised to be more powerful. Hallur (not Þórólfur) gets himself another man to replace Þórólfur on the ship.'

[MIcePaHC: 1275, Laxdaela.6721–6727]

Further examples of proper name subjects which are subsectional topics occurring to the right of *þá*/*þar*/*nú* are shown in (26).

(26) a. *Eftir það* *skiljast þeir* *konungur og* *Kjartan* *með*
after that.ACC part they.NOM king.NOM and Kjartan.NOM with
miklum *kærleik.* *Gengur* ***þá*** *[Kjartan]* *út* *á* *skip.*
great.DAT love.DAT goes ÞÁ Kjartan.NOM out to ship.ACC

'After that, the King and Kjartan part with great love. Kjartan (not the King) goes out to the ship.' [MIcePaHC: 1275, Laxdaela.3167–3168]

b. *Síðan gerðist* *Þórólfur* *handgenginn konungi* *ok* *gekk þar í*
then became Þórólfr.NOM retained.NOM king.DAT and went ÞAR in
hirðlög, *en* *Berðlu-Kári* *ok* *Eyvindur* *lambi,*
company.laws.ACC but Berðlu-Kári.NOM and Eyvindr.NOM Lamb.NOM
sonr *hans,* *fóru* *suðr* *með skip* *þat,* *er* *Þórólfr*
son.NOM he.GEN went south with ship.ACC that.ACC REL Þórólfur.NOM
hafði norðr haft. Fór ***þá*** *[Kári]* *heim* *til búa* *sinna*
had north had went ÞÁ Kári.NOM home to homestead.GEN his.GEN
ok *þeir* *Eyvindur* *báðir.*
and they.NOM Eyvindur.NOM both.NOM

'Then Þórólfur became a retainer to the kind and joined his followers by law but Berðlu-Kári and Eyvindur Lamb, his son, went south with the ship which Þórólfur had had in the north. Kári went back to his homestead, and Eyvindur too.' [MIcePaHC: 1250, Egils.5432–5434]

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

3.2.3 Nominal subjects with definite marking

Definite-marked nominal subjects occur overwhelmingly to the left of *þá/þar/nú* in the (M)IcePaHC data, see Table 5, and are generally shift topics, e.g. (27).

(27) a. Previous context: *But Ásgerðr, who was married to Þórólfr Skalla-Grímsson, was then with Arinbirn, his cousin. She and Þórólf had one young daughter, whose name was Þórdís,...*
...ok var [mærin] ***þar*** *með móður sinni.*
and was maiden.NOM.DEF ÞAR with mother.DAT her.DAT
'...and the maiden was with her mother.'
[IcePaHC: 1250, Thetubrot.29]

b. Previous context: *Þorgils awakes and thinks what he will have there, and now he remembers that he gave Þór a certain calf a long time ago. Þorgils tells this to Þóreyja...*
...og var [þetta] ***þá*** *gamall uxi...*
and was this.NOM ÞÁ old.NOM OX.NOM
'...and this (calf) was an old ox...' [MIcePaHC: 1400, Floamanna.1025]

The 5 examples of definite-marked nominal subjects which occur to the right of *þá/þar/nú* involve proximal demonstratives which occur as the subject of "happen" predicates, e.g. (28a), or proximal demonstratives which co-occur with and are co-referential with a clausal argument later in the sentence, e.g. (28b) and (28c).

(28) a. *Og nú fyrir sakir vinsælda hans og hann var*
and now for sakes.ACC popularity.GEN he.GEN and he.NOM was
sínum mönnum ástfólginn, þá játtu þeir þessu, að
his.DAT men.DAT beloved.NOM then agreed they.NOM this.DAT that
konungur skyldi ráða, og fór ***nú*** *[þetta] fram.*
king.NOM should decide and went NÚ this.NOM forth
'And now because of his popularity and because he was beloved of his men, they agreed that the king should decide, and this was done.'
[IcePaHC: 1260, Jomsvikingar.61–62]

b. *Réðst* ***nú*** *[þetta]$_i$ [að þeir Þorsteinn og Lambi*
decided NÚ this.NOM that they.NOM Þorsteinn.NOM and Lambi.NOM
skulu ráðast með Þorgísli til ferðar]$_i$...
should set-forth with Þorgísl.DAT to journey.GEN
'It is decided that Þorsteinn and Lambi should set off with Þorgísl on the journey...' [MIcePaHC: 1275, Laxdaela.4435]

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

c. *Fór* ***nú*** *[þetta]$_i$ fram [að Ólöf var gefin Þórði]$_i$...*
went NÚ this.NOM forth that Ólöf.NOM was given Þórður.DAT
'Things proceeded such that Ólöf was given to Þórður...'
[MIcePaHC: 1400, Thordar.1882]

In these contexts, the proximal demonstrative appears to be an alternative dummy pronoun, alongside *það*, cf. (20) and (21) above. In addition to examples like (28), there are also examples in the (M)IcePaHC data where *þetta* as a dummy pronoun precedes *þá*/*þar*/*nú*, e.g (29). This mirrors the behaviour of the dummy pronoun *það*, which can also appear to the left or right of *þá*/*þar*/*nú*, cf. (20) and (21) above.

(29) *Fór [þetta]$_i$* ***þá*** *fram [að grið voru sett með*
went this ÞÁ forth that settlements.NOM were established with
mönnum þar til að hver kæmi til síns heima]$_i$.
men.DAT there to that each.NOM came.SBJV to his.GEN home.GEN
'Things proceeded such that settlements were established with the men so that each would go to his home.' [MIcePaHC: 1300, Eyrbyggja.2596]

3.2.4 Nominal subjects without definite marking

Nominal subjects which occur without definite marking generally prefer the position to the right of *þá*/*þar*/*nú*, see Table 5, where they are typically in information focus, e.g. (30).

(30) a. Previous context: *But as soon as she saw the church at Skálaholt, she felt herself become more comfortable than she had previously ever been in her pain. She came there some nights before the church day...*
En þar var ***þá*** *[fjölmenni mikið]...*
but there was ÞÁ crowd.NOM great.NOM
'But there was a great crowd there...' [IcePaHC: 1210, Jartein.680]

b. Previous context: *Þorleifur shot from a bow and he was very likely to wound but nothing came of Jósteinn's shots and Þorleifur was able to hurt him. Deacon Þórður Símonarson held the shield for Þorleifur.*
Var ***þar*** *[harður bardagi]...*
was ÞAR hard.NOM battle.NOM
'There was a hard battle...' [IcePaHC: 1250, Sturlunga.391.123]

However, there are also examples where such a nominal subject without definite marking occurs to the left of the adverb, see Table 5. Such examples fall

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

into two broad types. Firstly, there are contexts where the subject is an occupation noun which is semantically definite though not morphologically marked for definiteness. Such subjects can be continuing topics, e.g. (31a) or shift topics, e.g. (31b), but they can also be subsectional topics, e.g. (32). In particular, the examples in (32) show that subsectional topics can appear to the left of *þá*/*þar*/*nú*, as well as to the right, cf. (25) and (26) above.

(31) a. Preceding context: *The king went north to Trondheim in the autumn. Then Þórólfur asks for leave to travel north to Halogaland to visit the gifts he had received in the summer from Bárður, his kinsman. The king grants that and gives him a message and token that Þórólfur should have all the property that Bárður gave him, letting it be known that the gift was made on the King's advice and that he wants it to be left so.*
Gerir [konungur] ***þá*** *Þórólf lendan mann…*
makes king.NOM ÞÁ Þórólfur.ACC landed.ACC man.ACC
'The king makes Þórólfur a landholder…' [MIcePaHC: 1250, Egils.353]

b. Preceding context: *Grettir fights with Skeggi, and Grettir strikes his head with an axe.*
Féll [húskarl] ***þá*** *dauður til jarðar.*
fell man.servant.NOM ÞÁ dead.NOM to ground.GEN
'The man servant (= Skeggi) falls dead to the ground.'
[IcePaHC: 1310, Grettir.273]

(32) a. *Og nú ganga þeir í braut af stefnunni, konungur*
and now go they.NOM away from meeting.DAT.DEF king.NOM
og jarl, og nú eftir það var jarl ***þar*** *með honum*
and earl.NOM and now after that.ACC was earl.NOM ÞAR with he.DAT
nokkora hríð í mikilli sæmd.
some.ACC time.ACC in great.DAT honour.DAT
'And now they depart from the meeting, the King and the Earl, and now after that the Earl was with him for some time in great honour.'
[IcePaHC: 1260, Jomsvikingar.299–300]

b. Preceding context: *A farmer and Þormóður are in conversation with one another.*
Snýr [bóndi] ***þá*** *utar eftir hlöðunni…*
turns farmer.NOM ÞÁ outside after storehouse.DAT.DEF
'The farmer turns back towards the storehouse…
[MIcePaHC: 1305, Fostbraedra.3301]

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

The second broad group of subjects which are annotated as nominals which lack definiteness marking in the (M)IcePaHC data and which can occur to the left of *þá*/*þar*/*nú* are impersonal pronouns, e.g. (33).

(33) a. *Riðu [menn]* ***þá*** *heim af þinginu…*
rode men.NOM ÞÁ home from assembly.DAT.DEF
'People rode home from the assembly…' [IcePaHC: 1310, Grettir.1411]

b. Previous context: *One autumn, Urðarköttur went out each evening and wouldn't come in before the night was late.*
Vita [menn] ***nú*** *eigi hvað hann gerir.*
knew men.NOM NÚ NEG what he.NOM does
'People didn't know what he was doing.'
[MIcePaHC: 1350, Finnboga.310]

However, this is not the only position available to impersonal pronouns; they can also occur to the right of *þá*/*þar*/*nú*, e.g. (34).

(34) a. *…og fóru* ***þá*** *[menn] meðal þeirra og voru þá grið sett.*
and went ÞÁ men.NOM among they.GEN and was ÞÁ truces.NOM set
'and people went among them and there were set truces.'
[IcePaHC: 1250, Sturlunga.436.1665–1666]

b. *Gjöra* ***nú*** *[menn] mikinn róm að máli Barða…*
make NÚ men.NOM great.ACC noise.ACC at case.DAT Barði.GEN
'People make a lot of noise about Barði's case…'
[MIcePaHC: 1300, Heidarviga.1080]

3.2.5 Quantified subjects

Quantified subjects, which cannot generally be topical (e.g. Frey 2004), overwhelmingly favour the position to the right of *þá*/*þar*/*nú*, see Table 5 and the examples in (35).

(35) a. *"Það munu* ***þá*** *[sumir menn] mæla," segir Höskuldur,*
that.ACC will ÞÁ some.NOM men.NOM say says Höskuldur.NOM
"að eg flýi þaðan fyrir hræðslu sakir og vil eg
that I.NOM flee.SBJV thence for fear.GEN sakes.ACC and will I.NOM
það eigi."
that.ACC NEG
"'Some men will say," says Höskuldur, "that I am fleeing from there for fear's sake, and I don't want that."' [MIcePaHC: 1300, Njals.6602]

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

b. *Eru* ***nú*** *[fáir slíkir menn] í yðvarri ætt sem*
are NÚ few.NOM such.NOM men.NOM in your.DAT generation as
Bolli er.
Bolli.NOM is.
'There are few men of your generation who are such as Bolli is.'
[MIcePaHC: 1275, Laxdaela.10072]

Nevertheless, there are 10 examples in the (M)IcePaHC data where a quantified subject precedes *þá*/*þar*/*nú* in the postfinite domain, see Table 5. On closer inspection, these either involve misannotations or a quantified subject which broadly translates as 'people' (impersonal pronoun), e.g. (36).[7]

(36) a. Preceding context: *The King calls a meeting at Eyrar and urges the local people to convert. Local farmers prepare to do battle with the King rather than convert. The King scares the farmers and they surrender to him.*
...og var [margt fólk] ***þá*** *skírt.*
and was much.NOM people.NOM ÞÁ baptised
'...and many people were baptised' [MIcePaHC: 1275, Laxdaela.2890]

b. Preceding context: *There was a man called Þórólfur who lived under Spákonufell. He had a fortuneteller, Þórdís, who was mentioned earlier.*
Þóttust margir ***þar*** *traust mikið eiga er*
seemed many.NOM ÞAR confidence.NOM much.NOM have where
hún var.
she.NOM was
'Many seemed to have much confidence wherever she was.'
[MIcePaHC: 1350, Kormaks.1636]

3.2.6 Summary

In sum, on the basis of the (M)IcePaHC data, the correspondence between syntax and information structure in the postfinite domain can be schematised as in (37). (37) shows a mixed picture, whereby certain information-structure roles are associated with a single position (e.g. information foci, right of *þá*/*þar*/*nú*;

[7] I do not rule out the possibility that quantified subjects in the different positions relative to *þá*/*þar*/*nú* have potentially distinct information-structural characteristics, since quantified noun phrases can have partitive or existential readings (see e.g. Jónsson 2002: 79 on Modern Icelandic). However, the number of examples with a quantified subject to the left of *þá*/*þar*/*nú* (which are not misannotated) are too few to draw any meaningful conclusions here.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

continuing/shift topics, left of *þá/þar/nú*) whereas others, such as subsectional topics, can appear freely in either position. The finding for subsectional topics is particularly striking, as they can occur both in the position associated with continuing/shift topics, and also in the position associated with information focus. One explanation for this behaviour could be that subsectional topics, which roughly equate to the notion of contrastive topics, are often assumed to involve a combination of topic and focus (Krifka 2007: 267). The finding that both dummy pronounsdummy pronoun and impersonal pronouns are not restricted in terms of their position with respect to *þá/þar/nú* is also interesting, since these elements are neither straightforwardly analysable as topic nor focus in information-structural terms.

(37) $$V_{\text{finite}} - \left\{ \begin{array}{c} \text{continuing topic} \\ \text{shift topic} \\ \\ \text{subsectional topic} \\ \text{dummy pronoun} \\ \text{impersonal pronoun} \end{array} \right\} - \textit{þá/þar/nú} - \left\{ \begin{array}{c} \text{information focus} \\ \\ \\ \text{subsectional topic} \\ \text{dummy pronoun} \\ \text{impersonal pronoun} \end{array} \right\}$$

Overall, clause-internal *þá/þar/nú* serve an information-structural watershed function in the postfinite domain in Old Icelandic (cf. Krivonosov 1977, Grosz 2016 on modern German; van Kemenade & Los 2006, van Kemenade et al. 2008, van Kemenade 2009, van Kemenade & Milićev 2011, van Kemenade & Links 2020 on Old English). However, this effect appears to be rather subtle in Old Icelandic. Whereas in modern German clause-internal discourse particles have been argued to partition topic and focus (cf. (8) above), and in Old English clause-internal adverbs have been argued to separate presupposition from focus (cf. (5) above), in Old Icelandic the watershed function appears to apply specifically to information focus and particular subtypes of topic. In the next section, I explore how these subtle correspondences between position and information structure can be modelled within Lexical Functional Grammar.

4 An LFG approach

4.1 The LFG architecture

Lexical Functional Grammar (LFG, Kaplan & Bresnan 1982, Bresnan et al. 2016, Dalrymple et al. 2019) is a "declarative" approach to grammar (cf. Levine & Meurers 2006, Sells 2021). This means that LFG does not commit to any procedural mechanism for deriving linguistic representations, and that all information in the

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

model is simultaneously present in parallel ("model-theoretic" approach, cf. Pullum & Scholz 2001). Different types of linguistic information are represented at independent dimensions which are related to each other within an overall projection architecture, see Figure 1. Each dimension differs in terms of its formal representation and must satisfy certain constraints. The core components of syntactic representation are c(onstituent)-structure, which captures information about category and constituency, and f(unctional)-structure, which captures abstract functional information. A third dimension which is relevant to this paper is i(nformation)-structure.

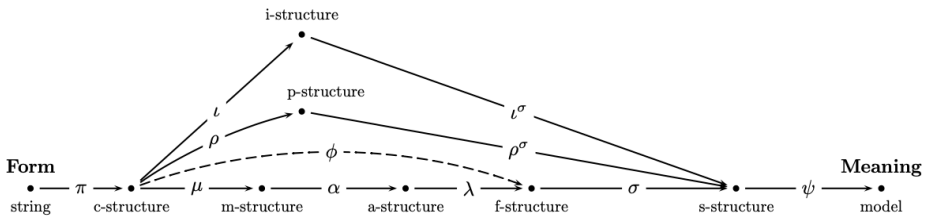


Figure 1: Parallel projection architecture of LFG (Asudeh 2006: 369)

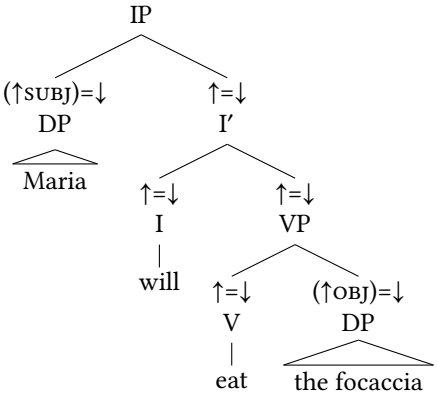
LFG's c-structure is represented as a tree diagram, while f-structure and i-structure are represented as attribute-value matrices. An example c-structure/f-structure/i-structure triad for Modern English is shown in (38). The abstract functional information represented at f-structure includes grammatical functions (GFs), e.g. SUBJ(ect) and OBJ(ect), as well as grammatical features, e.g. TENSE and DEF(initeness). A special type of functional feature is PRED, which is a pointer into the semantics of a predicate, takes a semantic form as its value, and captures the argument(s) (if any) a predicate requires in terms of grammatical function. i-structure represents information-structural information such as the discourse functions TOPIC and FOCUS.[8] Since i-structure is an underdeveloped part of the LFG architecture, it remains an open question as to what precise discourse function should be captured here (see e.g. Dalrymple et al. 2019: chapter 10, Zaenen Forthcoming). In this paper, I will assume that specific subtypes of topic (e.g. continuning topics/shift topics/subsectional topics) and focus (e.g. information foci) are represented as distinct discourse functions at i-structure.

[8]The PRED-FN notation in the i-structure in (38), as opposed to the standard PRED feature from f-structure, indicates that instead of projecting the entire PRED feature, including its argument structure, only the basic meaning of the PRED is projected to i-structure (cf. Kaplan & Maxwell 1996, King 1997, Butt et al. 2016).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(38) Q: What will Maria eat?
A: Maria will eat the focaccia.

c-structure:



f-structure:

$$\begin{bmatrix} \text{PRED} & \text{`EAT <SUBJ, OBJ>'} \\ \text{TENSE} & \text{FUT} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{`MARIA'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{`FOCACCIA'} \\ \text{DEF} & + \end{bmatrix} \end{bmatrix}$$

i-structure:

$$\begin{bmatrix} \text{TOPIC} & \begin{bmatrix} \text{PRED-FN} & \text{`MARIA'} \end{bmatrix} \\ \text{FOCUS} & \begin{bmatrix} \text{PRED-FN} & \text{`FOCACCIA'} \end{bmatrix} \end{bmatrix}$$

Within LFG's projection architecture, see Figure 1, all linguistic dimensions are present in parallel and a number of functions map between the various dimensions. With respect to the correspondence between c-structure and f-structure, I uncontroversially assume that this is formally handled via the correspondence function ϕ, whereby c-structure nodes are related to f-structures (Bresnan et al. 2016, Dalrymple et al. 2019). In a language like English where grammatical functions (e.g. subject, object) are structurally specified, functional annotation on specific c-structure positions exclusively associate a position with a specific grammatical function at f-structure, see (38). In the c-structure tree in (38), ↓ and ↑ are metavariables over f-structure variables and serve to relate every node in the c-structure to its corresponding f-structure. ↓ denotes the f-structure corresponding to that node itself, and ↑ denotes the f-structure corresponding to that node's mother node. As such, SpecIP bears an annotation which relates that node to the SUBJ function of the maximal f-structure; the annotation on the complement of V relates that node to the OBJ function.[9]

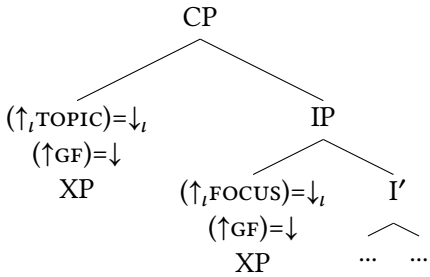
In a language like English, discourse functions such as TOPIC and FOCUS are not generally associated with fixed positions but are expressed via e.g. prosody. However, many languages exhibit "discourse configurationality" (cf. Vilkuna 1989, Kiss 1995), where discourse functions are associated with specific positions. In

[9]Additionally, two c-structure nodes in a mother-daughter relation may correspond to the same f-structure, in which case they are annotated as ↑=↓: this indicates that the functional information associated with a given node is the same as the functional information associated with that node's mother node.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

LFG, this can be modelled by annotations which exclusively associate certain c-structure positions with discourse functions at i-structure, e.g. (39); the arrows annotated with ι indicate projection to i-structure, and GF stands for any grammatical function (e.g. SUBJ, OBJ).[10]

(39)



4.2 Modelling the information-structural watershed

The relevant question in the context of this paper is how the information-structural watershed function of clause-internal *þar/þá/nú* in Old Icelandic can be modelled within LFG. For Old Icelandic, I model the postfinite domain in terms of a flat structure within I', see the c-structure in (40), following previous accounts of Icelandic clause structure within LFG, including Old Icelandic (Sells 2001, 2005, Booth & Schätzle 2019, Booth 2021).[11] One possibility for capturing the information-structural partitioning of the postfinite domain, and the function of clause-internal *þar/þá/nú*, is to assume discourse configurationality (see Section 4.1), whereby the position to the left of *þar/þá/nú* is uniquely associated with TOPIC at i-structure, and the position to the right of *þar/þá/nú* with FOCUS. This possibility is shown in (40).[12]

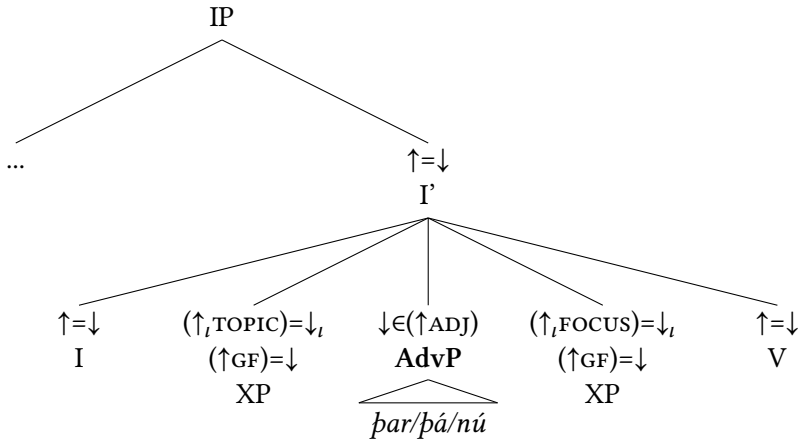
[10] In this paper, I follow Butt & King (1997) in assuming a model where i-structure projects from c-structure as defined by the function ι, which can be considered as a parallel to the ϕ function which relates c-structure nodes to f-structures. This is in line with Asudeh (2006), see again Figure 1. For a different proposal, see Dalrymple et al. (2019) who present a model where i-structure instead projects from s-structure.

[11] In (40) the adverb position bears the annotation ↓∈(↑ADJ), which specifies that the f-structure contributed by the adverb is a member of a set of values for the grammatical function ADJ(unct) in the maximal f-structure; this allows for multiple adjuncts in a clause.

[12] A similar discourse-configurational approach is employed by Mahowald (2011) in LFG in relation to the Old English postfinite domain and *þa/þonne*, based on the findings by van Kemenade & Los (2006) (see Section 2).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(40)



However, capturing the syntax-information structure correspondence in the postfinite domain in terms of dedicated topic and focus positions either side of *þar/þá/nú* is unsatisfactory in many ways, given the findings from the (M)IcePaHC data outlined in Section 3. Firstly, the position to the left of *þar/þá/nú* is not an exclusive topic position, since it can also host elements which cannot be construed as topics, such as dummy pronouns and impersonal pronouns, see (37) above. Secondly, the position to the right of *þar/þá/nú* is not an exclusive focus domain, but can also host subsectional topics, and again dummy pronouns and impersonal pronouns which cannot either be construed as foci. As such, an alternative to straightforward discourse configurationality as in (40) appears to be motivated in this context.

The alternative I suggest in this paper takes advantage of the fact that LFG provides for alternative ways of expressing ordering dependencies, besides strictly c-structure relations. One such alternative which has been developed and applied in various contexts is f(unctional)-precedence, which essentially constrains the ordering of c-structure constituents in terms of the f-structures they map onto, (cf. Bresnan 1984, Kaplan 1987, Kaplan & Zaenen 1989). f-precedence is formally defined as in (41).

(41) f-precedence
f_1 f-precedes f_2 if and only if all c-structure nodes corresponding to f_1 precede all c-structure notes corresponding to f_2
(Dalrymple et al. 2001: ex. (11))

For example, Zaenen & Kaplan (1995: 229–230) employ f-precedence in verb-final West Germanic varieties (e.g. Zürich German), where verbs in infinitival constructions cannot precede their nominal arguments, but where the verb and its

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

nominal arguments do not necessarily have to be adjacent. In such contexts, Zaenen & Kaplan (1995: 230) employ the c-structure rule in (42), where the V node is associated with an f-precedence constraint which states that the f-structure associated with the V node cannot f-precede its mother node's nominal grammatical functions (NGF), i.e. the nominal arguments of the particular verb.

(42) $\text{V'} \longrightarrow \underset{\downarrow \not\prec_f (\uparrow \text{NGF})}{\text{V}}$ (Zaenen & Kaplan 1995: 230)

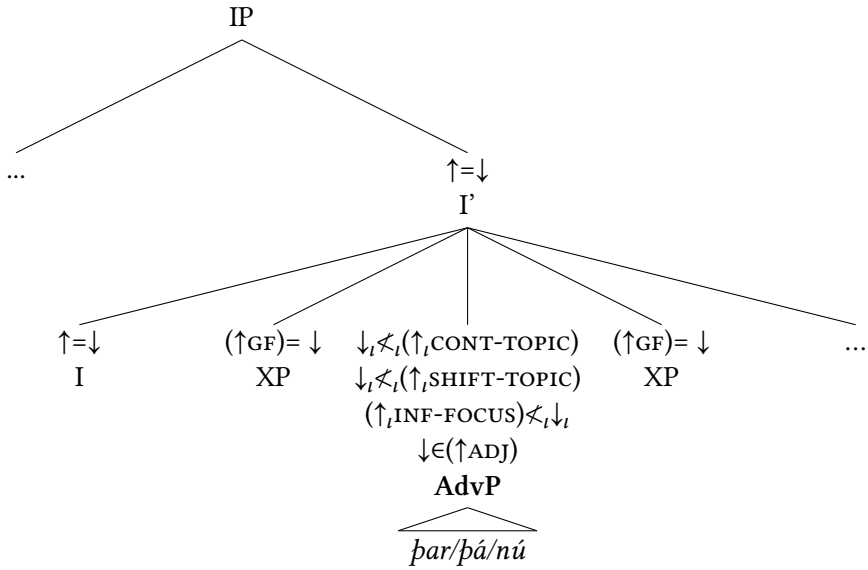
In order to model the mapping between syntax and information structure in the Old Icelandic postfinite domain, and the specific watershed function of clause-internal *þar/þá/nú*, I suggest a potential i-structure parallel to f-precedence, which I call i-precedence. i-precedence essentially allows one to constrain the ordering of c-structure constituents in terms of the i-structures they map onto, and can be defined as in (43) (cf. the definition for f-precedence in (41) above).

(43) i-precedence
ι_1 i-precedes ι_2 if and only if all c-structure nodes corresponding to ι_1 precede all c-structure notes corresponding to ι_2

I specify three i-precedence constraints on the position which hosts *þar/þá/nú*, see (44). The first i-precedence constraint – $\downarrow_\iota \not\prec_\iota (\uparrow_\iota \text{CONT-TOPIC})$ – states that the i-structure associated with the daughter node (i.e. associated with *þar/þá/nú*) does not i-precede the mother node's CONT(inuing)-TOPIC at i-structure, i.e. the clause's CONT(inuing)-TOPIC. This prevents continuing topics from occurring in the postfinite domain to the right of *þar/þá/nú*, as is the generalisation borne out from the (M)IcePaHC data, see (37) above. Likewise, the second i-precedence constraint – $\downarrow_\iota \not\prec_\iota (\uparrow_\iota \text{SHIFT-TOPIC})$ – prevents shift topics from occurring to the right of *þar/þá/nú*, another generalisation from the (M)IcePaHC data, see again (37). The third i-precedence constraint on the clause-internal adverb position – $(\uparrow_\iota \text{INF-FOCUS}) \not\prec_\iota \downarrow_\iota$ – states that the mother node's INF(ormation)-FOCUS at i-structure, i.e. the INF(ormation)-FOCUS of the overall clause, cannot i-precede the daughter node (i.e. the position of *þar/þá/nú*). This thus prevents information-foci from occurring in the postfinite domain to the left of *þar/þá/nú*, as indicated by the (M)IcePaHC data, see (37) above.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(44)



i-precedence, as a parallel to 's f-precedence, thus allows one to model the subtle information-structural watershed function of *þar/þá/nú*, which is more nuanced than standard discourse configurationality. The three i-structure constraints specify ordering restrictions specifically for continuing topics, shift topics and information foci relative to *þar/þá/nú*, but leave the position of other types of topic (e.g. /contrastive topics), and constituents which are neither topics nor foci (e.g. dummy pronouns and impersonal pronouns) underspecified with respect to *þar/þá/nú*, as borne out in the (M)IcePaHC data, see (37). In addition, these three i-precedence constraints are all relative to the I' node in terms of their domain of application (cf. Zaenen & Kaplan 1995, "relativized f-precedence"). The constraints thus specify orderings in the postfinite domain, but crucially still allow for various types of topic or foci to appear in the clause-initial prefinite position (prefield), as is attested in Old Icelandic (cf. the findings in Booth 2021).

5 Diachronic developments

Although clause-internal *þar/þá/nú* have been acknowledged to be a dominant feature of Old Icelandic in the previous literature (e.g. Faarlund 1990: 71), their status in later stages of Icelandic has not been addressed in specific detail. For Modern Icelandic, Thráinsson (2007: 38–39) states that place and time adverbs, including *þar/þá/nú*, typically occur either in the prefield or in a position after

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

the verb phrase, but cannot occur in a clause-internal position immediately after the finite verb, providing the data in (45).

(45) Modern Icelandic:

a. ***Þar/þá/nú*** *hefur Jón lokið þessu.*
there/then/now has Jón.NOM finished this.DAT
'There/then/now Jón has finished this.'

b. *Jón hefur lokið þessu* ***þar/þá/nú****.*
Jón.NOM has finished this.DAT there/then/now
'Jón has finished this there/then/now.'

c. * *Jón hefur* ***þar/þá/nú*** *lokið þessu.*
Jón.NOM has there/then/now finished this.DAT
Intended: 'Jón has there/then/now finished this.'

(Thráinsson 2007: 39)

However, Thráinsson (2007: 40) also states that, when used as a "discourse particle", presumably bleached of its temporal semantics, *nú* typically occurs clause-internally and cannot occur in the prefield, providing the data in (46). Thráinsson (2007) does not however discuss the status of *þar* or *þá* in this context.

(46) Modern Icelandic:

a. *Jón hefur* ***nú*** *lokið þessu.*
Jón.NOM has PTCL finished this.DAT
'Well, Jón has finished this.'

b. * ***Nú*** *hefur Jón lokið þessu.*
PTCL has Jón.NOM finished this.DAT
Intended: 'Well, Jón has finished this.' (Thráinsson 2007: 40)

At the same time, the dedicated study of *nú* as a particle in Modern Icelandic spontaneous conversation by Hilmisdóttir (2010) presents a rather different picture. Hilmisdóttir (2010: 276–277) observes that, while *nú* as a particle can occur clause-internally (see (46)), it in fact shows a very strong tendency to occur in the prefield (93% of all instances, compared to just 6% clause-internally), apparently going against the data from Thráinsson (2007: 40).

When one examines the full diachronic sweep of the data in IcePaHC, which spans 1150–2008 (all text types), it is clear that the relative frequency of clause-internal *þar/þá/nú* decreases over the course of the history of Icelandic, though it remains a possibility in the latest period (1901–2008), see Table 6. Note that the

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

period 1751—1900 contains the text *Hellismanna saga*, which alone provides 153 examples with clause-internal *þar/þá/nú*. This text is a "newly written" 19th century saga which aims to reproduce an earlier saga style and is a known outlier in this subsection of the corpus with respect to syntactic properties (cf. Schätzle & Booth 2019). The data for the period 1551–1749 also comes with an important caveat, since this subsection of IcePaHC contains a lot of translation texts and for the most part text types other than narratives, which dominate in the other periods (cf. Booth 2018: chapter 1). These caveats aside, the key finding in Table 6 is that, by the latest period (1901–2008), the frequency of clause-internal *þar/þá/nú* has dramatically declined compared to the previous centuries.

Table 6: Frequency of clause-internal *þar/þá/nú* across all matrix declaratives in IcePaHC (all text types, 1150–2008)

	clause-internal *þar/þá/nú*		all matrix declaratives
	n	%	
1150–1350	1,385	10.5%	13,321
1351–1550	1,272	11.4%	11,205
1551–1750	561	6.5%	8,590
1751–1900	747	8.6%	8,645
1901–2008	**249**	**2.9%**	8522

Taking a narrower view, Booth (2018: 260) observed the high frequency of clause-internal *þar/þá/nú* specifically in V1 presentational constructions in early stages of Icelandic, citing the examples in (47) (cf. also the corpus study in Booth & Schätzle 2019).

(47) Early Icelandic (Booth 2018: 260):

a. *Voru* ***þar*** *tvö skip í búnaði.*
were ÞAR two.NOM ships.NOM in preparation.DAT
'There were two ships in the preparations.'
[IcePaHC: 1250, Sturlunga.408.710]

b. *Verður* ***nú*** *mannfall ógurligt.*
becomes NÚ man.loss.NOM terrible.NOM
'There comes to be a terrible loss of men.'
[IcePaHC: 1480, Jarlmann.381]

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

c. *Kom* ***þá*** *veður á móti þeim.*
came ÞÁ wind.NOM towards they.DAT
'There came wind towards them.'
(Supplementary manual data: [1250, Eirik.9.2])

presentational constructions are a well known example of a clause which lacks a topic-comment articulation ("thetic", cf. Sasse 1987). In such clauses, which lack a topic, the clause-internal adverb occurs immediately after the finite verb, and before the subject which is in information focus. In this sense, *þar/þá/nú* serve to "close off" the topic domain in such contexts, and thus can be viewed as markers of a topicless clause.

Strikingly, this very function of signalling a topicless construction in presentational constructions has also been attributed to the prefield expletive *það* (e.g. Rögnvaldsson & Thráinsson 1990: 29; Sells 2005; Sigurðsson 2007: 145), which represents a later development in the language (Hróarsdóttir 1998, Rögnvaldsson 2002, Booth 2018, 2019, 2020), e.g. (48).

(48) a. 19th Century Icelandic:
Það *rísu upp tveir nýir kaupmenn.*
EXPL stoop up two.NOM new.NOM merchants.NOM
'There stood up two new merchants.' [IcePaHC: 1888, Grimur.126]

b. Modern Icelandic:
Það *komu nokkrir vopnaðir menn af næstu bæjum…*
EXPL came some.NOM armed.NOM men.NOM from next.DAT farms.DAT
'There came some armed men from the nearby farms…'
[IcePaHC: 2008, Ofsi.634]

Another investigation of the full diachronic sweep of IcePaHC reveals that, in presentational constructions, clause-internal *þar/þá/nú* are virtually in complementary distribution with the prefield expletive *það*, see Table 7. Once the expletive becomes established in presentationals, clause-internal *þar/þá/nú* are scarcely attested (n=3). I interpret these results to suggest that, in presentationals, the role of marking a thetic, topicless construction shifts diachronically from clause-internal *þar/þá/nú* to clause-initial (prefield) *það*.

This diachronic finding is further supported by Modern Icelandic data provided and discussed by Rögnvaldsson (1982: 78–79), shown here in (49). For presentationals, a V2 clause with expletive *það* in the prefield is possible (49a), as is a V1 clause with clause-internal *þá* (49b), provided other pragmatic conditions for V1 declaratives are fulfilled. However, a V1 clause without the clause-internal adverb is ruled out (49c), indicating that either a clause-internal adverb

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Table 7: Distribution of presentational constructions over three configurations in IcePaHC (all text types, 1150–2008)

	V-*þá*/*þar*/*nú*-SUBJ		EXPL-V-SUBJ		EXPL-V-*þá*/*þar*/*nú*-SUBJ	
	n	%	*n*	%	*n*	%
1150–1350	22	100.0%	0	0.0%	0	0.0%
1351–1550	22	81.5%	5	18.5%	0	0.0%
1551–1749	24	85.7%	4	14.3%	0	0.0%
1750–1900	17	32.1%	35	66.0%	1	1.9%
1901–2008	1	1.1%	86	96.6%	2	2.2%
all	86	39.3%	130	59.4%	**3**	**1.4%**

(e.g. *þar/þá/nú*) or the prefield expletive is required to render a thetic clause well-formed.

(49) Modern Icelandic:

a. ***Það*** *kóm einhver maður til mín.*
EXPL came some.NOM man.NOM to I.GEN
'Some man came to me.'

b. *Kóm* ***þá*** *einhver maður til mín.*
came ÞÁ some.NOM man.NOM to I.GEN
'Some man came to me.'

c. * *Kóm einhver maður til mín.*
came some.NOM man.NOM to I.GEN
Intended: 'Some man came to me.' (Rögnvaldsson 1982: 79)

I suggest that the rise of the prefield expletive *það* in presentationals and attendant decrease in V1 presentationals with clause-internal *þar/þá/nú*, as indicated in Table 7, reflects a broader change whereby Icelandic gradually shifts its structural specification of discourse functions from the postfinite domain to the prefield over time. Indeed, the fact that clause-internal *þar/þá/nú* has decreased in frequency in matrix declaratives in general, beyond specifically presentationals (see Table 6) indicates a wider change in the mapping between clause structure and information structure, likely connected with recent observations that topics are becoming increasingly firmly associated with the prefield in Icelandic (see previous work by Booth et al. 2017, Booth & Beck 2021).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

6 Conclusion

In this paper I have shown that, in Old Icelandic saga narratives, clause-internal adverbs (specifically *þar/þá/nú*) act as an information-structural watershed in the postfinite domain, contributing to the structural encoding of different types of information-structure roles in a highly specialised way. This phenomenon is in line with other Germanic varieties (e.g. Old English, Modern German), where similar effects have been observed in the postfinite domain, and suggests that the role of adverbs in the structural management of information structure is a relatively common phenomenon, at least within Germanic, and one which deserves more attention both empirically and theoretically. The need for further crosslinguistic research in this area is especially strong given the findings for Old Icelandic presented here, which indicate a complex system whereby the watershed adverbs are sensitive to different subtypes of topics and foci, and at the same time irrelevant for content which is neither topical or in focus. I argued that this nuanced and complex system calls for an alternative model other than straightforward discourse configurationality with designated topic and focus positions, and showed that the LFG architecture and in particular the notion of i-precedence can offer such an alternative. More broadly, the paper serves as a showcase for the rich empirical and theoretical insights which can be gained from the investigation of the morphosyntax-discourse interface in historical narrative texts, and in particular Old Icelandic saga narratives, which offer ample opportunities for future research in this area.

Abbreviations

ACC	accusative
ADJ	adjunct
CONT-TOPIC	continuing topic
DAT	dative
DEF	definite
EXPL	expletive
GEN	genitive
GF	grammatical function
INF-FOCUS	information focus
NEG	negation
NGF	nominal grammatical function
NOM	nominative
PTCL	particle
REL	relativiser
SBJV	subjunctive
SUBJ	subject
V	verb
V1	verb-initial
V2	verb-second

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Acknowledgements

References

Asudeh, Ash. 2006. Direct compositionality and the architecture of LFG. In Miriam Butt, Mary Dalrymple & Tracy Holloway King (eds.), *Intelligent linguistic architectures: Variations on themes by Ronald M. Kaplan*, 363–387. Stanford, CA: CSLI Publications.

Axel, Katrin. 2007. *Studies on Old High German syntax: Left sentence periphery, verb placement and verb-second*. Amsterdam: Benjamins.

Axel, Katrin. 2009. The verb-second property in Old High German: Different ways of filling the prefield. In Roland Hinterhölzl & Svetlana Petrova (eds.), *Information structure and language change: New approaches to word order variation in Germanic*, 17–44. Berlin: De Gruyter. DOI: 10.1515/9783110216110.1.17.

Barðdal, Jóhanna & Thórhallur Eythórsson. 2003. The change that never happened: The story of oblique subjects. *Journal of Linguistics* 39(3). 439–472. DOI: 10.1017/S002222670300207X.

Barðdal, Jóhanna & Thórhallur Eythórsson. 2009. The origin of the oblique subject construction: An Indo-European comparison. In Vit Bubenik, John Hewson & Sarah Rose (eds.), *Grammatical change in Indo-European languages*, 179–193. Amsterdam: Benjamins. DOI: 10.1075/cilt.305.

Barðdal, Jóhanna, Thomas Smitherman, Valgerður Bjarnadóttir, Serena Danesi, Gard B. Jenset & Barbara McGillivray. 2012. Reconstructing constructional semantics: The dative subject construction in Old Norse-Icelandic, Latin, Ancient Greek, Old Russian and Old Lithuanian. *Studies in Language* 36(3). 511–547. DOI: 10.1075/sl.36.3.03bar.

Berman, Judith, Stefanie Dipper, Christian Fortmann & Jonas Kuhn. 1998. Argument clauses and correlative *es* in German – deriving discourse properties in a unification analysis. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '98 Conference, University of Queensland, Brisbane*. Stanford, CA: CSLI Publications.

Booth, Hannah. 2018. *Clause structure and expletives: Syntactic change in Icelandic*. Manchester: University of Manchester. (Doctoral dissertation).

Booth, Hannah. 2019. Cataphora, expletives and impersonal constructions in the history of Icelandic. *Nordic Journal of Linguistics* 42(2). 139–164. DOI: 10.1017/S0332586519000143.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Booth, Hannah. 2020. Expletives in Icelandic: A corpus study. In Bridget Drinka (ed.), *Historical Linguistics 2017: Selected papers from the 23rd International Conference on Historical Linguistics, San Antonio, Texas, 31 July–4 August 2017*, 363–384. Amsterdam: Benjamins. DOI: 10.1075/cilt.350.17boo.

Booth, Hannah. 2021. Revisiting the configurationality issue in Old Icelandic. *Glossa: A journal of general linguistics* 6(1). 130. DOI: 10.16995/glossa.5804.

Booth, Hannah & Christin Beck. 2021. Verb-second and verb-first in the history of Icelandic. *Journal of Historical Syntax* 5(28). 1–53. DOI: 10.18148/hs/2021.v5i28.112.

Booth, Hannah & Christin Schätzle. 2019. The syntactic encoding of information structure in the history of Icelandic. In Miriam Butt, Tracy Holloway King & Ida Toivonen (eds.), *Proceedings of the lfg'19 conference, australian national university*, 69–89. Stanford, CA: CSLI Publications.

Booth, Hannah, Christin Schätzle, Kersti Börjars & Miriam Butt. 2017. Dative subjects and the rise of positional licensing in Icelandic. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG'17 Conference, University of Konstanz*, 104–124. Stanford, CA: CSLI Publications.

Börjars, Kersti, Pauline Harries & Nigel Vincent. 2016. Growing syntax: The development of a DP in North Germanic. *Language* 91(1). e1–e37. DOI: 10.1353/lan.2016.0002.

Bresnan, Joan. 1984. *Bound anaphora on functional structures*. Presented at the *Tenth Annual Meeting of the Berkeley Linguistics Society*.

Bresnan, Joan, Ash Asudeh, Ida Toivonen & Stephen Wechsler. 2016. *Lexical-functional syntax*. 2nd edn. Oxford: Blackwell. DOI: 10.1002/9781119105664.

Butt, Miriam, Farhat Jabeen & Tina Bögel. 2016. Verb cluster internal wh-phrases in Urdu: Prosody, syntax and semantics/pragmatics. *Linguistic Analysis* 40(3–4). 445–487.

Butt, Miriam & Tracy Holloway King. 1997. Null elements in discourse structure. In Karumuri V. Subbarao (ed.), *Papers from the NULLS seminar*, 1–23. Delhi: Motilal Banarsidass.

Carroll, Mary & Monique Lambert. 2003. Information structure in narratives and the role of grammaticised knowledge: A study of adult French and German learners of English. In Christine Dimroth & Marianne Starren (eds.), *Information structure and the dynamics of language acquisition*, 267–287. Amsterdam: Benjamins. DOI: 10.1075/sibil.26.13car.

Catasso, Nicholas, Marco Coniglio, Chiara De Bastiani & Eric Fuß. 2021. *He then said...*: (Understudied) deviations from V2 in early Germanic. *Journal of Historical Syntax* 5(16-25). 1–39. DOI: 10.18148/hs/2021.v5i17-26.55.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Cichosz, Anna. 2022. Old English V-initial and *þa*-VS main clauses: Independent constructions or allostructions? *Constructions and Frames* 14(2). 301–336. DOI: 10.1075/cf.21002.cic.

Clover, Carol J. 1974. Scene in saga composition. *Arkiv för nordisk filologi* 89. 57–83. DOI: 10.7591/9781501740510.

Dalrymple, Mary, Ronald M. Kaplan & Tracy Holloway King. 2001. Weak crossover and the absence of traces. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG'01 Conference*, 66–82. Stanford, CA: CSLI Publications.

Dalrymple, Mary, John J. Lowe & Louise Mycock. 2019. *The Oxford reference guide to lexical functional grammar*. Oxford: Oxford University Press. DOI: 10.1093/oso/9780198733300.001.0001.

Daneš, Frantisek. 1974. Functional sentence perspective and the organization of the text. In Frantisek Daneš (ed.), *Papers on functional sentence perspective*, 106–128. The Hague: Mouton. DOI: 10.1515/9783111676524.106.

Dimroth, Christine, Cecilia Andorno, Sandra Benazzo & Josje Verhagen. 2010. Given claims about new topics: How Romance and Germanic speakers link changed and maintained information in narrative discourse. *Journal of Pragmatics* 42(12). 3328–3344. DOI: 10.1016/j.pragma.2010.05.009.

Donhauser, Karin & Svetlana Petrova. 2009. Die Rolle des Adverbs *tho* bei der Generalisierung von Verbzweit im Deutschen. In Monika Dannerer, Peter Mauser, Hannes Scheutz & Andreas E. Weiss (eds.), *Gesprochen—geschrieben—gedichtet: Variation und Transformation von Sprache*, 11–24. Berlin: Erich Schmidt.

Egerland, Verner. 2013. Fronting, background, focus: A comparative study of Sardinian and Icelandic. *Lingua* 136. 63–76. DOI: 10.1016/j.lingua.2013.07.011.

Enkvist, Nils Erik. 1986. More about the textual functions of the Old English adverbial *þa*. In Dieter Kastovsky & Aleksander Szwedek (eds.), *Linguistics across historical and geographical boundaries: In honour of Jacek Fisiak on the occasion of his fiftieth birthday*, vol. 1, 301–310. Berlin: De Gruyter. DOI: 10.1515/9783110856132.301.

Enkvist, Nils Erik & Brita Wårvik. 1987. Old english *tha*, temporal chains and narrative structure. In Anna Giacalone Ramat, Onofrio Carruba & Giuliano Bernin (eds.), *Papers from the 7th International Conference on Historical Linguistics*, 221–237. Amsterdam: Benjamins. DOI: 10.1075/cilt.48.17enk.

Faarlund, Jan Terje. 1990. *Syntactic change: Toward a theory of historical syntax*, vol. 50 (Trends in Linguistics. Studies and Monographs [TiLSM]). Berlin: De Gruyter. DOI: 10.1515/9783110854947.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Foster, Robert. 1975. The use of *Þa* in Old and Middle English narratives. *Neuphilologische Mitteilungen* 76(3). 404–414.

Frascarelli, Mara & Roland Hinterhölzl. 2007. Types of topics in German and Italian. In Susanne Winkler & Kerstin Schwabe (eds.), *On information structure, meaning and form*, 87–116. Amsterdam: Benjamins. DOI: 10.1075/la.100.

Frey, Werner. 2004. A medial topic position for German. *Linguistische Berichte* 198. 153–190.

Fuß, Eric & Roland Hinterhölzl. 2023. On the historical development of pronouns referring to situations: The rise of pre-finite 'expletives' in German. *Journal of Historical Syntax* 7(2). 1–54. DOI: 10.18148/hs/2023.v7i2.162.

Götze, Michael, Thomas Weskott, Cornelia Endriss, Ines Fiedler, Stefan Hinterwimmer, Svetlana Petrova, Anne Schwarz, Stavros Skopeteas & Ruben Stoel. 2007. Information structure. In Steffi Dipper, Michael Götze & Stavros Skopeteas (eds.), *Information structure in cross-linguistic corpora: Annotation guidelines for phonology, morphology, syntax, semantics, and information structure*, 147–187. Potsdam: Universitätsverlag Potsdam.

Grosz, Barbara J., Scott Weinstein & Aravind K. Joshi. 1995. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics* 21. 203–225.

Grosz, Patrick G. 2016. Information structure and discourse particles. In Caroline Féry & Shinichiro Ishihara (eds.), *The Oxford handbook of information structure*, 336–358. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199642670.013.36.

Hendriks, Herman. 1996. Information packaging: From cards to boxes. In Teresa Galloway & Justin Spence (eds.), *Proceedings of the 6th Semantics and Linguistic Theory Conference (SALT VI)*, vol. 6, 75–92. Ithaca, NY: Cornell University. DOI: 10.3765/salt.v6i0.2765.

Hilmisdóttir, Helga. 2010. The present moment as an interactional resource: The case of *nú* and *núna* in Icelandic conversation. *Nordic Journal of Linguistics* 33(3). 269–298. DOI: 10.1017/S0332586510000211.

Hinterhölzl, Roland & Svetlana Petrova. 2010. From V1 to V2 in West Germanic. *Lingua* 120(2). 315–328. DOI: 10.1016/j.lingua.2008.10.007.

Holmberg, Anders. 2000. Scandinavian stylistic fronting: How any category can become an expletive. *Linguistic Inquiry* 31(3). 445–483. DOI: 10.1162/002438900554406.

Hooper, Joan B. & Sandra A. Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4(4). 465–497.

Hrafnbjargarson, Gunnar Hrafn. 2004. Stylistic fronting. *Studia Linguistica* 58(2). 88–134. DOI: 10.1111/j.0039-3193.2004.00111.x.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Hróarsdóttir, Thorbjörg. 2009. OV languages: expressions of cues. In Roland Hinterhölzl & Svetlana Petrova (eds.), *Information structure and language change: New approaches to word order variation in Germanic*, 67–90. Berlin: De Gruyter. DOI: 10.1515/9783110216110.1.67.

Hróarsdóttir, Þorbjörg. 1998. *Setningafræðilegar Breytingar á 19. Öld: þróun þriggja málbreytinga*. Reykjavík: Málvísindastofnun Háskóla Íslands.

Ingason, Anton Karl. 2020. Machine-Parsed IcePaHC, version 0.1. https://github.com/antonkarl/micepahc.

Jónsson, Jóhannes Gísli. 2002. S-adverbs in Icelandic and the feature theory of adverbs. *Leeds Working Papers in Linguistics and Phonetics* 9. 73–89.

Kaplan, Ron & Joan Bresnan. 1982. Lexical-functional grammar: A formal theory for grammatical representation. In Joan Bresnan (ed.), *The mental representation of grammatical relations*, 173–281. Cambridge, MA: MIT Press.

Kaplan, Ronald M. 1987. Three seductions of computational psycholinguistics. In P. Whitelock, M. M. Wood, H. L. Somers, R. Johnson & P. Bennett (eds.), *Linguistic theory and computer applications*, 149–188. London: Academic Press.

Kaplan, Ronald M. & John T. Maxwell. 1996. *Grammar writer's workbench*. Technical report, Xerox PARC.

Kaplan, Ronald M. & Annie Zaenen. 1989. Long-distance dependencies, constituent structure, and functional uncertainty. In Mark R. Baltin & Anthony S. Kroch (eds.), *Alternative conceptions of phrase structure*. Chicago, IL: University of Chicago Press.

van Kemenade, Ans & Bettelou Los. 2006. Discourse adverbs and clausal syntax in Old and Middle English. In Ans van Kemenade & Bettelou Los (eds.), *The handbook of the history of English*, 224–248. Oxford: Wiley Blackwell. DOI: 10.1002/9780470757048.ch10.

van Kemenade, Ans, Tanja Milicev & R. Harald Baayen. 2008. The balance between syntax and discourse in Old English. In Maurizio Gotti, Marina Dossena & Richard Dury (eds.), *English Historical Linguistics 2006: Selected papers from the fourteenth International Conference on English Historical Linguistics (ICEHL 14), Bergamo, 21–25 August 2006*, 3–21. Amsterdam: Benjamins. DOI: 10.1075/cilt.295.04kem.

van Kemenade, Ans. 2009. Discourse relations and word order change. In Roland Hinterhölzl & Svetlana Petrova (eds.), *Information structure and language change: New approaches to word order variation in Germanic*, 91–120. Berlin: De Gruyter. DOI: 10.1515/9783110216110.1.91.

van Kemenade, Ans & Tanja Milićev. 2011. Syntax and discourse in Old and Middle English word order. In Dianne Jonas, John Whitman & Andrew Garrett

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(eds.), *Grammatical change: Origins, nature, outcomes*, 239–255. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199582624.003.0012.

van Kemenade, Ans & Meta Links. 2020. Discourse particles in early English: Clause structure, pragmatics and discourse management. *Glossa: A journal of general linguistics* 5(1). DOI: 10.5334/gjgl.1020.

King, Tracy Holloway. 1997. Focus domains and information-structure. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG'97 Conference*. Stanford, CA: CSLI Publications.

Kiss, Katalin É (ed.). 1995. *Discourse configurational languages*. Oxford: Oxford University Press.

Klein, Jared S. 1994a. Gothic *þaruh, þanuh* and *-(u)h þan*. *Indogermanische Forschungen* 99. 253–276. DOI: 10.1515/9783110243406.253.

Klein, Wolfgang. 1994b. *Time in language*. London: Routledge. DOI: 10.4324/9781315003801.

Krifka, Manfred. 2007. Basic notions of information structure. In Caroline Féry & Manfred Krifka (eds.), *Interdisciplinary studies on information structure*, 13–56. Potsdam: Universitätsverlag Potsdam. DOI: 10.1556/aling.55.2008.3-4.2.

Krivonosov, Aleksej. 1977. Deutsche Modalpartikeln im System der unflektierten Wortklassen. In Harald Weydt (ed.), *Aspekte der Modalpartikeln: Studien zur deutschen Abtönung*, 176–216. Tübingen: Niemeyer.

Leiss, Elisabeth. 2000. *Artikel und Aspekt: Die grammatischen Muster von Definitheit*, vol. 55 (Studia Linguistica Germanica). Berlin: De Gruyter. DOI: 10.1515/9783110825961.

Leiss, Elisabeth. 2007. Covert patterns of definiteness/indefiniteness and aspectuality in Old Icelandic, Gothic, and Old High German. In Elisabeth Stark, Elisabeth Leiss & Werner Abraham (eds.), *Nominal determination: Typology, context constraints, and historical emergence*, 73–102. Amsterdam: Benjamins. DOI: 10.1075/slcs.89.

Levine, Robert & Detmar Meurers. 2006. Declarative models of syntax. In Keith Brown (ed.), *Encyclopedia of language and linguistics*, 2nd edn., vol. 3, 377–384. Oxford: Elsevier. DOI: 10.1016/B0-08-044854-2/01960-X.

Linde, Sonja. 2009. Aspects of word order and information structure in Old Saxon. In Roland Hinterhölzl & Svetlana Petrova (eds.), *Information structure and language change: New approaches to word order variation in Germanic*, 367–389. Berlin: De Gruyter. DOI: 10.1515/9783110216110.3.367.

Los, Bettelou. 2000. *Onginnan/beginnan* with bare and *to*-infinitive in Ælfric. In Olga Fischer, Anette Rosenbach & Dieter Stein (eds.), *Pathways of change: Grammaticalization in English*, 251–274. Amsterdam: Benjamins.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Mahowald, Kyle. 2011. *An LFG approach to Old English constituent order*. Oxford: University of Oxford. (MA thesis).

Maling, Joan. 1988. Variations on a theme: Existential sentences in Swedish and Icelandic. *McGill Working Papers in Linguistics: Special Issue on Comparative Germanic Syntax* 6(1). 168–191.

Maling, Joan. 1990. Inversion in embedded clauses in Modern Icelandic. In Joan Maling & Annie Zaenen (eds.), *Syntax and semantics: Modern Icelandic syntax*, 71–91. San Diego, CA: Academic Press. DOI: 10.1163/9789004373235_004.

Matić, Dejan, Rik Van Gijn & Robert D. Van Valin Jr. 2014. Information structure and reference tracking in complex sentences. In Rik van Gijn, Jeremy Hammond, Dejan Matić, Saskia van Putten & Ana Vilacy Galucio (eds.), *Information structure and reference tracking in complex sentences*, 1–42. Amsterdam: Benjamins. DOI: 10.1075/tsl.105.01mat.

Matsuda, Kenjirô. 1998. On the conservatism of embedded clauses. In Monika S. Schmid, Jennifer R. Austin & Dieter Stein (eds.), *Historical linguistics 1997: Selected papers from the 13th international conference on historical linguistics, düsseldorf, 10–17 august 1997*, 255–268. Amsterdam: Benjamins. DOI: 10.1075/cilt.164.18mat.

Ólason, Vésteinn. 2004. Family sagas. In Rory McTurk (ed.), *A companion to Old Norse-Icelandic literature and culture*, 101–118. Oxford: Wiley Blackwell. DOI: 10.1002/9780470996867.ch7.

Petrova, Svetlana. 2011. Modeling word order variation in discourse: On the pragmatic properties of VS order in Old High German. *Oslo Studies in Language* 3(3). 209–228. DOI: 10.5617/osla.46.

Phelpstead, Carl. 2017. Time. In Ármann Jakobsson & Sverrir Jakobsson (eds.), *The Routledge research companion to the medieval Icelandic sagas*, 187–197. London: Routledge.

Platzack, Christer. 1983. Existential sentences in English, German, Icelandic and Swedish. In Fred Karlsson (ed.), *Papers from the 7th scandinavian conference of linguistics*, 80–100. Helsinki: Department of General Linguistics, University of Helsinki.

Pullum, Geoffrey K. & Barbara C. Scholz. 2001. On the distinction between model-theoretic and generative-enumerative syntactic frameworks. In Philippe de Groote, Glyn Morrill & Christian Retoré (eds.), *Logical aspects of computational linguistics*, 17–43. Berlin, Heidelberg: Springer. DOI: 10.1007/3-540-48199-0_2.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.

Reinhart, Tanya. 1981. Pragmatics and linguistics: An analysis of sentence topics. *Philosophica* 27(1). 53–94. DOI: 10.21825/philosophica.82606.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Riester, Arndt. 2015. Analyzing questions under discussion and information structure in a Balinese narrative. In *Proceedings of the second international workshop on information structure of Austronesian languages*, 1–26. Tokyo: Research Institute for Languages, Cultures of Asia & Africa (ILCAA).

Rögnvaldsson, Eiríkur. 1982. *Um orðaröð og færslur í íslensku*. Reykjavík: University of Iceland. (MA thesis).

Rögnvaldsson, Eiríkur. 2002. ÞAÐ í fornu máli - og síðar. *Íslenskt mál* 24. 7–30.

Rögnvaldsson, Eiríkur & Höskuldur Thráinsson. 1990. On Icelandic word order once more. In Joan Maling & Annie Zaenen (eds.), *Syntax and semantics: Modern Icelandic syntax*, 3–40. San Diego, CA: Academic Press. DOI: 10.1163/9789004373235_002.

Sasse, Hans-Jürgen. 1987. The thetic/categorical distinction revisited. *Linguistics* 25(3). 511–80. DOI: 10.1515/ling.1987.25.3.511.

Schätzle, Christin. 2018. *Dative subjects: Historical change visualized*. Konstanz: University of Konstanz. (Doctoral dissertation).

Schätzle, Christin & Hannah Booth. 2019. DiaHClust: An iterative hierarchical clustering approach for identifying stages in language change. In *Proceedings of the 1st International Workshop on Computational Approaches to Historical Language Change*, 126–135. Florence, Italy: Association for Computational Linguistics.

Schätzle, Christin, Miriam Butt & Kristina Kotcheva. 2015. The diachrony of dative subjects and the middle in Icelandic: A corpus study. In Miriam Butt & Tracy Holoway King (eds.), *Proceedings of the LFG'15 conference, Tokyo, Japan, 18–20 July 2015*, 357–377. Stanford, CA: CSLI Publications.

Sells, Peter. 2001. *Structure, alignment and optimality in Swedish*. Stanford, CA: CSLI Publications.

Sells, Peter. 2005. The peripherality of the Icelandic expletive. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG'05 Conference*, 408–428. Stanford, CA: CSLI Publications.

Sells, Peter. 2021. The view from declarative syntax. In Nicholas Allott, Terje Lohndal & Georges Rey (eds.), *A companion to Chomsky*, 243–266. Oxford: Wiley Blackwell. DOI: 10.1002/9781119598732.ch16.

Sigurðsson, Halldór Ármann. 2007. Argument features, clausal structure and the computation. In Tammoy Bhattachayra, Eric Reuland & Giorgos Spathas (eds.), *Argument structure*, 121–158. Amsterdam: Benjamins. DOI: 10.1075/la.108.10sig.

Thráinsson, Höskuldur. 1979. *On complementation in Icelandic*. New York: Garland.

Thráinsson, Höskuldur. 2007. *The syntax of Icelandic*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511619441.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Trips, Carola & Eric Fuß. 2009. The syntax and semantics of the temporal anaphor *"then"* in Old and Middle English. In Artemis Alexiadou, Jorge Hankamer, Thomas McFadden, Justin Nuger & Florian Schäfer (eds.), *Advances in comparative Germanic syntax*, 171–196. Amsterdam: Benjamins. DOI: 10.1075/la.141.08the.

Vallduví, Enric. 1992. *The informational component*. New York: Garland Press.

Vilkuna, Maria. 1989. *Free word order in Finnish: Its syntax and discourse functions*. Helsinki: Suomalaisen Kirjallisuuden Seura [Finnish Literature Society].

Wallenberg, Joel C., Anton Karl Ingason, Einar Freyr Sigurðsson & Eiríkur Rögnvaldsson. 2011. *Icelandic Parsed Historical Corpus (IcePaHC) 0.9*. http://linguist.is/icelandic_treebank/Icelandic_Parsed_Historical_Corpus_(IcePaHC). Version 0.9. DOI: 10.21248/jlcl.26.2011.153.

Wårvik, Brita. 2013. Participant continuity and narrative structure: Defining discourse marker functions in Old English. *Folia Linguistica Historica* 47(34). 209–242. DOI: 10.1515/flih.2013.008.

Zaenen, Annie. Forthcoming. Information structure. In Mary Dalrymple (ed.), *Forthcoming: The handbook of lexical functional grammar* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Language Science Press.

Zaenen, Annie & Ronald M. Kaplan. 1995. Formal Devices for Linguistic Generalizations: West Germanic word order in LFG. In Mary Dalrymple, Ronald M. Kaplan & Annie Zaenen (eds.), *Formal issues in lexical-functional grammar*, 215–239. Stanford, CA: CSLI Publications.

Zaenen, Annie, Joan Maling & Höskuldur Thráinsson. 1985. Case and grammatical functions: The Icelandic passive. *Natural Language and Linguistic Theory* 3(4). 441–483. DOI: 10.1007/BF00133285.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Chapter 6

In consideration of Age and Genre: A diachronic corpus analysis of deverbal nouns in German verbonominal constructions and complex prepositions

Fabian Fleissner [a]
[a]Université de Neuchâtel

This paper deals with the formation and the diachronic development of German complex prepositions (CPs) following the general syntactic pattern [*in* N_{dev} P/GEN] and verbonominal constructions (VNCs) following the pattern [*in* N_{dev} V]. It sets out to answer two general questions: What are the discursive functions fulfilled by CPs and VNCs across different genres, and how does the age of their nominal core noun influence their genesis and persistence? It will be shown that both constructions exhibit different distributional and preferential patterns in these respects. [*in* N_{dev} P/GEN] shows clear genre-specific characteristics of 19th century scientific texts which are reflected in the emergence of a REFERENCE discourse schema, stimulated by the coinage of new nouns with this semantics in New High German. [*in* N_{dev} V] displays a broader, more open structure with diverse semantics and therefore shows a less genre-specific distribution. This construction primarily contains nouns from Early New High German, but repel those from New High German tailored to their functional niche.

1 Introduction

A prevailing observation in the contemporary history of German is its notable inclination towards a nominal style which is characterized by the frequent use

Fabian Fleissner. 2025. In consideration of Age and Genre: A diachronic corpus analysis of deverbal nouns in German verbonominal constructions and complex prepositions. In Ulrike Demske & Barthe Bloom (eds.), *Discourse structure and narration: A diachronic view from Germanic*, 173–210. Berlin: Language Science Press. DOI: ??

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

of nouns derived from verbs, adjectives, or other nouns.[1] This style is commonly found in academic writing, technical texts, and bureaucratic language (cf. Eggers 1973; Möslein 1981; Admoni 1973,1990; von Polenz 1987, von Polenz 2013). This paper sheds light on two constructions associated with this tendency and, in some cases, held responsible for it: verbonominal constructions (henceforth VNC) of the syntactic pattern [P (Det) (Mod) N_{dev} (P/GEN) V][2] as in (1)–(2) and complex prepositions (henceforth CP) of the syntactic pattern [P (Det) N_{dev} P/GEN] as in (3)–(4) (cf. e.g. Seppänen et al. 1994 and Stefanowitsch et al. 2020).

(1) *Es kommt aber noch ein zweiter Gegenſtand in Betrachtung.*
'However, a second issue **comes into consideration**.' (Clausewitz, 1834)[3]

(2) *Diese Hütte **stand** bis 1866 [...] **im Betrieb**.*
'This hut **was in operation** until 1866.' (Beck, 1899)

(3) *[...] der [wolle] ihn **in Betrachtung solchen Gehorsams** der gnädigen Rücksicht des Himmelssohnes empfehlen.*
'[...] He wishes to commend him, **in consideration of such obedience**, to the gracious care of the Son of Heaven.' (Berg, 1873)

(4) ***In Bezug auf den Monotheismus** der Indoeuropäer bemerke ich kurz Folgendes.*
'**With regard to the monotheism** of the Indo-Europeans, I briefly note the following.' (Chamberlain, 1899)

Both VNCs and CPs usually contain a deverbal noun and are assumed to be used in similar discursive contexts. Nevertheless, it has not yet been possible to trace functional similarities or even an interdependent development of both constructions. One obvious reason for this is that the usually assumed 'information condensation' of deverbal nouns only comes into play with CPs that potentially shorten subordinate clause structures. VNCs, on the other hand, are

[1]This study is part of the project *Zusammenspiel von Wortbildung und Syntax: Nominalisierungsstrategien in verbonominalen Konstruktionen und sekundären Präpositionen* 'The interaction of word formation and syntax: Nominalization strategies in verbonominal constructions and secondary prepositions' funded by the Swiss National Science Foundation (GNr. 197124).

[2]In research literature, constructions without prepositions are often considered part of the same class of VNCs. In this study, I exclusively discuss VNCs with prepositions to facilitate a comparison with complex prepositions.

[3]Throughout the paper, examples from the DTA-corpus are cited according to the following pattern: [author, year].

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

characterized by a complex predicative structure which, at first sight, does not offer any advantages over the use of verbs of the derivational base. Much attention has been paid to both constructions, especially from a diachronic and morphosyntactic perspective: Longstanding interest in the study of German VNCs has led to various insights in this respect (cf. e.g. Van Pottelberge 2001; Seifert 2004; Kamber 2008; Kabatnik 2020; Harm 2021; Fleischhauer & Hartmann 2021; Fleischhauer 2022). It is generally accepted that German *Funktionsverbgefüge* (lit. 'function verb constructions') are characterized by a 'desemanticized' verb and an eventive noun that constitutes the core of the predication. There is also a certain consensus that, from a semantic point of view, individual constructions display different degrees of compositionality and idiomaticity. Whereas older studies usually treat German VNCs as one homogenous category (von Polenz 1987; Helbig 1979; Kamber 2008), more recent studies focus on individual (families of) constructions and their specific properties (Zeschel 2008; Kabatnik 2020; Fleischhauer 2022; Smirnova & Stöber 2022). There has been a growing acceptance that they do not uniformly exhibit identical characteristics in terms of their morphosyntactic structure and function. However, both of these perspectives are yet to offer a comprehensive and unifying definition of the phenomenon itself or a clear distinction of its subclasses. The term *Funktionsverbgefüge* implies that these constructions are primarily described in relation to the status of the verb, which, in contrast to its 'heavy' use, is considered restricted or devoid of its 'full' semantic weight. Consequently, VNCs have been repeatedly examined and analyzed from the perspective of grammaticalization theory.

Complex prepositions in German are often described as an open class of constructions (cf. e.g. Di Meola 2000; Szczepaniak 2009; Stefanowitsch et al. 2020). Defining this class precisely has proven to be challenging due to the blurred boundaries between simple prepositions and free syntactic combinations (cf. e.g. Lehmann & Stolz 1992: 17; Meibauer 1995). The majority of CPs in present-day German are formed according to the pattern [P (Det) N P/GEN] as shown in (5).

(5) *in*$_P$ *Bezug*$_N$ *auf*$_P$, *in/im*$_P$ *Zusammenhang*$_N$ *mit*$_P$, *im*$_P$ *Vergleich*$_N$ *zu*$_P$
'in relation to, in connection with, in comparison with'

The principal characteristics of CPs of the syntactic pattern [P (Det) N P/GEN] in German have been researched in numerous ways (for the complete list of criteria see Beneš 1974: 34–35; Lehmann & Stolz 1992; Lindqvist 1994; Di Meola 2000; Szczepaniak 2009; Stefanowitsch et al. 2020; Ruf & Smirnova (forthcoming)). In CPs, the first preposition is usually assumed to be invariable, and the second, postnominal preposition or the genitive case of the noun phrase is also widely

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

taken as being fixed. Regarding the nominal core [(Det) N] of the pattern, as highlighted in the referenced literature, it is generally observed that the noun cannot undergo syntactic modifications such as adjectives and determiners, or morphological changes like plural marking. Generally, the characteristic properties of CPs are associated with the formal fixedness of these multi-word expressions. From a historical perspective, this formal rigidity is often interpreted as an outcome of gradual conventionalization, wherein regular syntactic combinations evolve into fixed expressions over time. This process involves the reanalysis of postnominal attributive phrases as complements of a new CP, as outlined by Lehmann (1998). Throughout this diachronic evolution, the constituent elements of the previously freely combined syntactic structure progressively lose their autonomy, eventually merging into a unified and distinctive structure. But none of the attributes examined in the literature can be considered a defining feature of a CP, as individual instances typically exhibit a combination of these properties, sometimes to varying degrees. According to an alternative hypothesis, CPs are formed directly after a general schema by filling its open slots with lexical material (see Hoffmann 2005). The mechanism at hand in the traditional scenario, along with reanalysis, is *routinization* and conventionalization of a syntagmatic string that results from the frequent and repeated use in the linguistic community. However, these characteristics remain crucial for delineating the category of CPs from syntactic combinations formed in a regular manner.

Apart from these individual morphosyntactic observations of particular patterns, German nominal style has been associated with the development of modern scientific language in discourse-pragmatic approaches. These approaches generally assume a shift of information from the verbal to the nominal domain in recent scientific texts of German (see Heinz & Weinrich 1995: 31). Whereas in narrative texts verbs with great semantic intensity play a major role as action carriers, the shift of information from the verbal to the nominal domain in non-narrative texts leads not only to a higher frequency of nouns in general, but also to a higher frequency of semantically 'weak' verbs such as auxiliary verbs, copulative verbs and VNCs. Modern scientific linguistic style, like modern scientific methodology, has its roots in the 17th century. The emancipation of the empirical natural sciences, once despised as *Realwissenschaften*, from an outsider position within the scholastic understanding of science to their modern role as leading sciences was promoted during this period, especially in England. This century also saw the beginnings of the transition from Latin to German in academic writing. Scholars like Gottfried Wilhelm Leibniz, Christian Thomasius and Christian Wolff are considered to be the forerunners of this protracted and laborious transition to German as the language of science. Leibniz is the one

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

who points to a language crisis throughout his writings. However, he does not see the improvement of the state of language in poetry, as the language societies of that time did, but in the development of *Sachprosaschriften* (cf. Pörksen 1998: 90). These formed the basis for the German language of science. One of the essential factors of German history in the absolutist era was the extraordinary extent of foreign language use, especially French. While Latin-German bilingualism existed in the state system, in the field of science the transition to German was delayed for a long time due to the internationally renowned French or the adherence to Latin. While in Italy the transition from Latin to the 'vernacular' as a language of science was already taking place gradually in the 16th century, and in France and England in the 17th century, academic writers in Germany remained predominantly with Latin until the late 18th century (cf. von Polenz 2013: 53-58). It is precisely during this period that the most quantitatively striking diachronic developments in CPs and VNCs occur, as our data show. The following two Figures illustrate this with the development curves of CPs Figure 1 and VNCs Figure 2 with certain selected individual constructions.

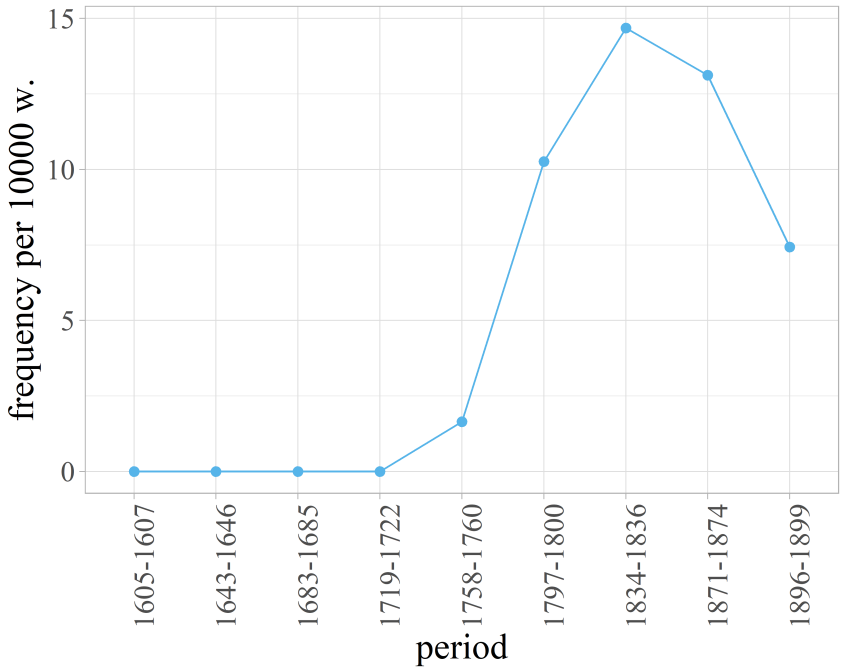


Figure 1: CPs of the pattern [in N P/GEN]; N = *Hilfe, Beihilfe, Bezug, Beziehung, Hinsicht, Rücksicht*

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

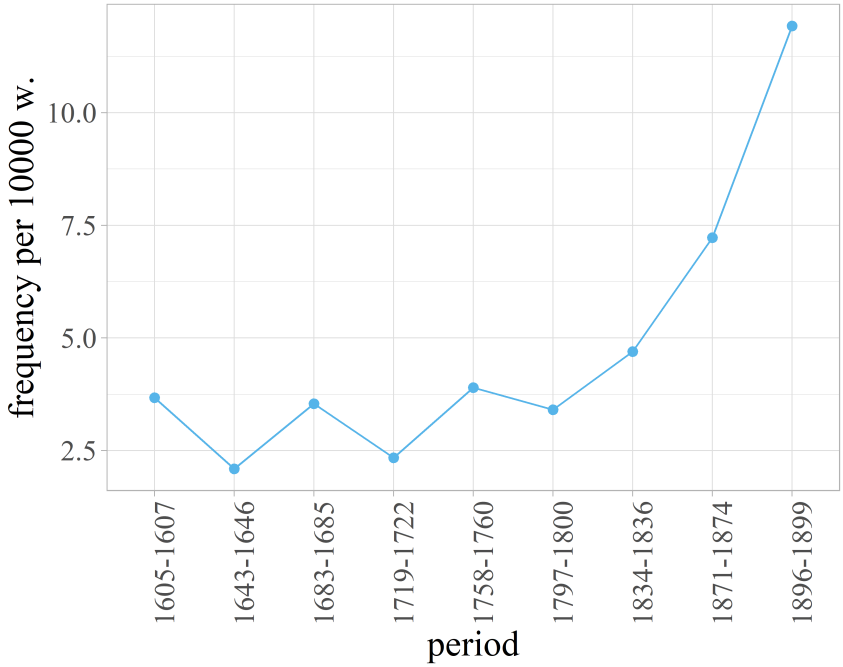


Figure 2: VNCs of the pattern [*in N kommen*]

In the present study, I will show that the developmental history of the two structures under investigation is related to the extralinguistic developments addressed. I argue that both the rise of nominalizations in German and the rise of CPs and VNCs can be explained discourse-pragmatically. The constructions feed on nominal material available at diachronically different points in time. The gateways are assumed to be certain texts that make corresponding stylistic and discursive demands, which can be fulfilled through the constructions under investigation.

The paper is structured as follows. In §2, I will introduce my research objects and theoretical background. §3 presents data and methods. §4 presents the results of the corpus analysis which are discussed in §5. §6 summarizes and consolidates the findings of the study and provides an outlook on future points of connection.

2 Research objects and theoretical background

The present study investigates multi-word expressions which are formed following the pattern [*in* N _]. This pattern gave rise to different constructions at vari-

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

ous points in the history of German, with CPs of the pattern [*in* N_{dev} P/GEN] and VNCs of the pattern [*in* N_{dev} V] being the most striking constructions. In contrast to previous approaches, I will not examine the development of individual morphosyntactic patterns to examination. Instead, I will approach it functionally from semantic clusters, which need to be identified for the entire construction. The most promising starting point in this case is the nominal core of the structures, i.e. deverbal nouns. The study starts out with two simple questions: Why are some nominalizations more prone to being used in the relevant constructions than others? And how do these preferences change over time? I will focus on 2 parameters:

(i) AGE. Following the tendency for nominalizations to become "nounier" over time, younger and "verbier" nominalizations are believed to be more often used in both types of constructions. Although the data show that this is true in principle, there is a non-negligible number of important exceptions that require explanation.

(ii) GENRE. The explosive spread of CPs and VNCs does not take place before the late 18th century, and it is assumed to initially occur in certain non-narrative texts, especially in scientific texts.

From the diachronic point of view, the category of CPs and LVCs in German are usually conceived of as an expanding open class. As mentioned in §1 above, it is generally assumed that both stem from regular syntactic combinations with compositional semantics. However, this perspective faces a challenge, as highlighted by Hoffmann (2005) with respect to English CPs, suggesting that for numerous CPs, there is insufficient empirical support for a transition from a freely combined structure to a fixed expression.

In English, numerous constructions of this nature emerge suddenly in written records, typically in the format [P (Det) N P/GEN]. This hints at the possibility that certain CPs might have arisen through analogy to pre-existing units that likely evolved previously in a manner akin to that outlined in (5). Hoffmann (2005) refers to this process as 'grammaticalization by analogy' and he suggests the presence of an 'abstract construct':

> [...] the sequence 'preposition–noun–preposition' appears to be available as a grammaticalized yet abstract construct which under certain circumstances can be filled by new lexical entities to form a new complex preposition. (Hoffmann 2005: 171)

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Hoffmann proposes what, in constructionist terms, could be termed as a schematic construction which is used to generate new constructions spontaneously by filling the variable slots [P] and [N] with lexical content. This appears to be the case for numerous CPs in English. In Ruf & Smirnova (forthcoming), a first sketch is been offered of how the category of CPs in German may be modeled in terms of a constructional network with several levels of abstraction and with different sub-schemas within the general schema [P (Det) N P] (see Figure 3).

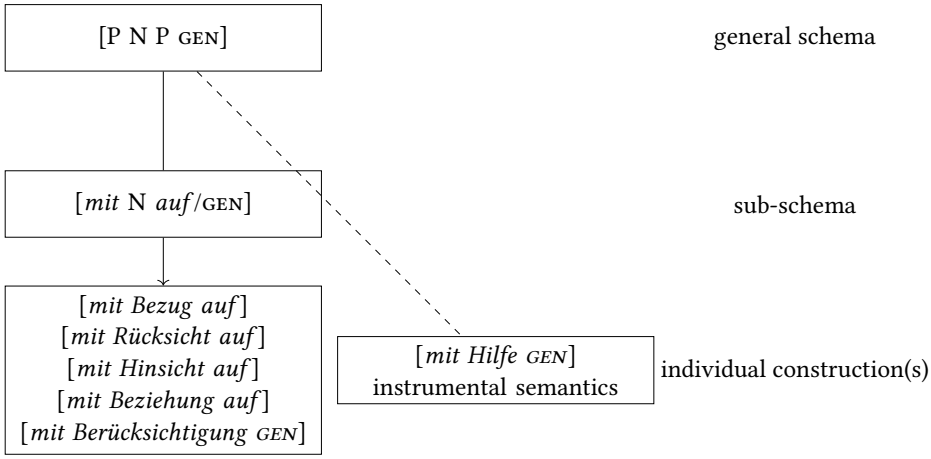


Figure 3: Partial network of German complex prepositions (Ruf & Smirnova forthcoming)

The examined data indicate that the structure [*mit* N_{dev} *auf*/GEN], involving nouns such as *Bezug* 'relation', *Berücksichtigung* 'consideration', *Rücksicht* 'consideration', *Hinsicht* 'regard', and *Beziehung* 'relation', qualifies as an abstract and productive constructional schema. This schema exhibits the capacity to attract various nouns sharing similar semantics of REFERENCE, and it maintains consistent across most instances. Of course, the class of complex prepositions in German is much more diverse and consists of many more members following different structural patterns, as for example [*in* N *mit*] as in *in Zusammenhang mit* 'in connection with' or [*im* N *von*/GEN] as in *im Zuge* 'in the course of'. In the case of [*mit Hilfe* GEN] 'with the help of', there is a conventionalized lexical item that emerged by a gradual process of routinization, whereby the internal structure became more and more fixed, and the meaning became less compositional and shifted towards the instrumental semantics. Some other synonymous nouns which, given their semantic content, had the potential to follow the analogical model of [*mit Hilfe* GEN], such as *Beistand* 'assist' or *Unterstützung* 'support',

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

did not succeed in this way. Whether the constructional network above applies to other complex prepositions, remains an open question. Unlike the preposition *mit*, the preposition *in* appears in both CPs and VNCs, and is therefore best suited to investigate any common developments between the two construction types.

3 Data

The present study focuses on the two aforementioned patterns with the preposition *in* 'in' [*in* N_{dev} P/GEN] and [*in* N_{dev} V], see e.g. (6) and (7) below, as this syntagmatic pattern is recurrent in both CPs and VNCs, while other prepositions like *mit* or *zu* usually occur in just one structure. The study is based on data from the DTA (Deutsches Textarchiv, version 2018). For the purposes of our research project, we selected a smaller data sample from the whole DTA that covers nine periods and amounts to about 10% of the entire corpus. Table 1 represents the composition of our data with respect to the individual periods.

Table 1: Investigated periods

period	years	corpus size
1	1605–1607	544,752
2	1643–1646	286,361
3	1683–1685	367,164
4	1719–1722	1,282,914
5	1758–1760	1,154,723
6	1797–1800	2,172,629
7	1834–1836	4,366,178
8	1871–1874	2,256,498
9	1896–1899	4,411,414
	Total	**16,842,633**

From this corpus, we created a data set containing all occurrences of deverbal nouns, based on a list that was compiled using the studies by Kamber (2008), Hartmann (2016) and Smirnova & Stöber (2022).[4] The data set from the last period is

[4]We restrict ourselves to deverbal nominalizations because of their high affinity to the structures under investigation. Though one can find CPs and VNCs with nouns of non-verbal origin, deverbal nouns are much more frequent in the structures [*in* N_{dev} P/GEN] and [*in* N_{dev} V].

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

not only the largest proportion of our corpus, but it also contains the most occurrences of CPs and VNCs.[5] Each individual nominal token was examined with regard to these syntactic structures. For this study, Period 9 was selected as the main research area because, in this segment, all diachronic developments concerning the parameters of AGE and GENRE are concluded and discernible in our data. A comprehensive overview across all time periods would obscure the focus on discursive functional characteristics of individual constructions, as written language patterns also undergo changes. Furthermore, lexical diversity is greatest at the end of the 19th century, as only a few nouns fell out of use during the observed periods. Table 2 provides an overview of the absolute numerical ratios of all deverbal nouns in Period 9.

Table 2: Overview of the token and type frequency of [in N_{dev} _]

	CP	VNC	ZERO
n (tokens)	2148	3553	68393
n (types)	191	231	346

The number of structures with a verb exceeds that of structures without significantly, which is not a surprising finding, since VNCs, with their additional slot, allow for higher syntactic variability and are more strongly integrated into the general argument structure of German as more or less ordinary predicate-forming units. As is immediately evident, the vast majority of nouns are used outside of the two structures (ZERO). At least, this further suggests the hypothesis that the increase in the use of nominal structures in German is not necessarily attributable to the increasing use of the two most conspicuous morphosyntactic patterns. Rather, it seems that both structures draw from a comprehensive lexical pool that can be utilized for specific discourse functions. A synchronic point at the end of the 19th century serves as a suitable culmination of all tendencies in preceding centuries. However, this does not alter the fact that the data must be interpreted in terms of their diachronic dynamics. Therefore, each noun was annotated based on the non-morphosyntactic parameter of AGE, as explained in §2. The assignment of AGE had to be done manually, since this parameter has not played a role in previous research and accordingly no corpus provides relevant information. The productivity of different word formation patterns at the same

[5]A complete list of the nouns as well as the data used for this study can be found at https://osf.io/yancd.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

time and numerous analogical formations complicated the access via the mere form of individual lexemes for diachronic sorting. For this reason, a conscious decision was made to omit this simpler, objectively measurable parameter of word formation type in favor of AGE. Various historical dictionaries and reference corpora served as the basis, occasionally requiring searches for reconstructed forms. For the majority of nouns, this method was sufficient to assign a specific age to an individual lexeme. The categorization of different stages was done according to traditional conventions of historical linguistics: Proto Germanic (PROTOG), Old High German (OHG), Middle High German (MHG), Early New High German (ENHG) and New High German (NHG). Table 3 provides an overview of the token frequency of each noun and the time of their emergence.

Table 3: Overview of the token frequency of [in N_{dev} _] by AGE

age	CP	VNC	ZERO	total
PROTOG	249	436	9653	10338
OHG	659	710	15159	16528
MHG	335	934	17832	19101
ENHG	490	1245	16421	18156
NHG	412	225	8664	9301
NA[a]	3	3	664	670

[a]These are mainly nominalized infinitives whose age are difficult to determine.

The mere comparison of absolute frequencies is not very informative, but it suggests a closer examination of the respective ratios and the underlying lexical types. Both analyzed structures not only behave differently in relation to each other but also each in relation to the lexical pool outside the specific constructions.

The underlying corpus is balanced in terms of the genre affiliation of its individual texts. As in the DTA, four different text genres are distinguished: *Belletristik* 'Fiction', *Zeitung* 'Newspaper', *Gebrauchsliteratur* 'Reference literature'[6], and

[6]'Reference literature' refers to publications designed to provide concise and specific information on a wide range of topics. Examples include dictionaries, encyclopedias, atlases, almanacs, handbooks, cookbooks, decency literature, travel literature and directories. Reference literature is structured for easy access to specific pieces of information. Readers typically consult reference materials for quick answers, definitions, or explanations rather than engaging in a continuous narrative. Reference literature occupies a position between more narrative genres such as fiction and newspapers on one side and clearly non-narrative scientific texts on the other.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Wissenschaft 'Science'.

Table 4: Overview of the token frequency of [in N_{dev} _] by GENRE

genre	CP	VNC	ZERO	total
fiction	390	697	4583	5670
newspaper	99	156	1227	1482
reference	400	784	4212	5396
science	1259	1916	11812	14987

The absolute distribution already suggests that, as anticipated, both constructions predominantly manifest in in scientific texts. But in comparison to the usage of the nominal core elements outside the respective constructions, it also becomes apparent that these are at best tendencies, and no construction-specific patterns are readily discernible at first glance. The analyses conducted in the following subchapter aim to explain the emergence of the distributional patterns outlined here.

4 Analysis

In this section, I will present and discuss the results of the corpus study based on the data outlined in §3. I will initially provide the findings for the parameter AGE (§4.1) and subsequently for GENRE (§4.2), before proceeding to the synthesis in §4.3. Initially, the focus is on conducting statistical tests for the respective sub data sets to elucidate the distributional patterns hinted at. For this purpose, Pearson Residuals[7] are calculated. These residuals help identify deviations between observed and expected values in the data, aiding in the interpretation of distributional trends. Once specific and statistically valid distribution patterns have been identified, conclusions can be drawn regarding the discourse-pragmatic potential of the analyzed structures based on these findings.

4.1 AGE – N_{dev}

The deployment of Pearson residuals facilitates the identification of cells where the observed data significantly diverge from expected outcomes, assuming inde-

[7]Pearson Residuals are included in the results obtained from the chisq.test() function in R (R Core Team 2023).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

pendence between the categorical variables. In essence, Pearson Residuals normalize a table's distribution, placing less emphasis on raw frequencies. The numerical values, presented in Table 5, serve as a measure for the relative importance of cells.

Table 5: Pearson Residuals, CXN by AGE

age	CP	VNC	ZERO
PROTOG	-3.05	-2.86	1.20
OHG	8.02	-3.15	-0.70
MHG	-9.44	0.34	1.60
ENHG	-1.75	12.39	-2.52
NHG	8.51	-10.60	0.91

A positive residual indicates that the observed value surpasses expectations (and vice versa for negative values). Residuals within the range of 0 to ±2 suggest that the difference between observed and expected is small enough to be interpreted as random variation. Examining Table 5 for CPs, OHG nouns and NHG nouns significantly exceed their expected occurrence, while the reverse holds true for MHG nouns and PROTOG nouns. For VNCs, different distributional patterns are observed. There is a strong tendency towards ENHG nouns, while all other periods, especially NHG, are disfavored. For ZERO, only ENHG nouns reach statistical significance, as they slightly but significantly fall below the expected value.

Employing residuals as a proportion-based method for data analysis is more advantageous than relying solely on raw frequencies. However, interpreting patterns of association between variables can be challenging within the confines of a table format. To address this, complex tabular data can be visualized through extended association plots (Cohen 1980; Meyer et al. 2003). Grounded in the principle of residuals, these plots provide an intuitive interpretation of multiple pieces of information in a single visual representation. Furthermore, they enhance the identification of trends across a table. Figure 4 shows the association plot for the data set shown in Table 5, plotting AGE by construction. These plots are interpreted as follows: In the graphical representation of contingency tables, the presence of red and blue tiles serves a pivotal role. These colors are indicative of the direction and magnitude of the residuals. Red tiles signify positive residuals, denoting instances where observed frequencies surpass expectations. Conversely,

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

blue tiles represent negative residuals, signaling scenarios where observed frequencies fall below expectations. Light grey shades indicate the relative absence of a trend. The height of a tile corresponds to the absolute value of the residual, with tiles positioned above the line for positive residuals and below the line for negative residuals. Additionally, the width of a tile reflects the observed frequency, being wider for higher frequencies.

Figure 4: CXN by AGE in an extended association plot

It became clear that in this data set, as indicated by the examination of the tables, CPs show a significant positive association with NHG nouns, depicted as high, dark blue tiles above the line, whereas VNCs show a significant positive association with ENHG nouns. This is noteworthy because it suggests that both constructions may be subject to different diachronic dynamics. While CPs exhibit a particular affinity for the most recent available lexical units, VNCs seem to integrate them to a lesser extent. Given this background, the comparison of the developmental curves in Figure 1 and Figure 2 should also be reinterpreted: The significant increase in CPs in the 19th century correlates with the rise of

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

modern lexemes in the structure. In VNCs, the increase is more moderate and diachronically preceding. The striking association of CP with OHG nouns, on the other hand, will need to be explained differently, as it contradicts the diachronic trend. The question arises whether it is possible to harmonize the trends of lexical favoring in specific periods with discourse-functional characteristics. A first step in this direction is taken with the second parameter.

4.2 GENRE – N_{dev}

The examination of the parameter GENRE is based on the same principles as for AGE. Table 6 shows the residuals:

Table 6: Pearson Residuals, CXN by GENRE

genre	CP	VNC	ZERO
fiction	-2.49	-1.28	1.30
newspaper	-1.54	-2.55	-1.51
reference	-1.02	3.32	-1.02
science	2.63	-0.41	-0.66

The data reveal some interesting patterns. Firstly, it is notable that, for GENRE, there are generally no statistical fluctuations to the extent observed with AGE. Regarding CPs, there is an expected distribution pattern: They are particularly present in scientific texts, while there is a significant negative correlation in fiction. On the other hand, VNCs exhibit a stronger tendency towards reference literature, with other genres showing no positive values. In the case of ZERO, all residuals fall within the range of 0 to ±2, indicating that the disparity between observed and expected values is minimal and can be interpreted as random. This finding is far from trivial, as it indicates that the general use of deverbal nouns is not genre-driven, as often assumed. Even with lower expectations for significance values, it must be noted that nouns outside the constructions CP and VNC exhibit a slight tendency towards narrative texts. Figure 5 shows the association plot for GENRE.

In summary, specific patterns emerge, but they are confined to the individual construction level. Both CPs and VNCs exhibit a preference for nouns from specific historical periods. CPs show an affinity for NHG nouns, while VNCs favor ENHG nouns. Interestingly, there's a reciprocal disfavoring, where CPs tend to disfavor ENHG nouns, and VNCs disfavor NHG nouns. Additionally, both constructions display a distinct genre-specific distribution, favoring non-narrative

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

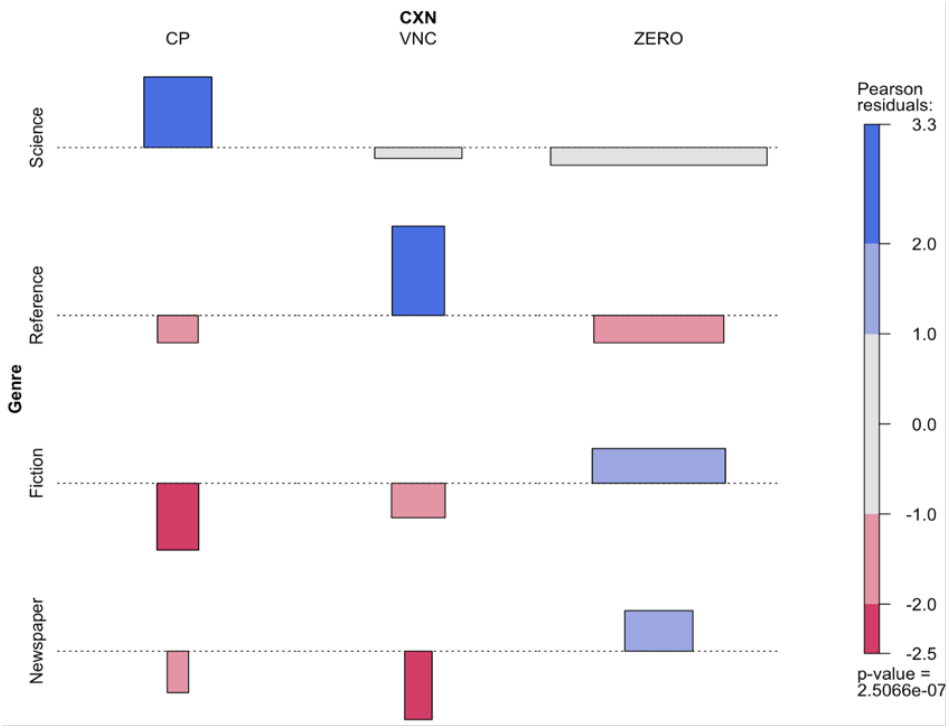


Figure 5: CXN by GENRE in an extended association plot

genres. The reciprocal preference extends to GENRE as well, since each construction is more prominent in a specific genre within the non-narrative category. This reciprocal pattern in both historical periods and different texts adds an intriguing layer to the analysis. The general difficulty in studying non-association lies in the challenge of providing clear evidence for absence, particularly when dealing with linguistic data. Therefore, the first step is to clarify the role that ENHG and NHG nouns play in their respective constructions.

4.3 The emergence of [*in* N_{dev}_]

To obtain a more precise understanding of the function of nouns of different ages within the structure [*in* N_{dev}_], I applied a Correspondence Analysis (Glynn 2014) as an explorative technique to identify patterns and dependencies among the categories and represent them graphically. It is particularly useful when dealing with large data sets with multiple categorical variables, providing insights into the structure and associations within the data. – Especially when dealing with

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

two variables, where one exhibits a higher level of complexity compared to the other. For this analysis, the two variables GENRE and NOUNS have been chosen. GENRE involves a relatively limited set, only four distinct genre types, while the high number of individual lexical types complicates a clear presentation in a contingency table: CPs are distributed across 191 different nouns in Period 9 (69 of them being of NHG origin), VNCs across 231 (of which 74 being ENHG).

For CPs, all the nouns that appeared only once in the data set where removed, leaving a set of 20 lexemes. The exclusion of hapaxes from the investigation is imperative in techniques such as CA, where the objective is to unveil relationships and associations between variables by exploring co-occurrence patterns. Hapaxes contribute disproportionately to the sparsity of the data matrix without providing substantial information regarding their associations with other variables. Their singular occurrences lack the consistency and repetition required to discern statistically significant patterns. Figure 6 displays the reduced data set and the corresponding CA plot.

The observable patterns for CPs reveal a predominant clustering of the majority of nouns and non-narrative genres like 'Reference literature' and 'Science' on the left side of the plot. The axis weighting within the graph indicates that over 83% of the data can be explained through the genre dichotomy 'narrative – non-narrative'. The size of the dots represents the frequency or weight of the categories or variables they represent. Larger dots indicate higher frequency, while smaller dots indicate lower frequency in the analyzed data.

When looking at the most frequent lemmas, certain discourse-functional patterns become apparent. The constructions [*im Verhältnis* P] 'in relation to', [*im Vergleich mit*] 'in comparison to' and [*im Zusammenhang mit*] 'in connection with' are used to establish relationships, make comparisons, and highlight connections between different elements or concepts. They are employed to provide context, clarify relationships between ideas in discussion, or emphasize the relevance of these constructions contribute to precision and clarity, helping readers understand the context or the comparative nature of the statements being made. Instead of explicitly emphasizing who is engaging in the discussion of the topics, there is a backgrounding of the agent, placing less focus on the individuals involved in the discourse. In addition, sentences can be formulated that do not have an agent at all. This strategy can contribute to a more objective or formal tone, making the text more impersonal or general in nature, see (6–6).

(6) *Ich bezweifle, ob man die Vergangenheit* ***in ihrem Verhältnis zur Gegenwart*** *überhaupt richtig zu beurteilen vermag.*

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

'I doubt whether one is able to judge the past **in its relation to the present** correctly at all.' (Chamberlain, 1899)

(7) *Wir wissen, welche kleine Zahl regierender Königinnen die Geschichte* ***im Vergleich zu den Königen*** *aufzuweisen hat.*

'We know what a small number of reigning queens history has **compared to kings**.' (Ichenhaeuser, 1898)

(8) *Umgekehrt können* ***im Zusammenhang mit dem Rechte*** *des Gemeingebrauchs dem Einzelnen auch besondere Nachteile erwachsen.*

'Conversely, **in connection with the right** of public use, the individual may also suffer special disadvantages.' (Mayer, 1896)

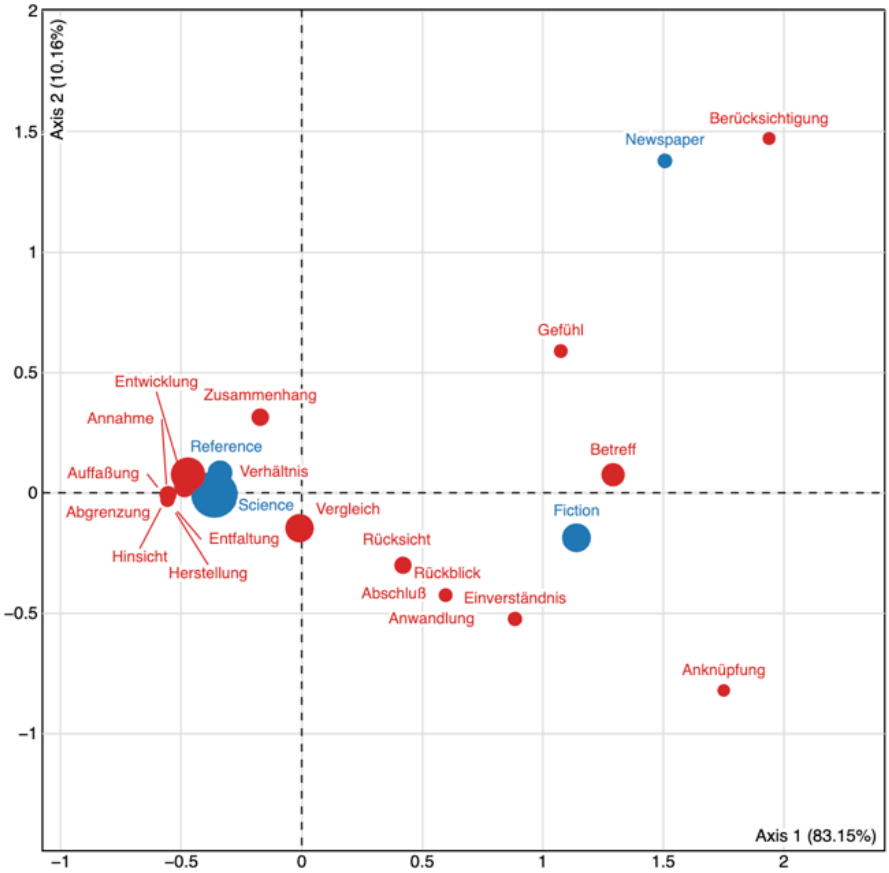


Figure 6: NHG nouns by GENRE in a CA plot (CP), see Table 9

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

This subgroup of nouns exhibits a commonality with regard to the concept of RELATION and suggest a thematic connection associated with analyzing relationships of different subjects.

We can distinguish a second major subgroup from RELATION, falling under the broad category of CONSIDERATION including constructions like [*in Rücksicht* P/GEN] 'in consideration of', [*in Hinsicht* P/GEN] 'in view of' and [*in Betreff* GEN][8] 'regarding' and others. The nouns used in these constructions suggest a thematic connection related to cognitive processes like consideration, understanding, agreement, and acknowledgment, see (9–11).

(9) *Dadurch war seiner Untersuchung* ***in Hinsicht der individuellen Tugenden*** *der Weg vorgezeichnet*

'Thus the way was marked out for his investigation **with regard to the individual virtues**.' (Natorp, 1899)

(10) *Nichtbefolgung der Anweisungen wird* ***in Rücksicht auf Strafen und Zwangsmittel*** *ebenso behandelt wie die Weigerung der Erfüllung überhaupt.*

'Non-compliance with the instructions shall be treated in the same way as refusal to comply at all, **with regard to penalties and coercives**.' (Mayer, 1896)

(11) *Beobachtungen* ***in Betreff der umgebenden Natur*** *(z.B. des gestirnten Himmels) können schon weit gediehen und ein mannigfaltiger Götter- und Dämonenkultus entstanden sein.*

'Observations **concerning the surrounding nature** (e.g. the starry sky) may already have developed far and a manifold cult of gods and demons may have arisen.' (Chamberlain, 1899)

As the constructions within the group of RELATION, the constructions in the given examples serve to shift the focus away from the specific agent or doer of the

[8]The noun *Betreff* is somewhat misleading in the plot. Essentially, it is equivalent to *Hinsicht* but seems to exhibit a tendency toward fictional texts. However, this inclination is solely due to the extensive use of the construction [*in Betreff* GEN] by Otto von Bismarck in his memoirs. Outside of this context, *Betreff* appears only in scientific texts. The less frequent constructions [*in Rückblick auf*] and [*in Rücksicht auf*] are also used by Bismarck, which accounts for their middle position within the plot. This illustrates the limitations of exploratory statistical methods, as they cannot account for the individual habits of specific authors. Not surprisingly, Otto von Bismarck leans towards a formal style even in his personal writings, reflecting his accustomed political demeanor.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

action, emphasizing instead the perspective, context, or considerations related to the action. In this case, the attention is directed towards the investigation and the path that was outlined in terms of individual virtues, rather than explicitly highlighting who conducted the investigation.

We can conclude that NHG nouns, which are preferred in the structure [*in* N_{dev} P/GEN], are used to form a REFERENCE construction to meet the requirements of non-narrative, especially scientific formal texts. Both RELATION and CONSIDERATION can be understood as discursive means of REFERENCE: Mental content is introduced into the discourse, whereby in the case of RELATION the aim is to compare different subjects, while in the case of CONSIDERATION individual subjects are established as a frame of REFERENCE within which various independent subjects are perspectivized. Table 7 summarizes these findings.

Table 7: Semantic groups within the general pattern [in N_{dev} P/GEN]

group	reference	
subgroup	relation	consideration
NOUNS	*Verhältnis, Vergleich, Entwicklung, Zusammenhang, Abschluss, Entfaltung*	*Betreff, Rücksicht, Hinsicht, Auffassung, Einverständnis, Rückblick, Berücksichtigung, Anknüpfung*
CHARACTERISTICS	connections, similarities, or associations between different subjects	cognitive aspects related to the author's perspective and consideration of the subject
CONSTRUCTION	[*in* $N_{reference}$ P/GEN]	

The subgroup RELATION is characterized by constructions that involve comparing or establishing relationships between different subjects or concepts, expressing connections, similarities, or associations between entities, without explicitly highlighting the agent or source of the comparison. On the other hand, the subgroup CONSIDERATION appears to involve cognitive aspects related to the author's perspective and consideration of the subject matter. These constructions suggest a focus on the author's examination, understanding, or acknowledgment of various aspects, and they may indicate a form of consideration or agreement with certain ideas or points. This subgroup reflects a discourse strategy where

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

the author engages in a more reflective and contemplative stance towards the subject without explicitly foregrounding themselves as the agent of the action.

The emergence of a REFERENCE construction explains the increase in use of the structure in the late 18th and 19th century, as it is mainly used in modern like scientific texts. This assumption is supported by the picture that emerges when contrasting the distribution of NHG nouns with those of other periods. Figure 9 shows the distribution of the 20 most frequent OHG, MHG and ENHG nouns in relation to their genre affinity.

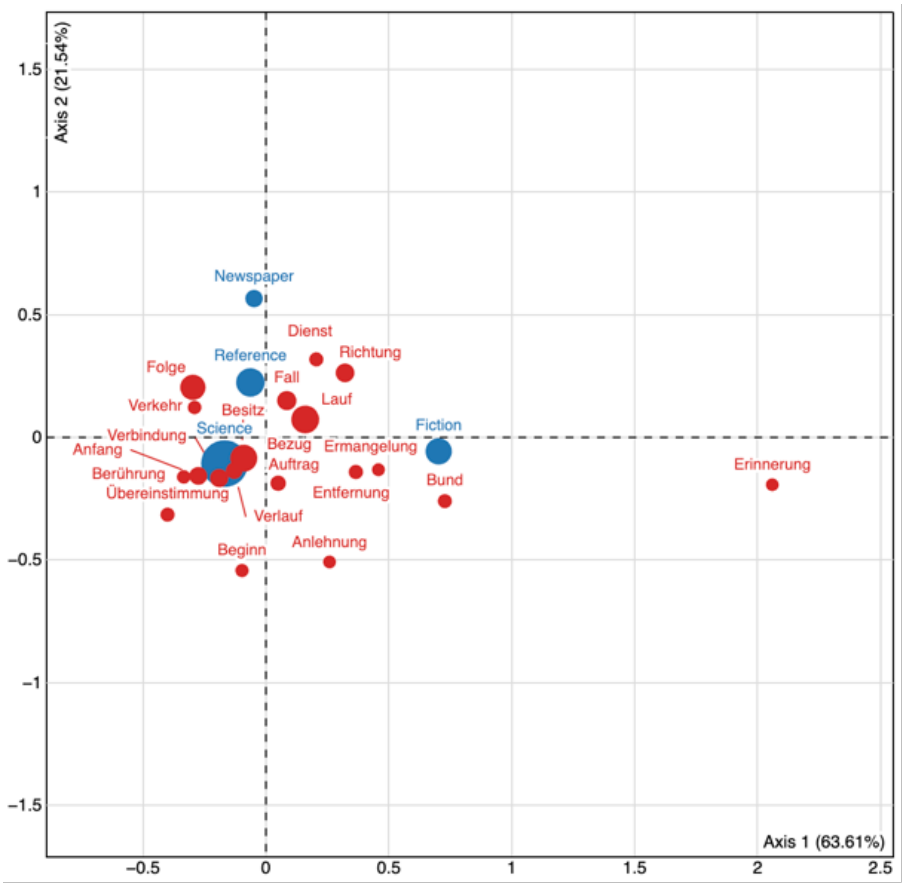


Figure 7: PROTOG, OHG, MHG and ENHG nouns byGENRE in a CA plot (CP), see Table 10

The distribution pattern presents a less structured image compared to NHG nouns. A clear distinction between narrative and non-narrative genres, as seen

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

in Figure 6, does not emerge to the same extent. While individual lexical representatives of the reference group like *Bezug* 'relation', *Verbindung* 'connection' or *Übereinstimmung* 'agreement' can be found here as well, they coexist with a variety of nouns with very heterogeneous semantics. *Bezug* shows most clearly that the existence of a productive REFERENCE construction is to be assumed for the period under examination. As an OHG word it is the prototypical and most used noun with REFERENCE semantics in German. Nevertheless, it appears in the structure [*in* N_{dev} _] only from the 18th century, see Figure 10. It seems to be a general trend in German that *ung*-derivations are increasingly being replaced by products of older word formations (see Demske 2000: 373; Hartmann 2016: 259). Schmidt (2007: 153) sees a tendency towards shortening from the 18th century onwards, which could be linked to the development of technical language. This would not only explain why *Bezug* possibly prevails over younger *Beziehung*, but also why *Lauf* prevails over *Verlauf*. These two frequent lexemes are responsible for the striking association of CPs with OHG nouns, which is shown in Figure 6. The appearance of *Bezug* further supports the assumption of a NHG REFERENCE construction. While *Rücksicht*, *Hinsicht*, and *Berücksichtigung* originated relatively recently, all these nouns are attested almost simultaneously together with *Bezug* within the pattern [*in* N_{dev} _]. This simultaneous emergence, despite the varied historical periods of their coinage, lends credence to the likelihood of a formation schema guiding their usage.

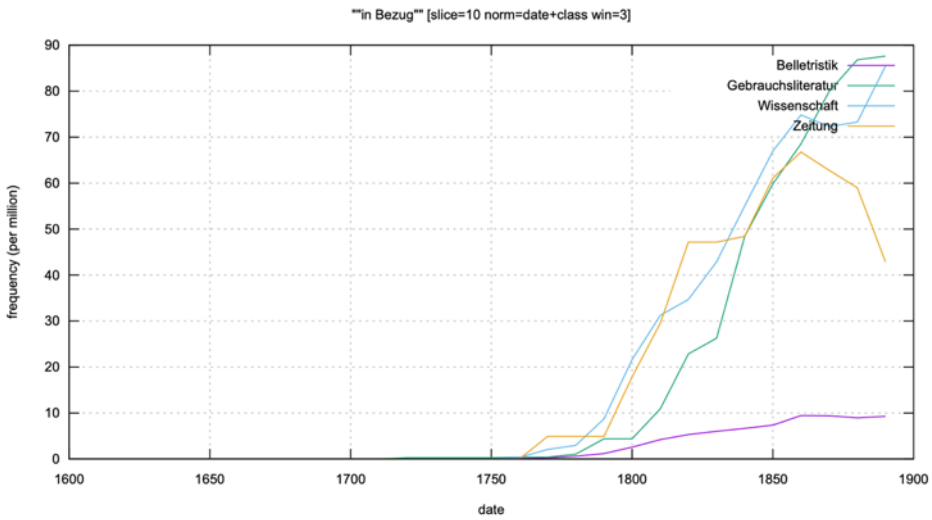


Figure 8: Token frequency of [in Bezug_] in the DTA Kernkorpus

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

When examining VNCs and their favored ENHG nouns with the highest frequency, two observations emerge. Firstly, the token count significantly exceeds that of CPs, as shown in Figure 6. Secondly, the organization of the semantic space with regard to GENRE appears less rigid. While CPs exhibited a two-parted system with a clear distinction between narrative and non-narrative genres, VNCs present a more triangular configuration in which the non-narrative genres in particular form different clusters, see Figure 9.

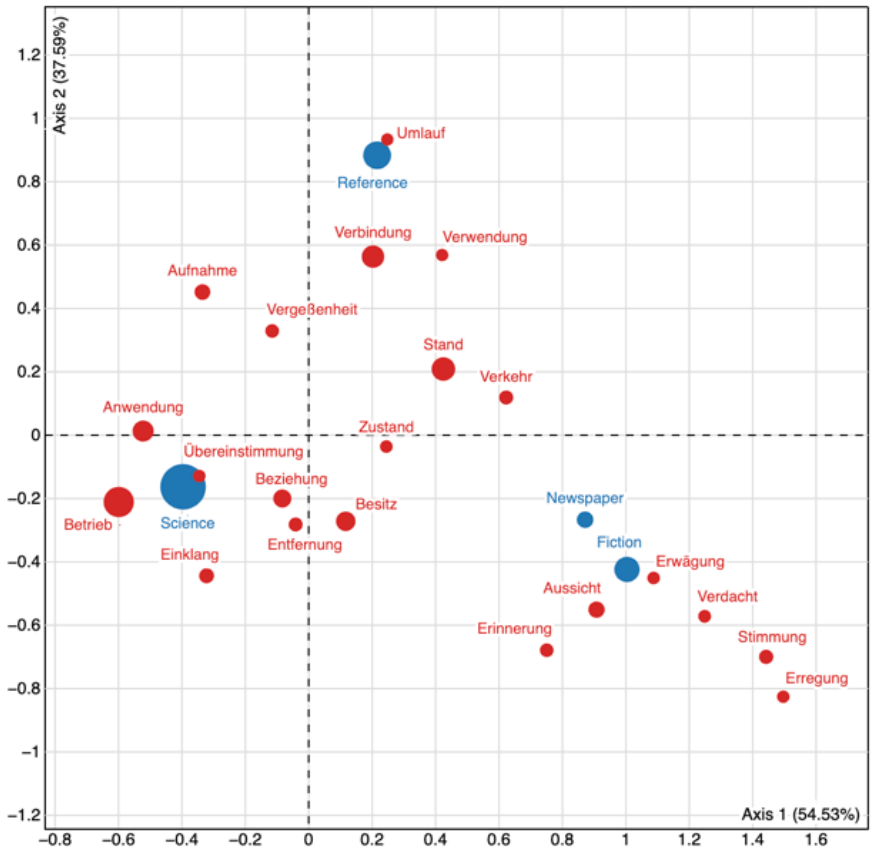


Figure 9: ENHG nouns by GENRE in a CA plot (VNC), see Table 11

The most significant difference compared to CPs is that non-narrative genres exhibit distinct behaviors concerning the structure [*in* N_{dev} V]. The sorting of individual functional clusters proves to be significantly more challenging, as no semantic group dominates in absolute frequency. Just as for CPs, a certain ten-

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

dency towards REFERENCE can also be observed for VNCs. Both the subgroup RELATION and the subgroup CONSIDERATION are represented, see (12)–(14).

(12) *Die vielfachen, namentlich in England gemachten Verſuche, denſelben mit der modernen Entwickelungslehre* ***in Einklang zu bringen****, ſind völlig fehlgeſchlagen.*
'The many attempts, especially in England, to **bring it in line** with the modern theory of development have completely failed.' (Haeckel, 1899)

(13) *Form und Jnhalt der Poesie* ***treten in Uebereinstimmung****.*
'The form and content of the poetry **coincide**.' (Wolff, 1899)

(14) *Er* ***hat*** *nämlich den Ort* ***in Erwägung gezogen****, von wo dieſe Wetterfahne ſtammt.*
'In fact, **he has considered** the place from where this weathervane comes.' (Fontane, 1899)

As can be seen from example (14), the agent-backgrounding function of VNC is not as distinct as it was in the case of CPs. This is due to the fact that the argument structure of VNCs requires the explicit mention of a subject, unless a passive construction is employed. It is therefore not coincidental that VNCs are used more frequently in narrative genres. Non-narrative scientific texts necessarily have fewer agentive subjects due to their focus on presenting objective information and analysis rather than on describing actions performed by specific agents. The emphasis is typically on conveying findings, observations, and interpretations rather than on detailing the actions of individuals or entities. As a result, the use of agentive subjects may be minimized in favor of passive constructions or nominalizations, which allow for a more impersonal and objective presentation of information. VNCs exhibit a greater openness towards different genres, reflected in expressions of emotional or psychological states with nouns like *Stimmung* 'mood' or *Erregung* 'excitement', which mainly require animate and agentive subjects, see (15).

(15) *Übrigens fange ich an,* ***in Stimmung zu kommen****.*
'By the way, I'm starting **to get in the mood**.' (Bierbaum, 1897)

Another noticeable subgroup comprises nouns that can be associated, in the broadest sense, with the concept USAGE. It includes nouns related to aspects of processes and activities. These nouns suggest a focus on how things operate, are used, controlled or function in a practical sense, see (16–17).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(16) *Der Union-Ofen bei Connelsville* ***kam*** *1791* ***in Betrieb****.*
'The Union Furnace at Connelsville **came into operation** in 1791.'
(Beck, 1897)

(17) *Das Verfahren* ***wurde*** *1810 in dem damaligen Königreich Westfalen bekannt gemacht und auch* ***in Anwendung gebracht****.*
'The method **was** introduced and **applied** in the Kingdom of Westphalia in 1810.' (Beck, 1899)

Table 8 provides an overview of the rough classification of Early New High German (ENHG) nouns in the structure [*in* N_{dev}V].

The variety of usage contexts of ENHG nouns the structure [*in* N_{dev}V] makes it clear that hardly any construction type is able to prevail over others. It is therefore not surprising that the semantic space differs little compared to that when considering all other nouns. Figure 10 shows the distribution of the 20 most frequent OHG, MHG and NHG nouns in relation to their genre affinity.

It must be noted that such an approach, which has proven fruitful for CPs, is misguided for VNCs. While *Betracht* 'consideration', *Gebrauch* 'usage', *Bewegung* 'motion', *Frage* 'question', and *Gang* 'motion' can be assigned to the identified subgroups, the semantic heterogeneity of the remaining lexical items is difficult to meaningfully categorize. At first, older nouns like *Betracht* seem to indicate that VNCs are also drawn into the pull of a REFERENCE construction, as from the 18th century, there is an increased use of [*in Betracht* _], see Figure 11.

Betracht is actually a MHG adjective that was rarely used as a noun before the 17th century. In the 18th century, it was able to establish itself against the lexical competitor *Betrachtung*. It is conceivable that the shorter *Betracht* was preferred in more fixed and specified constructions, as it does not exist as an autonomous noun outside of [*in Betracht* V]. Here, again, we encounter the aforementioned tendency to shortening. With the increasing use of *Betracht*, the decline of *Betrachtung* was initiated, see Figure 12.

While *Betrachtung* no longer appears in [*in* N_{dev} V] in contemporary German, it is still in use outside of the construction. [*in Betrachtung* _][9] was already an established construction in earlier periods. In summary, it can be said that VNCs exhibit a high degree of diversity in terms of GENRE and are challenging to cluster effectively with respect to their discourse-pragmatic functions. The increase in formations of this type can be explained from this perspective only insofar as

[9]In this case, [*in Betrachtung* _] and [*in Betrachtung* V] can be considered almost identical, as the structure is formed almost exclusively with a verb.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

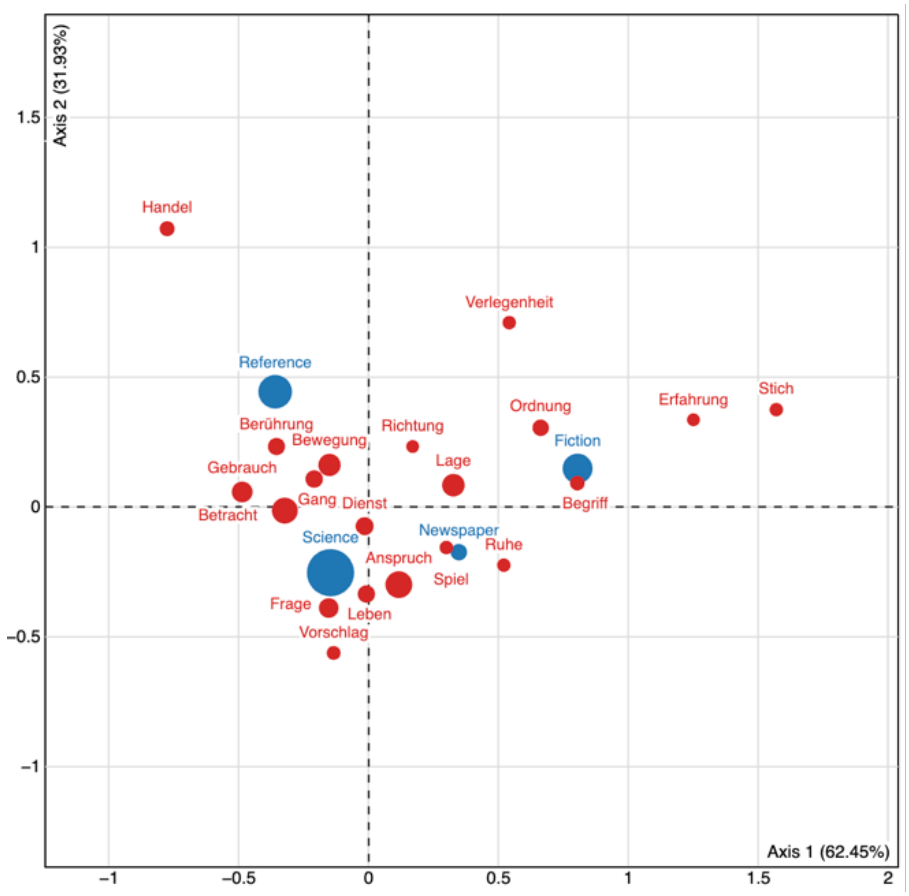


Figure 10: PROTOG, OHG, MHG and NHG nouns by GENRE in a CA plot (VNC), see Table 12

there was already an established structure that was additionally fueled by the increased availability of nominal material in Early New High German.

5 Discussion of the results

CPs are strongly associated with the function REFERENCE, reflecting a specialized syntactic and discourse-pragmatic development as required by evolving communicative needs and changes in textual conventions. This includes, first and foremost, agent backgrounding. Non-narrative texts, especially scientific papers or technical documents, prioritize the communication of information. Agent back-

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

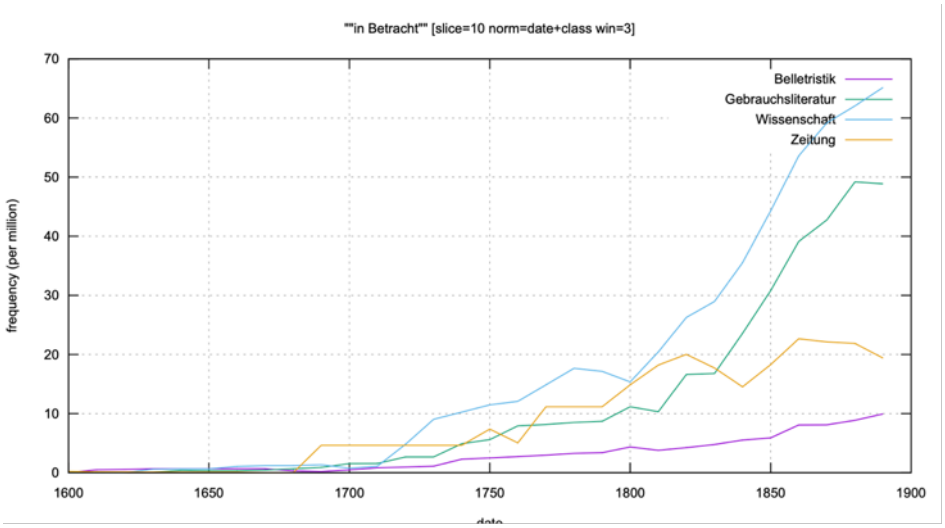


Figure 11: Token frequency of [in Betracht_] in the DTA Kernkorpus

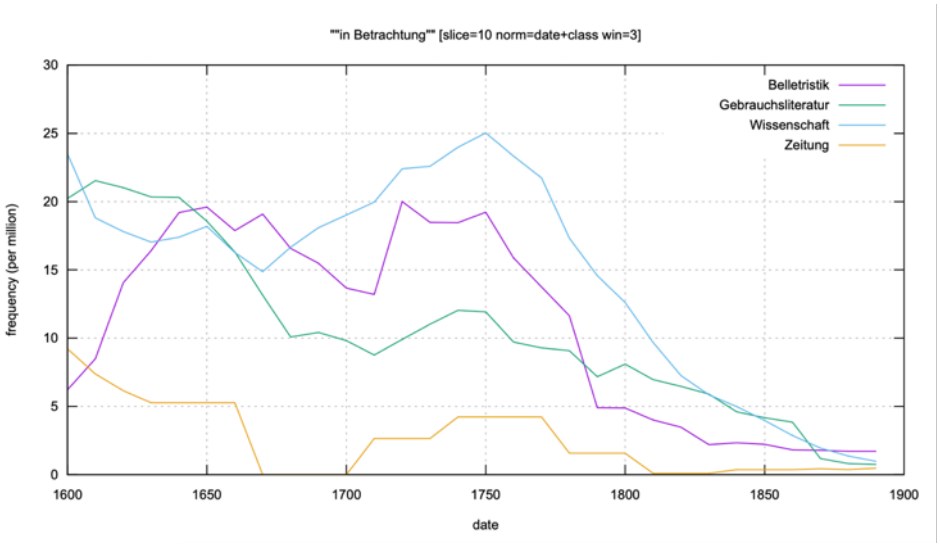


Figure 12: Token frequency of [in Betrachtung_] in the DTA Kernkorpus

grounding helps to keep the focus on the actions, processes, or concepts being discussed rather than on the individuals involved. REFERENCE constructions can be employed to establish a clear and precise connection between concepts or entities

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

without foregrounding the agent. CPs also play a crucial role in condensing information in a concise and efficient manner. Unlike whole subordinated clauses, complex prepositions allow for the expression of complex relationships and ideas within a compact structure. This conciseness is particularly valuable in scientific texts, where clarity and precision are paramount. In contrast to CPs, which excel at expressing relationships in a compact form, VNCs operate differently in terms of structure, since they form syntactic patterns that include a subject position. Therefore, VNCs lack the obligatory agent backgrounding, even though numerous constructions of the pattern [*in* N_{dev} V] allow for an agent-shifted perspective. It is not surprising that some constructions of this type can be found, which are also associated with a realm of RELATION and CONSIDERATION and thus contribute to the discourse-functional strengthening of the structure [*in* N_{dev} _]. In contrast to CPs, multi word predicates like VNCs are not suitable for information summarization, as they require significantly more material compared to simplex verbs to verbalize similar semantic content. Because of these differences, it would therefore not be justified to speak of a separate semi-schematic construction [*in* N_{dev} _] with the meaning REFERENCE. If one wishes to continue operating within this constructional framework, we would need to assume a more abstract schema at the discourse level, which is constituted by the CPs, but also by some VNCs. In recent years, many linguists have recognized that constructions encompass not only conventional patterns within individual syntactic patterns but also extend to broader linguistic units (for an overview, see Enghels & Sansiñena 2021). But the integration of these approaches into constructional frameworks is still up for debate, as they would necessarily entail loosening specific form-meaning dichotomies. Considering the existence of such discursive macrostructures, CPs and VNCs would at least be suitable functional bridging contexts, as they also exhibit a formal overlap establishing a potential associative relationship. Figure Figure 13 shows the partial network of such a REFERENCE discourse schema.

CPs contribute significantly to the existence of the schema. In contrast, the contribution of VNCs is considerably less strong, since the semantics of the nouns within the structure [*in* N_{dev} V] are too heterogeneous for a prototypical meaning to emerge to the same extent. There is, therefore, much to suggest that we are dealing with much more diverse and freer syntactic patterns in VNCs, which were already established earlier. An abstract discourse schema cannot be derived for this structure. This is also indicated by the development curves of individual nouns such as *Betrachtung*, which, as the oldest reference noun, shows a historically stable frequency. From New High German onwards, an increase in the frequency of VNCs can be observed, as an enlarged lexical pool is available for

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

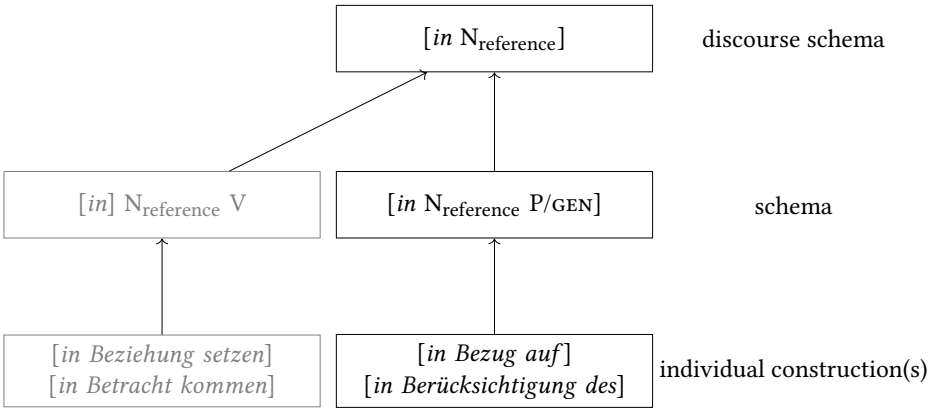


Figure 13: Partial network of [in N_{dev} _]

the established structure to draw on. Figure 14 shows the partial network for VNCs.

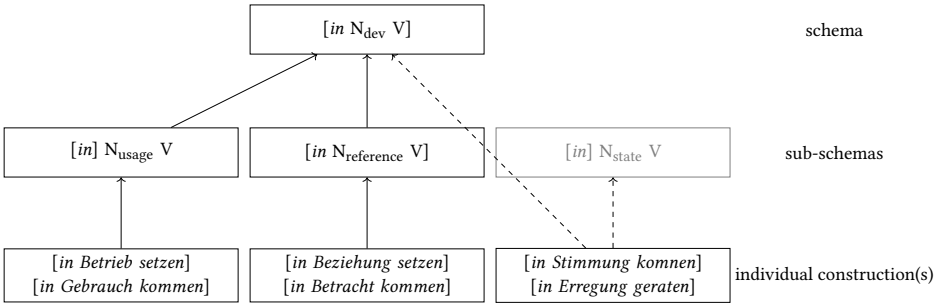


Figure 14: Partial network of German verbonominal constructions [in N_{dev} V]

It is evident that VNCs do not exhibit uniform discourse-specific behavior. The two identified and frequency dominant sub-schemas USAGE and REFERENCE reflect the functionality of the texts in which they predominantly appear. They differ primarily in the deictic origo shift with which they are prototypically associated. While the reference construction often fulfills the function of establishing an author-distant discourse perspective, this is not necessary for the usage construction. The agent involved, whether explicitly named or not, is usually not the author himself. It is not surprising that, unlike CPs, VNCs can be seen to have a more even distribution of the most frequent nouns across different genres, although there is some tendency towards non-narrative texts. In this respect, however, the nominal cores of VNCs differ little from autonomous deverbal nouns

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

occurring outside the structures under study. This is shown by a concluding look at the distribution of the 20 highest frequency ZERO nouns, see Figure 15.

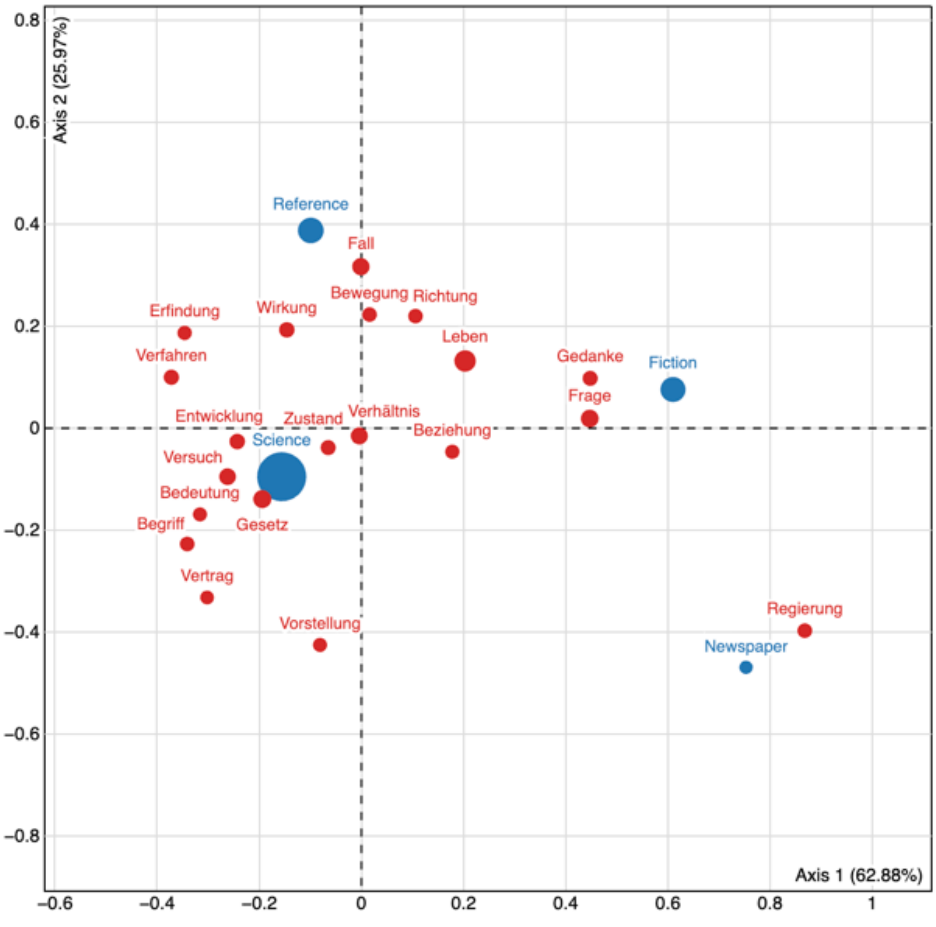


Figure 15: Most frequent nouns by GENRE in a CA plot (ZERO), see Table 13

The triangulated system, which we already encountered with the VNCs in Figure 11, shows up again here. It most closely reflects the general distribution of nouns in the 19th century. It seems that the assumption of a general nominal style in non-narrative text types seems to be confirmed, even independently of the two structures studied. VNCs of the pattern [*in* N_{dev} V] only deviate slightly from this genre-typical null model, but CPs of the pattern [*in* N_{dev} P/GEN] deviate significantly.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

6 Conclusion

This paper has offered a diachronic take on the formation of complex prepositions and verbonominal constructions in German. Two distinct syntactic patterns were analysed, [*in* N_{dev} P/GEN] exhibiting clear characteristics and [*in* N_{dev} V] displaying a broader, more open structure with diverse semantics. This leads me to some conclusions, which return to the observation mentioned at the beginning that German has developed a strong tendency towards nominal style. This phenomenon seems to be mainly related to the development of modern non-narrative text structures in German, as often assumed. Embedded in this diachronic trend is the development of a REFERENCE discourse schema, which is mainly supported by CPs. Interestingly, this construction shows a very strong tendency towards the most modern nouns available to 19th century German. The availability of these nouns thereby stimulates the discursive schematization of the pattern [*in* N_{dev} _]. Looking across constructional boundaries also shows that focusing on purely formal or structural characteristics of individual constructions in a discourse-functional context should not be overemphasized. CPs with other prepositions than *in* show similar behavior. This is exemplified by the schema [*mit* N_{dev} *auf*/GEN] with the nouns *Bezug* 'regard', *Beziehung* 'relation', *Berücksichtigung* 'consideration', *Hinsicht* 'regard', and *Rücksicht* 'consideration' in our data. Ruf & Smirnova (forthcoming) observe a rapid emergence of these five individual CPs following the same formation pattern and displaying REFERENCE semantics. This diachronic scenario does not necessarily depend on the gradual routinization and conventionalization of a particular syntagmatic string. Instead, it involves the creative utilization of resources available to the speakers of a speech community. In this process, new lexical items in a language are created almost instantly through the creative application of a REFERENCE schema.REFERENCE discourse schema Competing structures with the same function should prove fruitful as future fields of research, since it cannot be ruled out that different prepositions are also used to convey nuanced discourse functions. However, such functions are not foreseeable at this point. In addition to the traditional consideration of genre-specific distribution patterns, it also became clear that the parameter AGE in relation to different constituents of constructions should be considered in future studies. From a speaker's point of view, it must be assumed that established schemata are usually insensitive with respect to the age of their nominal elements. After all, language users are not usually aware of the age of a lexeme and nouns of all ages can be found in both constructions. Across the multitude of observable contexts, however, it should be noted that modern structures are also preferentially filled with modern material

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

during the initial phase of their emergence. Since VNCs can draw on established argument structures, they are available in an earlier phase of the expanding nominal pool. Despite their continued increase throughout the following centuries, they are particularly prone to reject those nouns that emerge during this time. CPs, on the other hand, are primarily a fashion of the 19th century and reflect the strongly genre-specific characteristics of scientific texts. They are therefore representatives of the delayed emancipation of German academic language mentioned initially. However, this cannot be extended to VNCs, and consequently, not to deverbal nominalizations as a whole.

Acknowledgements

Research reported in this paper results from the research project "Zusammenspiel von Wortbildung und Syntax: Nominalisierungsstrategien in verbonominalen Konstruktionen und sekundären Präpositionen" (2021–2025) supported by the Swiss National Science Foundation, Grant number 197124.

References

Admoni, Vladimir G. 1973. *Die Entwicklungstendenzen des deutschen Satzbaus von heute*. München: Max Hueber.

Beneš, Eduard. 1974. Präpositionswertige Präpositionalfügungen. In Ulrich Engel, Paul Grebe & Hugo Moser (eds.), *Sprachsystem und Sprachgebrauch: Festschrift für Hugo Moser*, 1st edn., vol. 33 (Sprache der Gegenwart), 33–52. Düsseldorf: Schwann.

Cohen, Ayala. 1980. On the graphical display of the significant components in a two-way contingency tables. *Communications in Statistics – Theory and Methods* 9(10). 1025–1041. DOI: 10.1080/03610928008827940.

Demske, Ulrike. 2000. Zur Geschichte der *ung*-Nominalisierung im Deutschen: Ein Wandel morphologischer Produktivität. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 122. 365–411. DOI: 10.1515/bgsl.2000.122.3.365.

Di Meola, Claudio. 2000. *Die Grammatikalisierung deutscher Präpositionen* (Studien zur deutschen Grammatik 62). Tübingen: Stauffenburg.

Eggers, Hans. 1973. *Deutsche Sprache im 20. Jahrhundert*. München: Piper.

Enghels, Renata & María. S. Sansiñena. 2021. Discourse-level phenomena in construction grammars. *Constructions and Frames* 13(1). DOI: 10.1075/cf.00045.int.

Fleischhauer, Jens. 2022. *Stehen unter* – Funktionsverbgefüge und ihre Familien. *Zeitschrift für Germanistische Linguistik*. 247–288. DOI: 10.1515/zgl-2022-2055.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Fleischhauer, Jens & Stefan Hartmann. 2021. The emergence of light verb constructions: A case study on German *kommen* 'come'. *Yearbook of the German Cognitive Linguistics Association* 9(1). 135–156. DOI: 10.1515/gcla-2021-0007.

Glynn, Dylan. 2014. Correspondence analysis: Exploring data and identifying patterns. In Dylan Glynn & Justyna A. Robinson (eds.), *Corpus methods for semantics: Quantitative studies in polysemy and synonymy*, 443–485. Amsterdam: Benjamins.

Harm, Volker. 2021. *Funktionsverbgefüge des Deutschen: Untersuchungen zu einer Kategorie zwischen Lexikon und Grammatik*. Berlin: De Gruyter.

Hartmann, Stefan. 2016. *Wortbildungswandel: Eine diachrone Studie zu deutschen Nominalisierungsmustern.* (Studia linguistica Germanica 125). Berlin: De Gruyter.

Heinz, Kretzenbacher L. & Harald Weinrich. 1995. *Linguistik der Wissenschaftssprache*. Berlin: De Gruyter.

Helbig, Gerhard. 1979. Probleme der Beschreibung von Funktionsverbgefügen im Deutschen. *Deutsch als Fremdsprache* 16. 273–285.

Hoffmann, Sebastian. 2005. *Grammaticalization and English complex prepositions: A corpus-based study* (Routledge advances in corpus linguistics 7). London: Routledge.

Kabatnik, Susanne. 2020. *Leistungen von Funktionsverbgefügen im Text: Eine korpusbasierte quantitativ-qualitative Untersuchung am Beispiel des Deutschen und Polnischen*. Tübingen: Narr.

Kamber, Alain. 2008. *Funktionsverbgefüge – empirisch: Eine korpusbasierte Untersuchung zu den nominalen Prädikaten des Deutschen* (Reihe Germanistische Linguistik 281). Tübingen: Niemeyer.

Lehmann, Christian. 1998. German abstract prepositional phrases. In Iørn Korzen & Michael Herslund (eds.), *Clause combining and text structure* (Copenhagen studies in language), 87–106. Frederiksberg: Samfundslitteratur.

Lehmann, Christian & Christel Stolz. 1992. *Bildung von Adpositionen im Deutschen* (Arbeitspapiere des Seminars für Sprachwissenschaft der Universität Erfurt 6). Erfurt: Seminar für Sprachwissenschaft.

Lindqvist, Christer. 1994. *Zur Entstehung von Präpositionen im Deutschen und Schwedischen* (Linguistische Arbeiten 311). Tübingen: Niemeyer.

Meibauer, Jörg. 1995. Komplexe Präpositionen – Grammatikalisierung, Metapher, Implikatur und division of pragmatic labour. In Frank Liedtke (ed.), *Implikaturen: Grammatische und pragmatische Analysen* (Linguistische Arbeiten 343), 47–74. Tübingen: Niemeyer. DOI: 10.1515/9783110958867.47.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Meyer, David, Achim Zeileis & Kurt Hornic. 2003. Visualizing independence using extended association plots. In Kurt Hornik, Friedrich Leisch & Achim Zeileis (eds.), *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*. Technische Universität Wien, Institut für Statistik u. Wahrscheinlichkeitstheorie.

Möslein, Kurt. 1981. Einige Entwicklungstendenzen in der wissenschaftlich-technischen Literatur seit dem Ende 18. Jahrhunderts. In Walter von Hahn (ed.), *Fachsprachen*, 276–319. Darmstadt: Wissenschaftliche Buchgesellschaft.

von Polenz, Peter. 1987. Funktionsverben, Funktionsverbgefüge und Verwandtes: Vorschläge zur satzsemantischen Lexikographie. *Zeitschrift für Germanistische Linguistik*. 169–189. DOI: 10.1515/zfgl.1987.15.2.169.

von Polenz, Peter. 2013. *Deutsche Sprachgeschichte vom Spätmittelalter bis zur Gegenwart. Band II. 17. und 18. Jahrhundert*. Berlin: De Gruyter. DOI: 10.1515/infodaf-1996-232-369.

Pörksen, Uwe. 1998. Deutsche Sprachgeschichte und die Entwicklung der Naturwissenschaften – Aspekte einer Geschichte der Naturwissenschaftssprache und ihrer Wechselbeziehung zur Gemeinsprache. In Werner Besch, Anne Betten, Oskar Reichmann & Stefan Sonderegger (eds.), *Sprachgeschichte – Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung* (Handbücher zur Sprach- und Kommunikationswissenschaft), 193–210. Berlin: De Gruyter.

Van Pottelberge, Jeroen. 2001. *Verbonominale Konstruktionen, Funktionsverbgefüge: Vom Sinn und Unsinn eines Untersuchungsgegenstandes*. Heidelberg: Winter.

R Core Team. 2023. *R: A language and environment for statistical computing*. R version 4.2.3 (2023-03-15 ucrt). Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Schmidt, Wilhelm. 2007. *Geschichte der deutschen Sprache: Ein Lehrbuch für das germanistische Studium* (10. Auflage erarbeitet von Helmut Langner und Norbert Richard Wolf). Stuttgart: Hirzel.

Seifert, Jan. 2004. *Funktionsverbgefüge in der deutschen Gesetzessprache* (Germanistische Linguistik 15). Hildesheim: Olms.

Seppänen, Aimo, Rhonwen Bowen & Joe Trotta. 1994. On the so-called complex prepositions. *Studia Anglica Posnaniensia* 24. 3–29.

Smirnova, Elena & Vanessa Stöber. 2022. Verbo-nominal constructions with *kommen* 'come' in German. *Constructions and Frames* 14(1). 121–149. DOI: 10.1075/cf.00060.smi.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Stefanowitsch, Anatol, Elena Smirnova & Matthias Hüning. 2020. Complex adpositions in three West Germanic Languages: German, Dutch, and English. In Benjamin Fagard, José Pinto de Lima, Dejan Stosic & Elena Smirnova (eds.), *Complex adpositions in European languages: A micro-typological approach to complex nominal relators*, vol. 65 (Empirical approaches to language typology), 65–138. Berlin: De Gruyter. DOI: 10.1515/9783110686647-003.

Szczepaniak, Renata. 2009. *Grammatikalisierung im Deutschen: Eine Einführung*. Narr Studienbücher. Tübingen: Narr.

Zeschel, Arne. 2008. Funktionsverbgefüge als Idiomverbände. In Anatol Stefanowitsch & Kerstin Fischer (eds.), *Konstruktionsgrammatik: Von der Konstruktion zur Grammatik*, 263–278. Tübingen: Stauffenburg.

Appendix

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Table 8: Semantic groups within the general pattern [in N_{dev} V]

group	reference		usage		(emotional) state
SUBGROUP	RELATION	CONSIDERA-TION	OPERATION	(TECHNICAL) STATE	
NOUNS	*Verbindung, Beziehung, Übereinstim-mung*	*Aussicht, Einklang, Erinnerung, Vergessenheit, Verdacht, Erwägung*	*Betrieb, Anwendung, Drehung, Verkehr, Umlauf, Verwendung, Verbindung*	*Stand, Besitz, Aufnahme, Entfernung, Zustand*	*Stimmung, Zustand, Erregung, Beziehung, Verbindung, Stand*
CHARACTER-ISTICS	connections, similarities, or associations between different subjects	cognitive aspects related to the author's perspective and consideration of the subject	nouns related to the practical aspects of processes and activities	nouns related to a controlled state by an agent	feelings or emotions experienced by an animated entity
CONSTRUC-TION			[*in* N_{dev}V]		

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Table 9: NHG nouns by GENRE in a CA plot (CP), see Figure 8

noun	n
Verhältnis 'relation'	125
Vergleich 'comparison'	81
Betreff 'regard'	49
Entwicklung 'development'	36
Zusammenhang 'connection'	20
Rücksicht 'consideration'	19
Hinsicht 'regard'	18
Auffassung 'conception'	8
Einverständnis 'agreement'	8
Rückblick 'regard'	6
Gefühl 'feeling'	6
Berücksichtigung 'consideration'	3
Herstellung 'production'	3
Abschluss 'conclusion'	2
Anknüpfung 'connection'	2
Entfaltung 'unfolding'	2
Anwandlung 'mood'	2
Annahme 'assumption'	2
Anpassung 'adaption'	2
Abgrenzung 'limitation'	2

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Table 10: PROTOG, OHG, MHG and ENHG nouns by GENRE in a CA plot (CP), see Figure 9

noun	n
Lauf 'course'	253
Bezug 'relation'	235
Folge 'consequence'	200
Fall 'case'	94
Richtung 'direction'	87
Verbindung 'connection'	79
Anfang 'begin'	75
Verlauf 'course'	57
Auftrag 'order'	42
Entfernung 'distance'	30
Übereinstimmung 'agreement'	28
Dienst 'service'	28
Bund 'tie'	27
Besitz 'property'	23
Berührung 'contact'	22
Verkehr 'traffic'	21
Beginn 'begin'	19
Ermangelung 'lack'	15
Anlehnung 'backing'	14
Erinnerung 'memory'	13

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Table 11: ENHG nouns by GENRE in a CA plot (VNC), see Figure 11

noun	n
Betrieb 'usage'	277
Stand 'condition'	149
Verbindung 'connection'	131
Anwendung 'usage'	110
Besitz 'property'	87
Beziehung 'relation'	69
Aussicht 'prospect'	47
Aufnahme 'uptake'	40
Einklang 'harmony'	31
Drehung 'turning'	25
Stimmung 'mood'	25
Verkehr 'traffic'	23
Entfernung 'distance'	22
Erinnerung 'memory'	19
Vergessenheit 'oblivion'	19
Verdacht 'suspicion'	11
Zustand 'state'	10
Erwägung 'contemplation'	9
Übereinstimmung 'agreement'	9
Erregung 'excitement'	8
Umlauf 'circulation'	7
Verwendung 'usage'	7

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Table 12: PROTOG, OHG, MHG and NHG nouns by GENRE in a CA plot (VNC), see Figure 12

noun	n
Anspruch 'claim'	222
Betracht 'consideration'	203
Lage 'situation'	144
Bewegung 'motion'	139
Gebrauch 'usage'	115
Frage 'question'	98
Dienst 'service'	76
Gang 'motion'	69
Leben 'life'	67
Berührung 'touch'	65
Ordnung 'order'	56
Handel 'trade'	39
Begriff 'concept'	35
Vorschlag 'suggestion'	27
Spiel 'game'	24
Stich 'stab'	23
Ruhe 'calm'	23
Verlegenheit 'embarrassment'	23
Erfahrung 'experience'	18
Richtung 'direction'	17

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Table 13: Most frequent nouns by GENRE in a CA plot (ZERO), see Figure 15

noun	n
Leben 'life'	2146
Gesetz 'law'	1404
Frage 'question'	1350
Fall 'case'	1282
Verhältnis 'relation'	1230
Versuch 'attempt'	1164
Entwicklung 'development'	944
Wirkung 'effect'	927
Verfahren 'procedure'	906
Gedanke 'thought'	869
Zustand 'state'	860
Regierung 'government'	808
Begriff 'concept'	802
Bewegung 'motion'	796
Richtung 'direction'	788
Erfindung 'invention'	753
Beziehung 'relation'	732
Vorstellung 'concept'	721
Bedeutung 'meaning'	706
Vertrag 'contract'	655

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Chapter 7

Argument realizing verb second clauses in Early New High German narratives: A discourse-pragmatic account

Malika Reetz[a]
[a]Universität Potsdam

This paper presents a corpus study of argument realizing verb second clauses in Early New High German narratives from a discourse pragmatic point of view. By means of a conditional inference tree, it will be shown that there are two pragmatic types of argument realizing verb second clauses at this point in time: One is licensed by mediated assertivity, the other by discourse progressiveness. Furthermore, the mapping between these discourse pragmatic properties and the syntactic form of a clause is - dependent on the respective text - considerably systematic, which will be proposed to be the relevant factor behind the assumed higher frequency of dependent V2 clauses in Early New High German, compared to present-day German.

1 Introduction

In present-day German (PDG), there is a formal distinction between asyndetic verb second structure (henceforth: V2) and syndetic verb final structure (henceforth: VF) in declarative clauses. V2 is broadly used in main clauses, but generally considered to be the non-canonical form of dependent clauses, especially in written language (Auer 1998, Axel-Tober 2012). However, regarding argument realizing clauses, there is variation between these two syntactic forms for some semantic classes of matrix verbs. Representatives of these are verbs that denote an act of assertion (1a), like *sagen* 'saying' or *denken* 'thinking'. In contrast, V2 is generally blocked under factive-emotive verbs (1b), which inherently presuppose the proposition in the argument clause.

Malika Reetz. 2025. Argument realizing verb second clauses in Early New High German narratives: A discourse-pragmatic account. In Ulrike Demske & Barthe Bloom (eds.), *Discourse structure and narration: A diachronic view from Germanic*, 211–231. Berlin: Language Science Press. DOI: ??

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(1) a. *Amanda sagte,* ***dass*** *sie Paul nicht gesehen* ***habe***
Amanda said COP she Paul NEG seen has.AUX.SBJV

b. *Amanda sagte, sie* ***habe*** *Paul nicht gesehen*
Amanda said she has.AUX.SBJV Paul NEG seen
'Amanda said that she hadn't seen Paul.'

(2) a. *Amanda bedauerte,* ***dass*** *sie Paul nicht gesehen* ***hatte****.*
Amanda regretted COMP she Paul NEG seen had.AUX

b. **Amanda bedauerte, sie* ***hatte*** *Paul nicht gesehen.*
Amanda regretted she had.AUX Paul NEG seen
'Amanda regretted that she hadn't seen Paul.'

The semantics of the respective matrix verb is generally assumed to be indicative of the pragmatic features controlling the use of argument realizing verb second clauses (arg.V2). Numerous people appeal to the notion of *assertion* as the relevant notion to account for arg.V2 in PDG (Reis 1997, Auer 1998, Gärtner 2002, Meinunger 2006 amongst others). Yet, it has also been observed that in PDG, there is no iconism between the syntax and pragmatics of a dependent clause: The arg.V2 is always an optional variant of the (canonical) VF-clause (Reis 1997).

Regarding earlier stages of German, especially the periods preceding Early New High German (ENHG), it has been shown that syntactic variation used to be affected by discourse pragmatic categories to a higher degree than in PDG (Hinterhölzl & Petrova 2010, Speyer 2010). However, the question on how argument realizing clauses fit into the picture is not taken into account in these discussions.

In ENHG, a period which is assumed to constitute a transitional state in the standardisation of German written language (Betten 2000, von Polenz 2000), we find a similar syntactic variation in argument clauses in prose texts:

(3) a. *die diener sagten zu dem graffen das er im ain cronen zů*
The servants said to the.DAT count COP he him a krona to
zerung gab̊
consumption gives.SBJV
'The servants told the count that he had given him a krona for consumption.' [FORT, 436]

b. *sagt er hab mir ainen grossen sold geben*
said he has.SUJV me.DAT a big pay given
'[He] said he had given me a big pay.' [FORT, 410]

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Operationalized by means of the arg.V2-selecting matrix verbs, it has been argued that the arg.V2 is historically old (Axel-Tober 2012), and its distribution in Old- and Middle High German "does not differ from that of PDG in a radical way" (Petrova 2020b: 556). Against the background that syntactic variation is stronger affected by discourse pragmatics in earlier stages of German than in PDG, the question arises whether arg.V2 is used much more systematically in ENHG written text than in PDG written text. Since arg.V2 can be assumed as mainly a phenomenon of spoken language (Auer 1998), this question is especially relevant from a language modality point of view.

This present study aims at adding to the theoretical field that is concerned with the mapping of pragmatical categories and syntactical (clause) form in earlier stages of German. By means of a conditional inference tree, argument clauses of ENHG narratives will be arranged according to the pragmatical features that have been considered to have an effect on their form in previous literature. Whereas assertion is often assumed to be the relevant notion to define the licensing requirements, theories on where the locus of the pragmatic assertion lies offer different perspectives to account for the V2 syntax in the argument clause: Seeing that it is generally matrix verbs denoting an act of assertion that embed arg.V2, mediated assertivity has been appointed the relevant pragmatic requirement for the V2 syntax (Reis 1997, Gärtner 2002). This term refers to the fact that the embedded proposition is mediated by a speaker who is not the originator of the assertion. In line with this, the possibility to mark a dependent clause with V2 in Germanic is often directly associated with the semantics of the matrix verb. Primarily, this goes back to Hooper & Thompson (1973). More recent approaches consider the reference level of the speaker/narrator (the mediator of the embedded assertion) as the relevant level on which – depending on the respective theory – the assertion is made (Meinunger 2006) or moved to in order to be marked as informatively relevant (Djärv 2022) to account for the V2 syntax. Note that from this point of view, the assessment of the distribution of arg.V2 on the basis of the matrix verb is only indirectly taking into account the pragmatic requirements of the dependent syntax.

Since both the "assertive levels" mentioned above come into question as being crucial for the licensing of arg.V2, they are both included into the analysis. Therefore, the analysis is not merely based on the matrix verb, but on the pragmatical features mentioned here.

The paper is organized as follows: In Section 2, the theoretical approaches are outlined in more detail and under consideration of some representative conceptions that will be fundamental for the further investigation. Section 3 moots the value of the historical approach to the topic and summarizes the state of research

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

on the diachrony of the arg.V2. Section 4 outlines the methodology, and in Section 5, the results of the corpus study are presented. In Sections 6 and 7, the results are summarized and discussed.

2 Theoretical approaches

As has been mentioned in the introduction, *assertion* is commonly appointed the relevant notion to account for arg.V2. In speech act theory, "[t]he point or purpose of the members of the assertive class is to commit the speaker (in varying degrees) to something's being the case, to the truth of the expressed proposition" (Searle 1975: 12). However, there is more than one possible level of an assertion that has to be taken into account with regard to argument realizing clauses. Consider (1b) (repeated here as (4))

(4) *Amanda sagte, sie* ***habe*** *Paul nicht gesehen*
Amanda said she has.AUX.SBJV Paul NEG seen
'Amanda said that she hadn't seen Paul.'

First, there is the level of the matrix subject *Amanda* that is the originator of the assertion, and second, there is the reference level of the speaker/narrator that functions, in any case, as the mediator of the assertion. Seeing that it is mainly verbs denoting an act of assertion that are able to select arg.V2, some researchers have drawn the conclusion that a *mediated* assertion yields the relevant driving force behind the V2 syntax (Reis 1997, Gärtner 2002). This notion refers to the fact that the asserted truth value of the dependent proposition is originally anchored at a matrix subject and is (only) mediated by the speaker. According to Reis (1997: 142), who coined the term *mediated assertion* ("vermittelte Assertion"), the declarative mode in arg.V2 complies with an "assertive restriction" that the matrix verb projects on its propositional complement. Correspondingly, the matrix verb sets an alternative mental world of the matrix subject (Reis 1997: 122). Reis concludes that arg.V2 are cases of non-structural theta-marking. Similarly, Gärtner (2002: 40) states that arg.V2 have "embedded assertional proto-force" that "can be 'absorbed' on arguments of predicates that denote acts of assertion". In line with these approaches, arg.V2 are commonly analyzed with regard to their selecting matrix verbs.

These approaches have frequently been adapted and further developed in the literature. From a more discourse-pragmatic perspective, it has been recognized that arg.V2 do not appear as flexibly under verbs of saying as the matrix-verb-approach would predict. For instance, Meinunger (2006) shows that V2 syntax

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

is unacceptable under direct repetition of a proposition. This holds true even if the proposition is selected by a verb of saying, which is illustrated in (5) (slightly adapted from Meinunger (2006: 7).

(5) A: *Bernd ist endlich gekommen*!
'Bernd has finally arrived!'
B:

a. *Ja, mir wurde schon gesagt, dass Bernd endlich gekommen ist.*
Yes me.DAT was already told COMP Bernd finally come.PST.PTCP is.AUX

b. **Ja, mir wurde schon gesagt, Bernd ist endlich gekommen*
Yes me.DAT was already told Bernd is.AUX finally come.PST.PTCP
'Yes, I have already been told that Bernd has finally arrived.'

It is restrictions like this that have led researchers to take on a more discourse-pragmatic approach as an explanatory framework of arg.V2 licensing. Meinunger (2006: 2) suggests, for instance, that V2 can optionally mark a dependent clause that the speaker "intends to assert", and furthermore, "these dependent verb second clauses have illocutionary force and mark new information" (:15). In reverse conclusion, Meinunger assumes discourse-old propositions to be obligatorily realized as syndetic VF-clauses (canonical dependent clauses), because "[t]hese belong to the background and are hence infelicitous candidates to make an assertion with" (ibid.).

Lohnstein & Staratschek (2020) present a somewhat weaker version of the speaker-assertion-hypothesis, claiming that "[t]he process of fronting finiteness has (...) exactly the same effect as in root clauses, with the exception that the speaker is not liable, but gives the information in the name of the matrix subject's referent" (Lohnstein & Staratschek 2020: 138). According to this, the illocutionary force is not on the side of the speaker, but the mediated proposition is still relevant with respect to the common ground between the speaker and the addressee. Similarly, Djärv (2022) hypothesizes that arg.V2 "represent discourse moves whereby the speaker adds the issue represented by the embedded clause to the conversational table as an issue for discussion".

Other researchers (to some degree in line with the speaker–assertion–hypothesis) assume the pragmatic distinction between the background and the

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

foreground of a discourse to be crucial for the syntactic variation of dependent clauses. Antomo (2022) argues that in German, V2 can only be applied to clauses that have so-called *at-issue* content, and that V2 is, reversely, excluded from all clauses that do not. Antomo aligns her conception of *at-issueness* with von Stutterheim & Klein (1992), Roberts (1996), Simons et al. (2010) who share the opinion that the primary goal of each discourse is to answer a *Question under Discussion* (in Roberts' words), which is typically implicit in a discourse. In line with this, Simons et al. (2010) state that all content that is relevant with regard to the current Question under Discussion is at-issue-content. Corresponding to this notion of at-issueness, Antomo (2022) states that all at-issue-content might possibly be marked by V2 in German.

Whatever pragmatic factor is taken to be the relevant requirement to account for arg.V2 does at the same time imply which factor is taken to be the 'blocking factor' of arg.V2. Meinunger (2006), who appoints discourse newness a relevant notion to account for arg.V2, makes it explicit by stating that discourse-old propositions are necessarily realized as VF clauses because "[t]hese belong to the background and are hence infelicitous candidates to make an assertion with" (Meinunger 2006: 15). Discourse-old information is sometimes defined as the epitome of the notion of *presupposition* (Lambrecht 1996). In Antomo's (2022) account, on the contrary, it is non-at-issue content that is diametrically opposed to at-issueness and that therefore blocks arg.V2. Once again, she refers to Simons et al. (2010), who state that non-at-issue content is information that is peripheral to the conversational goal. These different notions of "V2-blocking contexts" are not precisely congruent, but this does not entail that they have to be mutually exclusive. As Antomo (2022: 33) claims, "presuppositions form a subset of not-at-issue material". Since the aim of this study is to find out how different factors interact, it is important to keep the different conceptions of "V2 blockers" in mind.

3 A historical perspective

Some researchers assume that in historical German, even still in ENHG, syntactic variation is affected by discourse pragmatics to a higher degree than in PDG (Fleischmann 1973, Hinterhölzl & Petrova 2010, Speyer 2010). Hinterhölzl & Petrova (2010) model their theory within the frame of Segmented Discourse Representation Theory (SDRT) (Asher & Lascarides 2003). They argue that in OHG, an older system in which the verb functions as a means "to separate the aboutness topic plus other background elements from the focus domain of the clause"

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(Hinterhölzl & Petrova 2010: 5) is still clearly recognizable. According to them, V1-clauses served to introduce new discourse referents, whereas V2 clauses "serve to provide additional information about an already established discourse referent" (Hinterhölzl & Petrova 2010: 2). (This latter discourse relation is called *elaboration* in SDRT.) Speyer (2010) adds up to this and states that in the ENHG period, a shift happened to that system in that V2 clauses began to be used for purely "rhematic" sentences as well. In both these accounts, it is not explicitly considered how argument realizing clauses fit into the picture, but it implies that the VF-variant was historically reserved for clauses that are informationally completely backgrounded, and hence more restricted than their PDG structural equivalents. An account relevant for the situation in ENHG comes from Fleischmann (1973): He argues that the asymmetry of main- and dependent clauses is primarily reflective of pragmatical categories.

A hypothesis that is frequently discussed in the literature on the historical development of dependent verb second is that the syntactic variation between V2 and VF is attributable to a universal, hence diachronically robust, principle of the association between a pragmatic feature and a syntactic feature (amongst others, Auer 1998, Petrova 2020b,a). With regard to the "original" type of the PDG arg.V2, Axel-Tober (2012: 187) argues that already the earliest attestations of the arg.V2 in'Old High German seem to be restricted to matrix verbs denoting an assertion, which is indicative of a clause type that is historically old and diachronically unrelated to the verb final dependent clause (ibid).

As has been mentioned earlier, previous historical studies on the distribution of arg.V2 have primarily concentrated on the selecting matrix verbs. The arg.V2 is documented by Axel-Tober (2012) in the MHD period mainly with assertive verbs, which leads her to argue that the pragmatic licensing of arg.V2 might in fact be diachronically constant (:173). Petrova's (2020b, 2020a) data from Old and Middle High German support Axel-Tober's hypothesis, as the verbs that Petrova finds to predominantly select arg.V2 are verbs of saying and thinking, whereas factive-emotive verbs are only found with the VF-variant.

In fact, there is more than one possible origin scenario of the of the arg.V2. One would be that the original pattern was a clause pragmatically relatively independent of, and only loosely connected to the matrix clause. This scenario has been discussed in detail in previous literature. For example, Demske (2019) states that up until the sixteenth century, clauses of indirect speech that are only loosely connected to the speech markers in the preceding clause are frequently occurring in prose texts. She sees this type of "autonomous" indirect speech as a sign for a strong character perspective. The possibility that there has been a pattern

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

in which the dependent V2 clause was pragmatically more integrated can, however, not be excluded. The PDG V2-variant(s) of argument clauses might even go back to two diachronically different types of (more or less) dependent verb second: One that is indisputably integrated, and might historically be marked as integrated by non-syntactic means such as verbal mood (Gärtner & Eythórsson 2020, Coniglio et al. 2021), and one that is less integrated. From a syntactic perspective on PDG, Freywald (2016) suggests that there is a truly embedded type, on the one hand, and a (syntactically) unembedded type, on the other hand. According to her, this bipartition accounts for structural differences like a more fine grained structure of the left periphery of the unembedded type and its restriction to be placed to the right of the superordinate clause. Pragmatically, the coexistence of two types of arg.V2 could be reflected in the coexistence of arg.V2 that are attributed informative weight by the speaker/narrator, on the one hand, and assertions by a matrix subject that are only mediated, but are left unspecified in the course of the narration, on the other hand. Seeing that narratives "are multiperspectival since they prototypically integrate viewpoints of different characters and narrators and thus offer a set of possible alternate perspectives that allow for viewpoint switches" (Zeman 2020: 2), this assumed coexistence of arg.V2, that are driven by different pragmatic features, becomes even more plausible. I will come back to this scenario in more detail in Sections 4 to 6.

Summing up this section, it has to be noted critically that a focus on matrix verbs of arg.V2 provides a useful basis to analyze arg.V2 historically, but will only give us rather indirect insights into the pragmatic requirements of V2 (respective VF). Therefore, I am going to operationalize the pragmatic features of arg.V2 so far discussed to analyze the distribution of arg.V2 in ENHG narratives. How this is done is further explained in the following sections.

4 Methodological approach

4.1 Corpus

The data are collected from six narrative texts, five of which are part of the Romankorpus Frühneuhochdeutsch (Roko.UP) (Bloom et al. 2023) within the DFG-funded project *Wortstellung und Diskursstruktur in der Frühen Neuzeit* at the University of Potsdam. For the sixth text, the Huge Schapler, the Roko.UP provides a different edition than the one used here. The edition used here is an (incomplete) handwriting, which has been chosen for this study because it has a considerably higher incidence of indirect speech than the print which is provided by the Roko.UP. The texts that are used are given in Table 1. As can be seen, the

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

texts come from different narrative traditions, the Fortunatus being the only text that is an ENHG original. Schöne Magelone, Melusine, and Huge Schapler have a French template text, and Wigalois and Tristrant und Isalde are prose versions of Middle High German lyrical texts.

Table 1: Early New High German narratives

text	model text	dialect area
Huge Schapler (HUG)	French	Rhine Franconian
Schöne Magelone (MAG)	French	Upper German
Melusine (MEL)	French	Upper German
Tristrant u. Isalde (TRIST)	Middle High German	Upper German
Wigalois (WIG)	Middle High German	Upper German
Fortunatus (FORT)	no model text	Upper German

Prose texts lend themselves well for the investigation of the discussed pragmatical properties because one narrator tells a self-contained story, which makes the features discourse newness and at-issueness relatively easy to assess. How this is done is outlined in the section on the operationalization.

4.2 Preliminaries

To recapitulate: The study aims at shedding light on the pragmatic licensing conditions of arg.V2 in ENHG narratives. To extract the data, all argument realizing clauses have been subcategorized as either V2 (when they are unintroduced with the verb in the left periphery) or VF (when they are introduced by the complementizer *dass* 'that'). The implication of this is of course that in the Early New High German period, there is a bipartite pragmatic distinction between these two categories. This is based on the presumption that the complementizer *dass* was originally restricted to a specific (pragmatic) clause type (for a discussion of origin scenarios of this complementizer, see Axel-Tober (2012: 91–126). This procedure entails that syndetic clauses that have a post-verbal element will also fall into the VF-class.

My working hypothesis is that the distinction between VF and V2 clauses is pragmatically driven in ENHG. As has been mentioned, to test the hypothesis, I am choosing pragmatic categories as predictor variables. This is further explained in the next subsection. I am mainly focusing the study on verbs of saying as a semantic class where there is variation in the pragmatic function of the

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

argument clause, i.e., they are characterized by mediated assertivity by nature, but might be further influenced by informative weight. This has been explained in section 2.

To ensure that the main focus of the study is well-founded, I have conducted a preliminary step: I have collected data of argument clauses of verbs of saying on the one hand, and argument clauses of factive-emotive verbs on the other hand. This was to ensure that the "general distribution" of arg.V2 in ENHG is indeed similar to the distribution in PDG, meaning that arg.V2 are blocked under the same matrix verbs. If this is not the case, this would already imply a high flexibility of arg.V2 in ENHG that is not necessarily pragmatically restricted. Note that it has already been shown by Axel-Tober's (2012) and Petrova's (2020b) data that arg.V2 do not occur under factive-emotive matrix verbs in OHG and MHG, so it is highly unexpected that this drastically changes towards ENHG, and then changes back towards PDG.

After the extraction of argument clauses and the classification of their syntactic form, the clauses were annotated according to their matrix verbs, which I then grouped into different semantic classes (in line with, for example, Petrova 2020b). As the two semantic classes I am focusing on are factive-emotive verbs and verbs of saying, the grouping of their matrix verbs is exemplarily shown below.

1. Factive-emotive verbs, as they represent verbs that inherently presuppose the truth value of their dependent propositions and render them as discursively backgrounded. Verbs that fall under this class are for instance *bereuen* 'regret', *genießen* 'enjoy' or *wehtun* 'hurt'.

2. Verbs of saying, as representatives of verbs that denote an act of assertion, whereby the proposition can either be discursively foregrounded or backgrounded, and there might be varying degrees of speaker commitment. Verbs that fall under this class are *sagen* 'saying', *geloben* 'swearing', or *antworten* 'respond'.

The results of this survey show that the expectation holds true: We see striking syntactic variation under verbs of saying. Under factive verbs, on the contrary, only the VF variant can be accounted for, as Table (2) shows.

4.3 Operationalization

The second phase of the study focusses on argument clauses of verbs of saying in order to investigate their formal variation regarding their pragmatic properties. For this, the argument clauses have been annotated according to the concepts of

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Table 2: Distribution of arg.V2

	total	VF	V2
factive-emotive verbs	146	146	0 (0%)
verbs of saying	362	149	213 (58.9 %)

pragmatic licensing that I have referred to in Section 2. Argument clauses within sections of direct speech are not taken into account, as direct speech represents a different discourse situation.

In Section 2, I have referenced different hypotheses on arg.V2. I am categorizing these hypotheses on the basis of the (implicit) assumptions about the locus of the assertion. The strand of hypotheses that considers the assertion on matrix subject level as crucial to account for arg.V2, and that stands in distant operational relation to the matrix-verb-approach, I label mediated-assertivity-approach. The strand of hypotheses that considers the level of the speaker/narrator as relevant assertive level, I label relevance-approach. I have presented some conceptions that I consider to be representative for the relevance-approach, within which discourse newness (Meinunger 2006) and at-issueness (Antomo 2022) are two ideas with disagreeing assumptions: Whereas both refer to the informative relevance on discourse level, newness does not necessarily imply at-issueness, and vice versa. As the assumptions are not necessarily mutually exclusive, but might interact as influencing variables, the pragmatic licensing conditions can best be illustrated by means of a conditional inference tree, as this is a statistic model in which several effects can be considered simultaneously. The categories that are taken into account as being associated with arg.V2 are mediated assertivity, on the one hand, and informative relevance, on the other hand, with their different specifications. In the following, I will explain how I define and operationalize these specifications. The statistic model is further explained in Section 4.4.

Corresponding to the different informative-relevance-hypotheses, I consider two variables: Newness status (new vs old) and narrative relevance (progressive vs.non-progressive). Furthermore, I consider mediated assertivity as another variable. For newness status, I am appealing to a simple concept: If the truth value of the proposition in question can either be derived from the preceding story line, or has an explicit antecedent in the preceding text (by the same or a similar wording), it is labelled as "old" accordingly. Regarding the discourse antecedent, I do not differentiate between narrator-assertions and protagonist-assertions.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

For the category narrative relevance, I am appealing to the conception that propositions are characterized by varying degrees of informative relevance on the side of the speaker/narrator. What has been labelled as *speaker assertion*, *at-issue content*, and *discourse move* respectively - mostly for communication-based discourse models - must be operationalized for narrative texts. Therefore, I base my operationalization of this category on general principles of narrative coherence. I follow von Stutterheim & Klein (1992), who describe narrative texts as presentations of events that stand in a temporal relation to one another, but can also be connected by spatial or logical concepts. In von Stutterheim & Klein's (1992) sense, the superordinate question under discussion, that may be something like *What happened?* determines the main structure of the text. In conclusion, the juxtaposition of propositions that together form a (logical) progression of states and events can be understood as "main structure". Therefore, I am going to focus precisely on the discourse relation between the proposition expressed by the argument clause and the following proposition of the narrative segment: If the sequence in question can be interpreted in a way that the proposition in the arg.V2 is logically directly followed by the proposition in the next clause (i.e. they are connected by a consecutive relation), the label *progressive* is given. In such cases, the two propositions progressively contribute to answering the same question under discussion. By contrast, if the following clause connects to matrix clause or an earlier main clause, the arg.V2 is not understood as contributing to the same question under discussion, and as such is labelled "non-progressive".

Note that this procedure resembles the analytical criteria of Hinterhölzl & Petrova (2010) and Speyer (2010). However, I do not explicitly follow the discourse model of Asher & Lascarides (2003). This is because I do not agree with the assumption that a coherent story line unfolds on only one superordinate discourse level. In line with Langacker (2001: 177), I assume that "a discourse is not limited to the building of a single connected structure. Structures can be assembled and developed in distinct channels, worlds, or mental spaces, either successively or by shifting back and forth between them." Concretely, I assume that in narrative texts, the story line is not built on just one narrative level, but can yield new points of departure. This idea will become more clear in the discussion of the results.

The last pragmatic category is mediated assertivity. This receives no further label, as all propositions under verbs of saying are mediated assertions by nature. Note that this does not mean that it cannot be operationalized: If none of the named predictors come into question as licensing factors of arg.V2, mediated assertivity will be the last predictor variable that remains and will therefore be assumed to have explanatory power. Eventually, the category *text* will also be

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

taken into account as an predictor variable, in order to determine whether there is intertextual variation with regard to the distribution of arg.V2.

4.4 Modelling the data

I will illustrate the data on verbs of saying by means of a conditional inference tree, using the ctree function in R. The conditional inference tree algorithm produces a tree that can be seen as a visualization of how predictors combine in their effect size, for a specific data set. The tree uses recursive partitioning to estimate the effect of a predictor variable while the other variables are simultaneously taken into account. The significance of the effect of each variable is calculated based on permutation. The statement that each of the branches makes can be formulated as "if…then…"-statements.

5 Pragmatical distribution of arg.V2 under verbs of saying

Figure (1) shows the distribution of arg.V2 under verbs of saying by means of a conditional inference tree.

The tree gives out the variable "text" as the variable with the highest association with the response variable (the form of the argument clause). The texts are split into two subsets, one of which consists of the Melusine and the Wigalois, which have a relatively high proportion of verb-final argument clauses. In fact, all clauses that are labelled old and non-progressive simultaneously (11 out of 42 cases) are realized with VF. But as the proportion of VF clauses is relatively high in general, these hypothesized influence variables make up for no significant effect and hence, this node is not bisected any further.

For the second set of texts, consisting of the Fortunatus, Huge Schapler, Schöne Magelone, and Tristrant und Isalde, the tree gives out newness status as the category where the next best association is found. The data are sectioned into new and old propositions, according to the fact that in the group of old propositions, the frequency of arg.V2 is lower than in the set of new propositions. However, discourse-old status alone does not uniformly block V2. This can be derived from node 7 in the tree, where the response is once again sectioned according to discourse progression. This means that discourse progression plays a decisive role in predicting the form of an argument clause in the set of old propositions. If the proposition is progressive, it is very likely to be realized with V2, even though it is disourse-old at the same time. This is exemplified in (6).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

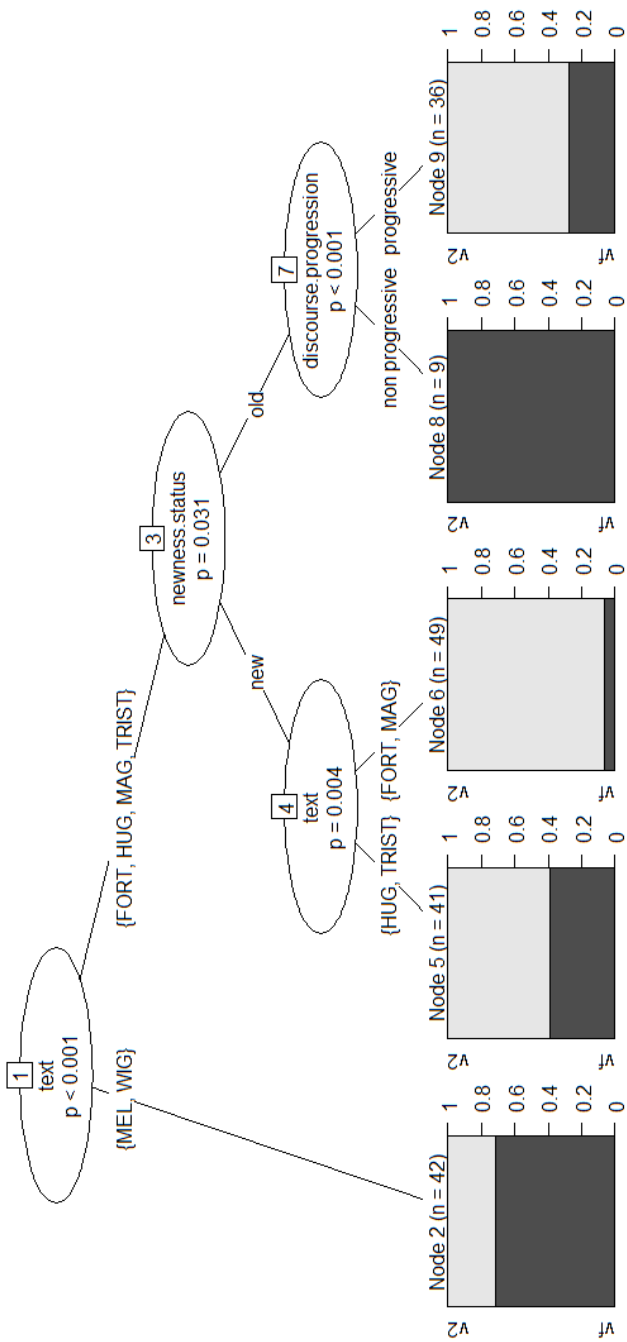


Figure 1: arg.V2 under verbs of saying

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(6) Context: After a battle, Tristrant is critically wounded. His ship drifts towards the coast of Ireland. There it is discovered by the king, who asks his servants to find out what is in the ship.

dye diener kamen vnd sagten. da wär ein man wundt in den
The servants came and said there was a man wounded in the
tod. der küng gieng selbs dar vnd fand als jm gesagt was
death. The king went himself there and found as him told was
'The servants came and said that there was a man who was critically wounded. The king went there himself and discovered what he had been told.' [TRIST, 19r]

According to the proposition having a discourse antecedent that positively decides its truth value, the proposition under question in (6) is old. At the same time, the proposition proceeds the story line, in that it contributes in causally bringing forth the event expressed in the following clause. Therefore, it is progressive. The perspectivation of the expressed state of affairs seems to play a decisive role for the "main clause character" of the arg.V2: In this case, the reader is following the stream of consciousness of the king, who inquires about a matter. For him, the state of affairs (that there is a wounded man in the ship) is new, informative and decisive for his further action. In this case, the V2 syntax might be attributable to the fact that the narrator "takes on the perspective of protagonist" in order to progress the story line in a coherent way. To contrast this, (7) shows a sentence in which the proposition in question is discourse-old and non-progressive (node 8 in Figure 1).

(7) Context: Rupert, a servant of the earl of Flanders, has deliberately convinced Fortunatus to leave the earl's court without his knowledge. The earl is now devastated about Fortunatus' disappearance. Rupert notices the earl's frustration and is frightened that one of his companions might reveal that he is the causer of Fortunatus' disappearance.

Und gieng zu yn allen/ und yedem in sonderhait/ bat sy/ das
and went to them all and everyone in particular asked them that
sy nyendert meldeten/ daz er ain ursacher wär/ sines hinwegschaidens/
they nobody told that he a causer was his disappearance
das gelobten sy im gar trüwlich
this promised they him very faithfully
'and (Rupert) went to all of them individually and asked them not to tell anybody that he was a causer of his (Fortunatus') disappearance. This they promised him faithfully.' [FORT, 403]

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

First note that in (7), the matrix verb of the proposition in question is negated. The truth value of a proposition is not affected by the scope of the negation in the matrix clause, which is commonly assumed to be a sign that the proposition is presupposed. This can be explained in discourse-pragmatic terms with regard to the two criteria under consideration: The proposition is old and non-progressive. The proposition is already incorporated in the reader's mental system, so for the reader, its truth value is decided. A privotal factor here is that the proposition is also already incorporated in the mental state of the protagonists of the following proposition (the companions). The action that they take does not consecutively follow from the proposition in question. Instead, the consecutive relation lies between a preceding proposition (Rupert asking his companions to do him a favor) and the clause following the arg.V2 (his companions agreeing). Note that when the proposition is old and non-progressive at the same time, it is realized as a VF clause in nine out of nine cases, which is highly significant with relation to the group of discourse old, progressive propositions, according to the model.

The effect of the discourse progression status on the form of the argument clause gets lost in the subset of discourse-new propositions. This corroborates the hypothesis that there is another licensing factor for arg.V2 besides progressiveness, which I assume to be mediated assertivity. In other words, mediated assertivity does yield V2 under the premise that the proposition is new in the discourse. This becomes especially clear if we look at a clause where the proposition in the argument clause is non-progressive, as it is the case in 7).

(8) Context: Magelone, who works in a hospital, is visited by a count and countess.
vnd baten die sie solt mit jnen zů morgen essen/ aber jr hertz
and asked her she shall with them to morning dine but her heart
vermocht jhn solchs nit zů zesagen/ darumb sie sprach sie hette
could them this NEG PRTCL to.promise therefore she said she had
daheim zů thůn (...) Nach dem kam die Magelona wider zů dem Peter
at.home to do (...) After this came the Magelona again to the Peter
'And (they) asked her to eat with them in the morning, but her heart could not promise that, therefore she told them she had something to do at home (...). After this, Magelone returned to Peter.' [MAG, 674]

In (8), the clause under question is obviously non-progressive, as it does not have any kind of consecutive discourse relation with the proposition following it. The mediated proposition might be new and informative for the characters involved in the described situation (the count and countess), but the reader does not

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

follow their stream of consciousness. Instead, the story line continues to build around Magelone (the originator of the proposition in the arg.V2) and for the stream of the narrated events, the proposition in question does not play a decisive role. In other words, there is no consecutive relation between what Magelone makes up (the proposition in the argument clause), and the proposition following it. Furthermore, what is important to note is that the data do not imply that the correlation between the newness status and the clause form is a spurious correlation. Instead, it rather suggests that newness status individually effects the form of the argument clause, to a somewhat lower degree than it does in combination with (non)-progressiveness. In the Fortunatus, for example, all instances of discourse-new propositions (28 in total) are marked by V2, but discourse-old propositions that are progressive have VF syntax in three out of 16 cases. Consider (9).

(9) Context: The count introduces Fortunatus to other people.
und sagt yn wie er so ain gůter jäger wäre/ die vogel in dem
and told them how he such a good hunter was.SUJV the birds in the
lufft und die thyer in den wälden wär kaines sicher vor ym/
air and the animals in the woods were.SBJV none safe in.front him
tzu dem das er sunst wol dienen kund
to this COMP he else well serve could
'And told them how he was such a good hunter, none of the birds in the air and the animals in the forest were safe from him, furthermore, that he could serve well.' (FORT, 394)

In (9), what is special with regard to the distribution shown in Figure 1 is that there is a formal variation between the coordinated dependent clauses in 8), although they all (as a configuration) bear the same discourse relation to the following event (consecutive). Nevertheless, only the discourse old (and positively decided) ones have VF syntax, whereas the one that is new (in that a state of affairs is mediated in an "embellished" way by using the subject referent's words) has V2 syntax.

6 Discussion

The findings of the present study can be summarized as follows: The mapping between pragmatics and syntax is not equally systematic in all of the invested texts. The Melusine and the Wigalois both have a relatively high frequency of

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

syndetic VF argument clauses, so that neither the newness status nor the discourse progression status of a proposition make up for a significant effect on the form of the clause. VF seems to be the default form of argument clauses here. I suspect this tendency to become more dominant diachronically towards PDG. This, however, needs further investigation.

In a larger group of texts, discourse-old status of a proposition ("old" within the mental state of the reader) reduces the likelihood for the proposition to be realized with arg.V2, but does not exclude the possibility. The discourse progression status seems decisive here: If a discourse-old proposition is at the same time progressive, argument clauses are still frequently realized with V2. On the contrary, if an old proposition is non-progressive, i.e., not decisive for the stream of events, VF is the only formal variant that is used. Hence, it can be seen in the category of old propositions that the discourse progression status is a decisive predictor variable with respect to the form of an argument clause. VF seems to be rather restricted. Only category of propositions that are discourse old and non-progressive seems to be rather (if not exclusively) restricted to VF syntax. Therefore, I propose the interaction of discourse old status and non-progressive status to be relevant for defining the blocking context of arg.V2 in ENHG narratives. Furthermore, it has also been shown that within the subset of discourse-new propositions, the progression status of a proposition has no significant effect on the form of the argument clause. In how far can this be linked to the different levels of assertion that have been discussed earlier? If in a narration, a proposition is consecutively followed by another, it is informatively heavier than when this is not the case, because it becomes emphasized on narrative level. We can say that the embedded proposition is "shifted" to the narrative level. This does not necessarily entail that it is an assertion on the narrator's side, but is assigned narrative relevance by the narrator. However, I have shown that propositions that are mediated, but left unspecified in the progression of events (i. e., have no or weaker narrative relevance), are also frequently realized with V2 syntax in ENHG narratives. As these clauses lie on a deeper narrative level, they can be classified as pragmatically embedded. How concordant this pragmatically embedded class is to the class of "true V2-complements" (Freywald 2016) cannot be derived with certainty. However, there are signs that V2-clauses can be embedded. Consider the argument clause in (7) again: Here, the finite verb is used in subjunctive form, although the proposition's truth value is clearly decided. This would be marked in PDG, where subjunctive mood is used to relativize truth claims (Fabricius-Hansen et al. 2019). This is a sign that the usage of subjunctive mood in ENHG still differs from that in PDG: In ENHG, subjunctive verbal mood can function as a means to mark syntactical dependence, as it is often proposed

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

to be the case in older stages of German (Schrodt 2004). While the finite verbs of arg.V2 are frequently used in subjunctive form in the ENHG narratives, there is also a considerable amount of subjunctive verbs in arg.V2. As subjunctive verbal mood can function as a marker of clausal dependency, it can be derived that the varying use of subjunctive is (at least occasionally) attributable to the syntactic (in)dependence of the respective clause. Consider (10) as a minimal pair.

(10) a. *sagten sy er wär der grost boswicht der vff ertrich lebt*
said they he was.SBJV the biggest villain that on earth lives
'They said he was the greatest villain to walk the earth.' [FORT, 499]

b. *Zů dem Anndolosia sagt ia er wolt ym geren dienen mit*
To which Andolosia said yes he wanted him willingly serve with
leib vnd mit gůt
body and with good
'To which Andolosia responded, yes, he wanted to serve him willingly, with body and good' [FORT, 515]

At this point, it is also important to note that in some - if not many - cases of arg.V2, there can be ambiguities with regard to the hypothesized progressive status. Hence, in written text, there might be a misconception of narrative relevance/irrelevance. As I said in the beginning of the discussion, I suspect the tendency of VF being the default form of argument clauses to become more dominant diachronically. The hypothesized ambiguity between narrative relevance and narrative irrelevance, as well as the change in the use of subjunctive mood (which is still ongoing in ENHG) might be a crucial factor for this change. Therefore, what will be highly relevant for upcoming studies is the question how exactly PDG proportions have prevailed, and which role the interaction of sentence mood and verb mood plays diachronically, form ENHG onwards. Whether discourse-old status becomes a stronger differentiating factor for the syntactic form of argument clauses (as Meinunger 2006 proposes it for PDG) is an open question as well. These questions on the motivations for changes are especially relevant from a language modality perspective: Auer (1998) finds 60 percent "dependent main clauses" in spoken language and 35 percent in written language. These numbers suggest that it is precisely the written variety that has "drifted away" from the original proportions. Although spoken language has often been argued to have a stronger progressive force than the written variety, literacy might drive forward linguistic innovations in the field of syntax through its inherently low degree of situational embedding, as Koch & Oesterreicher (1996)

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

propose. The relevance of ENHG narrative prose for the assessment of linguistic changes within the written language is especially justified against the background that the importance of this genre increased at that time, which is reflected in the rapid growth of the number of prose novels as prints and manuscripts. Hence, the scientific focus on the narrative genre can provide more insight not only into linguistic circumstances at this point in time, but also into what factors are crucial for the development of linguistic phenomena in the written language.

7 Summary and conclusion

In previous literature, different pragmatic factors have been discussed with respect to their relevance as licensing requirements for arg.V2. The main finding of the present study is that the inventory of arg.V2 in ENHG is best understood as the outcome of an interaction of factors in ENHG prose texts. Hence, the theoretical approaches on the licensing requirements do not necessarily contradict one another. It has been shown that V2 comes into effect not only if the assertion is assigned narrative relevance by progressing the story line, but also if an assertion is just mediated without being further specified on narrative level. In some texts, V2 is used considerably frequently in both these cases. Only propositions that are discourse-old and non-progressive at the same time seem to clearly block V2 syntax.

8 Abbreviations

V2	asyndetic verb second		
VF	syndetic verb final	ENHG	Early New High German
Arg.V2	argument realizing verb second clause	PDG	present-day German

Acknowledgements

This research was funded by the Deutsche Forschungsgemeinschaft (Project 456973946, „Wortstellung und Diskursstruktur in der Frühen Neuzeit“).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Primary sources

Die Schöne Magelone (1535). Müller, Jan-Dirk (ed.). 1990. Romane des 15. und 16. Jahrhunderts. Nach den Erstdrucken mit sämtlichen Holzschnitten, 587–678. Frankfurt am Main: Deutscher Klassiker Verlag.

Fortunatus (1509). Schmitt, Ludwig Erich and Renate Noll-Wiemann (eds.) 1974. Fortunatus: Von Fortunato und seynem Seckel auch Wünschhütlein. Hildesheim: Olms.

Huge Scheppel (1500). Huge Scheppel; Königin Sibille Saarbrücken. 1455-1456. 76 pages. Hamburg: Staats-und Universtitätsbibliotek. Codex 12 in scrinio. pp. 1ra-57vb. Fascimile available online: https://resolver.sub.uni-hamburg.de/kitodo/HANSh495

Melusine (1474). Müller, Jan-Dirk (ed.). 1990. Romane des 15. und 16. Jahrhunderts. Nach den Erstdrucken mit sämtlichen Holzschnitten, 9–176. Frankfurt am Main: Deutscher Klassiker Verlag.

Tristrant und Isalde (1484). Elsner, Helga (ed.). 1989. Tristan und Isolde (Augsburg bei Antonius Sorg, 1484). Hildesheim: Olms.

Wigalois vom Rade (1519). Melzer, Helmut (ed.). 1973. Wigalois. Hildesheim: Olms.

References

Antomo, Mailin. 2022. Marking (not-) at-issue content by using verb-order variation in German. In Ingo Reich & Augustin Speyer (eds.), *Co- and subordination in German (Sonderheft der Linguistischen Berichte 21)*, 21–54. Hamburg: Buske.

Asher, Nicolas & Alex Lascarides. 2003. *Logics of conversation*. New York: Cambridge University Press.

Auer, Peter. 1998. Zwischen Parataxe und Hypotaxe: ‚Abhängige Hauptsätze' im gesprochenen und geschriebenen Deutsch. *Zeitschrift für Germanistische Linguistik* 26(3). 284–307. DOI: 10.1515/zfgl.1998.26.3.284.

Axel-Tober, Katrin. 2012. *(Nicht-)kanonische Nebensätze im Deutschen*. Berlin: De Gruyter.

Betten, Anne. 2000. Zum Verhältnis von geschriebener und gesprochener Sprache im Frühneuhochdeutschen. In Werner Besch, Anne Betten, Oskar Reichmann & Stefan Sonderegger (eds.), *Sprachgeschichte. Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung*, vol. 1, 1646–1664. Berlin: De Gruyter.

Bloom, Barthe, Malika Reetz & Ulrike Demske. 2023. *Romankorpus Frühneuhochdeutsch (Roko.UP)*. https://zenodo.org/records/10210706.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Coniglio, Marco, Chiara De Bastiani, Roland Hinterhölzl & Thomas Weskott. 2021. In the right mood, in the right place. *Journal of Historical Syntax* 5. 1–27. DOI: 10.18148/hs/2021.v5i1-13.43.

Demske, Ulrike. 2019. Zur Autonomie indirekter Redewiedergabe – Eine diachrone Perspektive. *Zeitschrift für Germanistische Linguistik* 47(1). 70–101. DOI: 10.1515/zgl-2019-0003.

Djärv, Kajsa. 2022. On the interpretation and distribution of embedded main clause syntax: New perspectives on complex discourse moves. *Glossa: A journal of general linguistics* 7(1). DOI: 10.16995/glossa.5752.

Fabricius-Hansen, Cathrine, Kåre Solfjeld & Anneliese Pitz. 2019. Der Konjunktiv. Formen und Spielräume. *Zeitschrift für Sprachwissenschaft* 38(2). 305–309.

Fleischmann, Klaus. 1973. *Verbstellung und Relieftheorie: Ein Versuch zur Geschichte des deutschen Nebensatzes*. München: Fink.

Freywald, Ulrike. 2016. Clause integration and verb position in German. Drawing the boundary between subordinating clause linkers and their paratactic homonyms. In Ingo Reich & Augustin Speyer (eds.), *Co- and subordination in German (Sonderheft der Linguistischen Berichte 21)*, 181–220. Hamburg: Buske.

Gärtner, Hans-Martin. 2002. On the force of V2 declaratives. *Theoretical Linguistics* 28(1). 33–42. DOI: 10.1515/thli.2002.28.1.33.

Gärtner, Hans-Martin & Thórhallur Eythórsson. 2020. Varieties of dependent Verb Second and verbal mood: A View from Icelandic. In Rebecca Woods & Sam Wolfe (eds.), *Rethinking verb second*, 208–239. Oxford: Oxford Scholarship Online.

Hinterhölzl, Roland & Svetlana Petrova. 2010. From V1 to V2 in West Germanic. *Lingua* 120(2). 315–328. DOI: 10.1016/j.lingua.2008.10.007.

Hooper, Joan B. & Sandra A. Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4(4). 465–497.

Koch, Peter & Wulf Oesterreicher. 1996. Sprachwandel und expressive Mündlichkeit. *Zeitschrift für Literaturwissenschaft und Linguistik* 26(2). 64–96.

Lambrecht, Knud. 1996. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press.

Langacker, Ronald W. 2001. Discourse in cognitive grammar. *Cognitive Linguistics* 12(2). 143–188. DOI: 10.1515/cogl.12.2.143.

Lohnstein, Horst & Nathalie Staratschek. 2020. The left periphery as interface – On verb second and finiteness interaction. In Horst Lohnstein & Antonios Tsiknakis (eds.), *Verb second: Grammar internal and grammar external interfaces (interface explorations 34)*, 207–249. Berlin: De Gruyter.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Meinunger, André. 2006. On the discourse impact of subordinate clauses. In Valeria Molnár & Susanne Winkler (eds.), *The architecture of focus*, 459–487. Berlin: De Gruyter.

Petrova, Svetlana. 2020a. Embedded Verb Second in the history of German. In Rebecca Woods & Sam Wolfe (eds.), *Rethinking verb second*, 208–239. Oxford: Oxford Scholarship Online.

Petrova, Svetlana. 2020b. Variation and change in the licensing of dependent V2 in German. In Horst Lohnstein & Antonios Tsiknakis (eds.), *V2 in Germanic and Romance*, 251–276. Berlin: De Gruyter.

von Polenz, Peter. 2000. *Deutsche Sprachgeschichte vom Spämittelalter bis zur Gegenwart. Bd. 1: Einführung, Grundbegriffe: 14. bis 16. Jahrhundert*, vol. 2. Berlin: De Gruyter. DOI: 10.1515/9783110824889.

Reis, Marga. 1997. Zum syntaktischen Status unselbstständiger Verbzweit-Sätze. In Christa Dürscheid, Karl Heinz Ramers & Monica Schwarz (eds.), *Sprache im Fokus. Festschrift für Heinz Vater zum 65. Geburtstag*, 121–144. Tübingen: Niemeyer.

Roberts, Craige. 1996. Information structure in discourse: Toward a unified theory of formal pragmatics. *Semantics & Pragmatics* 5(6). 1–69.

Schrodt, Richard. 2004. *Althochdeutsche Grammatik II: Syntax*. Tübingen: Niemeyer.

Searle, John R. 1975. A taxonomy of illocutionary acts. *Language, Mind, and Knowledge. Minnesota Studies in the Philosophy of Science* 7. 344–369.

Simons, Mandy, David Beaver, Judith Tonhauser & Craige Roberts. 2010. What projects and why. *Proceedings of SALT* 20. 309–327. DOI: 10.3765/salt.v0i20.2584.

Speyer, Augustin. 2010. Die Markierung von Diskursrelationen im Frühneuhochdeutschen. *Sprachwissenschaft* 35(4). 409–442.

von Stutterheim, Christiane & Wolfgang Klein. 1992. Textstruktur und referentielle Bewegung. *Zeitschrift für Literaturwissenschaft und Linguistik* 86. 67–92. DOI: 10.1007/978-3-476-05420-3_7.

Zeman, Sonja. 2020. Parameters of narrative perspectivization: The narrator. *Open Library of Humanities* 6(2). 1–33. DOI: 10.16995/olh.502.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Chapter 8

Verb-initial declaratives in Old English prose: a functional analysis

Anna Cichosz[a]
[a]University of Łódź

The aim of this study is to conduct a comprehensive, corpus-based analysis of verb-initial main declarative clauses in Old English prose records in order to determine their discourse functions and to establish what linguistic and/or extralinguistic factors may be responsible for the highly uneven distribution of the structure. The analysis shows that Old English verb-initial declaratives perform two main functions: narrative and descriptive, with the latter restricted to just a handful of early compositions, and the former attested as a low-frequency item in the vast majority of texts. Both patterns are analysed in context and it turns out that verb-initial declaratives with and without an overt subject share many properties and functions, and that verb-initial declaratives with personal pronoun subjects are rare, which has some important implications for the interpretation of the results.

keywords: Old English, word order, syntax, verb-initial, corpus study

1 Introduction

Verb-initial (V1) main declarative clauses with an overt subject are a relatively well-attested structure present in numerous Germanic languages (Walkden 2014: 94; Ringe & Taylor 2014: 408). In the Old Germanic context, V1 was quite frequent in negated declaratives (Hopper 1975), whereas non-negated V1 declaratives were a marked pattern (Lehmann 2005), usually associated with some sort of emphasis or "dramatic force" (Smith 1971: 96). The special function of V1 is directly reflected in the term "narrative inversion", which is usually equated with a non-negated V1 declarative containing an overt subject, with the underlying assumption that the motivation behind the pattern was always to push the narration forward.

Anna Cichosz. 2025. Verb-initial declaratives in Old English prose: a functional analysis. In Ulrike Demske & Barthe Bloom (eds.), *Discourse structure and narration: A diachronic view from Germanic*, 233–262. Berlin: Language Science Press. DOI: ??

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

This study confirms the initial findings of Cichosz (2022), proving that in Old English (OE) this assumption is not fully confirmed by the data and V1 cannot be exclusively associated with a narrative function because there is an alternative structure, relatively frequent in early prose texts, where V1 declaratives are used to provide additional information about a known referent.

As far as other Old Germanic languages are concerned, V1 declaratives are attested e.g. in Old High German (OHG), where they are reported to appear in presentational and existential structures, with mutative verbs (i.e. verbs of change, including verbs of movement), passive or passive-like constructions, verbs of saying, and negated clauses (Axel 2007). Their frequency in OHG is quite robust (with the maximum at 14% in the OHG Tatian Gospel Harmony, cf. Cichosz et al. 2016: 127), while in Old Icelandic (OI) they constitute 18% of all main declaratives (Booth 2018). Most of these are classified as narrative (Booth 2018: 107) and narrative inversion in the context of OI is defined as "a particularly common type of V1" (Booth & Beck 2021: 41) alongside presentational structures, impersonal constructions and clauses without an overt subject. By contrast, in Old English (OE) the frequency of V1 declaratives is generally on the low side (ca. 2% in non-negated main clauses, cf. Cichosz 2020: 365), and the structure, if used, is reported to introduce a new storyline or mark some sort of transition in the narrative structure of the text (Mitchell 1985: §3933; Los 2000; Ohkado 2004; Calle-Martín & Miranda-García 2010). Its use is quite regular in OE poetry but the pattern is very unevenly distributed in the prose records, with the frequency of V1 declaratives relatively high in just a handful of OE prose texts, e.g. Bede's *Historia Ecclesiastica* (Ohkado 2000, Calle-Martín & Miranda-García 2010, Cichosz 2017b).

Thanks to the availability of the syntactically annotated York-Toronto-Helsinki Parsed Corpus of Old English Prose (YCOE, Taylor et al. 2003), it is possible to conduct detailed quantitative and qualitative studies of syntactic structures, including narrative inversion. The quantitative approach, however, faces a serious problem in terms of the adequate operationalisation of the structure. As signalled in Cichosz (2022), there are two parallel constructions taking exactly the same form in the OE data, illustrated by (1) and (2):

(1) *Wæs þa sum Godes þegen binnon þære byri*
was then some.NOM God.GEN nobleman.NOM in the.DAT city.DAT
'There was a man of God in the city.' [ÆCHom I, 27: 401.23-25]

(2) *Hæfde þæt deor þrie hornas on foran heafde*
had the.NOM animal.NOM three.ACC horns.ACC on forehead.DAT
'The animal had three horns on its forehead.' [coalex, Alex: 20.2.229]

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(1) is a classic example of narrative inversion: the clause introduces a new participant and it is used at the beginning of a story. (2), however, is quite different, providing additional information about a known referent without pushing the narration forward in any way. In Cichosz (2022), where samples from 4 different prose texts with the highest frequency of V1 declaratives are analysed, it turns out that narrative inversion is used by Ælfric (though not consistently in all of his works), but other prose texts (Bede, Vercelli Homilies and Alexander's Letter) opt for the construction shown in (2). Thus, it seems that in OE V1 declaratives cannot be equated with narrative inversion as at least some of them perform a drastically different discourse function. It is not uncommon in language for the same form to have distinct meanings or to perform different functions, but such a situation is easily missed if only aggregate numbers are taken into account. In this case, both structures involve exactly the same corpus annotation: the clause is main, non-conjunct and non-negated, the verb is indicative, and there is an overt subject in the clause. It is technically impossible to separate one from the other in any automated way: manual analysis in context is an indispensable component of the study.

The aim of this investigation is to conduct a comprehensive quantitative and qualitative analysis of all V1 declaratives in the YCOE corpus in order to establish the actual frequency of both narrative and descriptive V1 clauses in OE prose (and, if possible, to identify other functions of V1 declaratives), seeking the main reasons for the uneven distribution of the patterns in YCOE. On a more general level, the study shows the importance of qualitative, context-based analyses, which can and should complement any data-driven corpus studies.

2 V1 declaratives in Old English

The functions of V1 declaratives in OE are usually quite broadly defined as marking some sort of transition in the narrative structure of the text (Mitchell 1985: §3933; Calle-Martín & Miranda-García 2010), which is not necessarily as exceptional as the reported markedness of the pattern would suggest. OE had a large set of clause-initial discourse-sequencing adverbs, mostly *þa* 'then' and *þonne* 'then', which could and did mark transition, and *þa* was the most frequent and neutral device used for this function in OE prose (van Kemenade & Los 2006). It has been noted that narrative inversion tends to co-occur with particularly vivid or dramatic fragments of the narration (Stockwell 1977: 291; Kiparsky 1995: 163; Los 2000), though Petrova (2006: 168) reports that in Beowulf V1 clauses "occur exactly in cases where the plot enters a new stage of development or the nar-

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

ration needs to be pushed forward" and according to her study, V1 and *þa*-VS clauses (at least in OE poetry) are functionally interchangeable.

Nonetheless, existing research on narrative inversion in prose still treats V1 declaratives as a special discourse device and Ohkado (2004) is most explicit in his classification of all the functions fulfilled by the structure. His analysis of V1 declaratives used in Ælfric's Catholic Homilies I (30 clauses in total) reveals that they mark transition to a new action, summarise a discussion, introduce a new or contrasting character or type and open a new story or paragraph. Mitchell (1985: §3930) mentions that V1 declaratives are often presentational constructions (which would probably correspond to Ohkado's (2004) "new character" function) and Allen (1995: 34) suggests that V1 serves to "emphasise the (new information) subject". The last point is quite problematic since V1 clauses are known to employ personal pronoun subjects (Ringe & Taylor 2014: 408), and such constituents by definition cannot introduce new discourse referents, but Petrova (2006: 171) describes the function of V1 as "focusing not only the subject but the entire proposition". Los (2000) finds a difference between the two on the basis of her analysis of Ælfric's works: in her study both V1 and *þa*-VS clauses occur at episode boundaries but the former indicate "a thematic discontinuity", while the latter mark "the smooth flow of narrative" (Los 2000: 263).

As mentioned in Section 1, more recently it has been shown that V1 declaratives attested in OE prose texts have more functions, and not all of them may reasonably be covered by the umbrella term "narrative inversion" (Cichosz 2022). Attested patterns include both narration-progressing clauses as in (3) and descriptive structures giving additional information about previously introduced discourse referents shown in (4). Interestingly, the patterns are reported to show drastically different lexical preferences: while the former tend to co-occur with dynamic and punctual verbs (as in *wæs cumen* 'was come' in (3)), the latter is associated with verbs of state (*wesan* 'to be' or *habban* 'to have' as in (4)).

(3) *Þa foron forð oþ hi comon to Lundenbyrig & ða mid*
then went forth until they.NOM came to London.DAT and then with
þam burhwarum & mid þam fultume ðe him westan
the.DAT citizens.DAT and with the.DAT help.DAT that they.DAT from.west
com foran east to Beamfleote; ***Wæs ða Hæsten þær cumen mid***
came went east to Benfleet was then Hasten.NOM there come with
his herge þe ær æt Middeltune sæt
his.DAT army.DAT which before at Middletown.DAT sat
'Then they went forth until they reached London and then they went east to Benfleet with the citizens and with the reinforcements that came to

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

them from the west. **Then Hasten came there with his army, which had previously stayed at Middletown**'.
[ChronC (O'Brien_O'Keeffe): [038400 (894.35)]]

(4) *Þa eode he, ure dryhten Crist, ut beforan þa*
then went he.NOM our.NOM lord.NOM Christ.NOM out before the.ACC
Iudeas. ***Hæfde he þaþyrnenne coronan on his***
Jews.ACC had he.NOM the.ACCthorny.ACC crown.ACC on his.DAT
heafde, *& mid þam readan hrægle gegyred wæs. Þa cwæð*
head.DAT and with the.DAT red.DAT gown.DAT clothed was then said
he, Pilatus, to Iudeum:Ecce homo
he.NOM Pilate.NOM to Jews.DAT [Latin]
'Then he, our Lord Christ, went out before the Jews. **He had a crown of thorns on his head** and he was wearing a red robe. Then he, Pontius Pilate, said to the Jews: *Ecce homo*.'
[HomS_24_ (ScraggVerc_1): [009500 (159)]-[009700 (161)]]

This means that the same form (a non-negated V1 declarative clause with an overt subject) may represent both coordinating and subordinating discourse relations, as defined in the Segmented Discourse Representation Theory, where clauses are classified as coordinating if they belong to the main line of narrative, introducing main events and pushing the story forward, and subordinating when they are placed in a substructure of the text, presenting background, providing additional information, comments or explanations (Asher & Lascarides 2003; Asher & Vieu 2005). (4) is a clear example of elaboration or background, while (3) operates on the highest level of text hierarchy, presenting one of subsequent events in the story. In the OHG context, Hinterhölzl & Petrova (2005: 73) determine that "verb-initial structures establish coordinative discourse relations whereas verb-second clauses signal subordinating linkage to previous discourse part", which is in line with the conclusions reached by Los (2000) for OE, but examples such as (4) clearly diverge from this pattern as V1 declaratives are not consistently found on the main level of the narration.[1]

One of the main limitations of various functional studies on OE narrative inversion is that they are predominantly based only on the output of Ælfric of

[1]An example of a study applying the Segmented Discourse Representation Theory to the analysis of OE is Bech (2012), who concludes that there is no straightforward relation between a specific constituent order pattern (with the focus of the study on V-final and V-late clauses) and discourse function in OE. Thus, at least in OE, the situation seems quite complex, not only in the context of V1 declaratives.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Eynsham, a famous OE writer and translator, who is single-handedly responsible for ca. 25% of OE prose records (Cichosz & Pęzik 2021). What is more, Ælfric (c. 955–1010) was active at the very end of the OE period so his output, probably full of idiosyncrasies of his style, represents only the later part of period. As a result, such studies do not cover early OE, which is the most probably reason why the high frequency of the structure represented by (4) has gone largely unnoticed. Moreover, the analyses are usually based on relatively small datasets, e.g. Ohkado (2004) analyses 30 V1 clauses from Ælfric's Catholic Homilies, Los (2000) analyses variation between V1 and *þa*-VS in clauses with the verb *onginnan/beginnan* 'to begin' in all works by Ælfric (11 V1 clauses vs. 35 *þa*-VS clauses), and Cichosz (2022) takes samples of V1 clauses with 3 most frequent verbal collocates from 4 different YCOE texts (57 clauses in the qualitative part of the analysis). As a result, our understanding of the functional load of V1 declaratives and the division of labour between narrative V1 and other OE discourse devices is fragmentary and calls for a comprehensive analysis.

Finally, it should be noted that there are considerable differences in the distribution of V1 declaratives in the OE textual records. The structure is relatively rare in prose but frequent in poetry (while *þa*-VS is frequent in prose and rare in poetry), and there is one prose text where the frequency of V1 declaratives is exceptionally high, i.e. the OE translation of Bede's *Historia Ecclesiastica* (Mitchell 1985: §3932; Calle-Martín & Miranda-García 2010), though Afros (2022) shows that in later manuscripts the frequency of V1 declaratives is slightly lower (15% in T as opposed to 20% in earlier witnesses), with changes into the less marked verb-second, V-late or *þa*-VS. Since the OE Bede is based on a Latin source, Ohkado (2000) concluded that it is a foreign structure, but more detailed studies by Cichosz (2017b) and Cichosz et al. (2016) show the pattern is mostly of the Latin original. This study investigates whether the same level of independence is visible for all the functions of V1 declaratives attested in the OE Bede.

3 Study design

The aim of the study is to conduct a comprehensive quantitative and qualitative analysis of all V1 declaratives found in the YCOE corpus (Taylor et al. 2003) and find answers to the following research questions:

a. What are the functions of V1 declaratives in OE prose, and is the narrative function dominant?

b. What is the actual frequency of narrative V1 and how is it distributed in the OE prose corpus?

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

c. What is the relation between narrative V1 and descriptive V1 declaratives?

d. Which OE prose texts use the identified patterns and do these compositions form a homogeneous group?

e. Does Latin influence any particular function of V1 declaratives in the OE Bede?

f. What is the main reason for the uneven distribution of V1 declaratives in OE prose?

Since YCOE is a syntactically annotated corpus, it was possible to extract the clauses automatically by means of the following CorpusSearch 2 (Randall 2005-13) query:

node: IP-MAT*

query: ((IP-MAT* iDomsNumber 1 finite_verb)

AND (IP-MAT* iDoms NP-NOM*)

AND (NP-NOM* iDoms ! *con*|*pro*|*exp*))

The query searched for verb-initial main clauses, which automatically excludes conjunct clauses introduced by coordinating conjunctions since these are treated separately in studies of OE syntax because of their specific constituent order patterns, often overlapping with subordinate clauses (Mitchell 1985; van van Kemenade 1987; Pintzuk 1999; Olga Fischer & van der Wurff 2000; Ringe & Taylor 2014; Cichosz 2022). An obligatory element of the clause included in the query was an overt subject, though an additional mirror query searching specifically for clauses without overt subjects was used for the purposes of the case study described in Section 4.3. The definition file of the finite verb used in the query does not cover negated forms (negation is a known factor associated with V1 order in OE and research shows that it is not related to narrative inversion, cf. Cichosz 2020) and subjunctive forms[2] since the latter were predominantly directive, representing the so-called jussive subjunctive structures. Forms ambiguous between indicative and subjunctive were included in order not to narrow down

[2]The file includes the following verb labels:
RP+VBPI|RP+VBDI|RP+VBP|RP+VBD|RP+BEPI|RP+BEDI|RP+BED|RP+BEP|RP+HVPI|RP+HVDI|RP+HVP|RP+HVD|RP+AXPI|RP+AXDI|RP+AXP|RP+AXD|RP+MDPI|RP+MDDI|RP+MDP|RP+MDD|VBPI|VBDI|VBP|VBD|BEPI|BEDI|BED|BEP|HVPI|HVDI|HVP|HVD|AXPI|AXDI|AXP|AXD|MDPI|MDDI|MDP|MDD

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

the sample excessively; any examples of directives were manually excluded, and so were the examples with modal verbs since a closer analysis in context revealed that all of them could be interpreted as directive. Since imperative clauses in OE (as in any Germanic languages) are usually V1, morphologically indicative verbs in clauses with similar illocutionary force could follow the same pattern. Such clauses were not taken into account since their function is clearly different from narrative and descriptive V1 declaratives that this study is focused on. (5) and (6) are examples of clauses returned by the query which were discarded at this stage.

(5) *Weaxe ge nu*
grow you now
'Grow now.' [cootest,Gen:9.7.381]

(6) *Mæg gehyran se ðe wyle be þam halgan mædene*
may hear this.NOM who will by the.DAT holy.DAT maiden.DAT
Eugenian …
Eugenia.DAT
'Let everyone interested hear about the holy maiden Eugenia…'
[coaelive,+ALS_[Eugenia]:1.197]

The query returned 918 examples but after the manual cleanup the final number of V1 clauses included in the study was narrowed down to 840.

The qualitative analysis of V1 clauses and the division into particular functions required looking at the examples in broader context, and the Dictionary of Old English Web Corpus (DOEC) was used for this purpose. As a result, when V1 clauses are presented in isolation, they are accompanied by YCOE identifiers, but when broader context is presented, the examples are followed by DOEC identifiers. The functional classification was based on a bottom-to-top approach, i.e. the only pre-set category was narrative and the analysis was supposed to identify all the distinct patterns visible in the data.

The analysis of extralinguistic factors (genre, translation status, dialect) was based on YCOE metadata; texts were divided into early (o1-o2) and late OE (o3-o4) on the basis of their presumed date of composition.

The analysis of the relation between source Latin clauses and their OE equivalents in the case study of Bede's *Historia Ecclesiastica* (Section 4.2) was based on manual alignment of all 459 V1 declaratives from Bede to their most probable Latin sources. Since it is not certain which Latin manuscript was the direct source of the OE translation (Lemke 2015: 43) and it is very probable that this particular version of Latin Bede did not survive (Wallis 2013), the choice of the Latin

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

edition was based on its availability in electronic form (Plummer 1896). Clauses were classified as modelled on Latin if the source clause had a clause-initial verb form (finite or non-finite); presence of the subject was not considered necessary since pronominal subjects are rarely overt in Latin. If the verb was placed later in the Latin clause, or if the OE clause was a serious paraphrase of the original fragment, the clause was classified as . Examples provided in Section 4.2 hopefully make this methodology transparent.

4 Results

4.1 General results

Table 1 presents YCOE texts with the highest absolute number of V1 declaratives (with 10 clauses as a cut-off point). Rather unsurprisingly, and confirming the results of numerous studies reported in Section 2, Bede takes the first position in this classification, with the highest number of clauses covering more than a half of all the OE prose examples of the structure under investigation.

Table 1: Absolute frequency of V1 declaratives in different OE prose texts.

Bede's *History of the English Church*	459
Vercelli Homilies	67
Blickling Homilies	45
Alexander's *Letter to Aristotle*	42
Ælfric's Catholic Homilies I	35
Ælfric's Lives of Saints	21
Ælfric's Catholic Homilies II	20
Anglo-Saxon Chronicle D	13
Boethius' *Consolation of Philosophy*	11
Heptateuch	10
Anglo-Saxon Chronicle C	10
YCOE in total	840

What is more, it should be noted that 73% of the clauses in question come from 4 non-Ælfrician texts (Bede, Vercelli Homilies, Blickling Homilies and Alexander's Letter), while (as shown in Section 2) the only detailed and context-based functional studies of the structure in OE prose are primarily based on the output of Ælfric.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

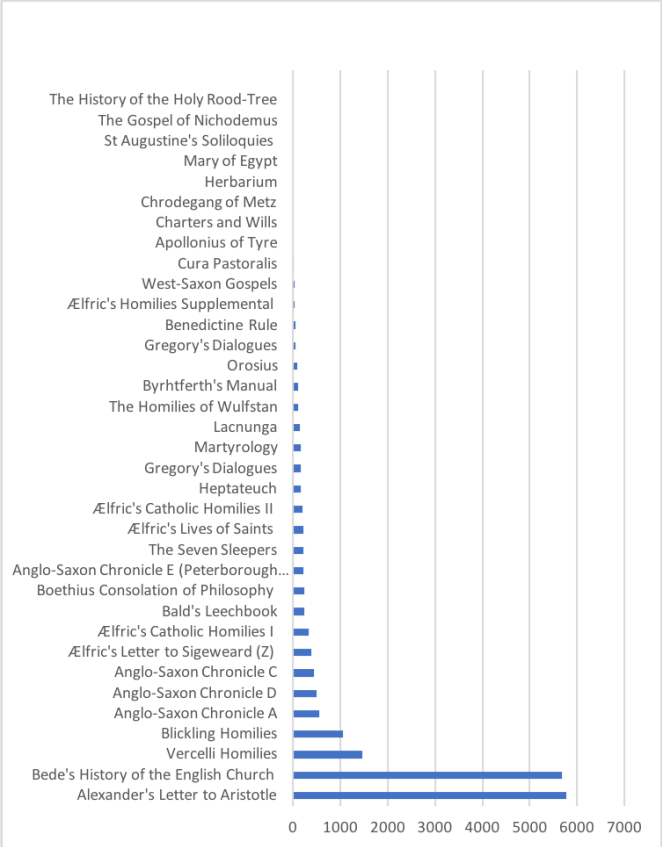


Figure 1: Normalised frequencies of V1 declaratives per 1,000,000 words in the longest YCOE texts.

When normalised frequencies are taken into account, the difference between individual texts becomes even more conspicuous, and it is actually Alexander's Letter, not Bede, which turns out to have the highest frequency of V1 declaratives in YCOE. Nevertheless, the top four remains the same (Bede, Alexander's Letter, Vercelli Homilies and Blickling Homilies) and Ælfric's compositions show very low results, which indicates that we do need functional studies of non-Ælfrician texts to understand the function of V1 declaratives in OE prose.

Functionally speaking, it is necessary to analyse the clauses in a larger context to get a clear idea of what a V1 clause does in a particular place in discourse. As explained in Section 3, this study followed a bottom-to-top approach, which made it possible to identity four clear patterns in the data, with some interrela-

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

tions between them. A typical context for a V1 declarative (in bold) is shown in (7).

(7) *Ða wæs sum þegen annanias gehaten. & his*
then was some.NOM nobleman.NOM Ananias called and his.NOM
wif saphira hi cwædon him betweonan þæt
wife.NOM Saphira.NOM they.NOM said them.DAT between that
hi woldon bugan to þæra apostola geferrædene namon
they.NOM would turn to the.GEN apostles.GEN fellowship.DAT took
þa to ræde þæt him wærlicor wære þæt hi
then to counsel.DAT that them.DAT wiser were that they.NOM
sumne dæl heora landes wurþes æthæfdon weald hu
some.ACC part.ACC their.GEN land.GEN worth.GEN withhold in case
him. getimode ***Com þa se þegen mid feo***
them.DAT befell came then the.NOM nobleman.NOM with riches.DAT
to þam apostolum*. Ða cwæð petrus: annania: deofol*
to the.DAT apostles.DAT then said Peter.NOM Ananias.NOM devil.NOM
bepæhte þine heortan. & þu hæfst alogen þam halgan
deceived your.ACC heart.ACC and you.NOM have lied the.DAT holy.DAT
gast.
ghost.DAT

'There was some nobleman called Ananias and his wife Saphira. They said to each other that they wanted to turn to the apostles' fellowship. (They) agreed that it would be better for them to withhold some part of their land's worth in case something happened to them. **Then the nobleman came with riches to the apostles**. Then Peter said: Ananias, devil deceived your heart and you have lied to the Holy Ghost.'

[ÆCHom I, 22, [004000 (357.88-92)]]

As can be seen, the V1 declarative is surrounded by *þa*-VS clauses (and there is also a V1 clause without an overt subject, *namon þa* ..., which may be an important part of the whole picture, cf. 4.3), so narrative V1 seems to be one of many tools OE writers had at their disposal when presenting a story. The clause *Com þa se þegen* undoubtedly introduces a subsequent event in the narration[3]. Whether this clause is more "dramatic" than the following *Ða cwæð petrus* is perhaps less clear, but the narrative function of the clause in bold is certain.

[3]An important factor here is also the subject shift (from the plural, indicating Annanias and his wife, to the nobleman himself), observed regularly in narrative V1 clauses with a nominal subject, cf. Section 4.3

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

The next pattern, related to the general narrative function and potentially a subtype of it, is the presentational structure, illustrated by (8).

(8) *Mid þy ic ða wolde georlicor þa þing geseon &*
when I.NOM then wanted more carefully the.ACC things.ACC see and
furðor eode þa geseah ic gyldenne wingeard trumlicne &
further went then saw I.NOM golden.ACC vine.ACC strong.ACC and
fæstlicne, & þa twigo his hongodon geond þa
firm.ACC and the.NOM twigs.NOM his.NOM hang over the.ACC
columnan Ða wundrode ic þæs swiðe. ***Wæron in þæm***
columns.ACC then wondered I.NOM this.GEN very were in the.DAT
wingearde gyldenu leaf *& his hon &*
vinebush.DAT golden.NOM leaves.NOM and his.NOM tendrils.NOM and
his wæstmas wæron cristallum & smaragdus, eac þæt
his.NOM fruit.NOM were crystal.NOM and emeralds.NOM also the.NOM
gimcyn mid þæm cristallum ingemong hongode.
gem.NOM with the.DAT cristals.DAT among hang
'When I wanted to see these things more carefully and went further, then I saw a golden vine, strong and firm, and the twigs were hanging over the columns. Then I wondered at this very much. **There were golden leaves in the vine bush** and the vine's tendrils and fruit were crystal and emerald, and also a gem was hanging there among the crystals.'
[Alex, 003700 (8.14-8.16)]

Such clauses would usually be translated into Modern English with existential *there*. In this case, the golden leaves of the vine are an important element of the story of all the riches that Alexander's army encountered in the city of Porres, so the clause represents coordinating discourse relations, functioning on the main level of the storyline. Since this seems to be the case with all the presentational structures identified in this study, unless signalled otherwise, presentational clauses (93 in total) will be included in the general statistics considering narrative V1.

In addition, there is a big group of V1 declaratives used in reporting structures to quote (mainly direct) speech, as in (9). While such reported dialogues undoubtedly push the story forward and they could also be treated as examples of narrative inversion, I will keep them separate in the statistics since the level of attraction between V1 and verbs of saying is exceptionally high (Cichosz 2017a).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(9) *Eft he frægn, hwæt seo þeod nemned wære, þe*
Again he.NOM asked what the.NOM people.NOM called were which
heo of cwomon. ***Ondswarede him mon*** *þæt heo*
they.NOM of came answered him.DAT one.NOM that they.NOM
Ongle nemde wæron. ***Cwæð he:*** *Wel þæt swa mæg:*
Angles.NOM called were Said he.NOM well that.NOM so may
forðon heo ænlice onsyne habbað, & eac swylce
because they.NOM angelic.ACC faces.ACC have and also so
gedafonað, þæt heo engla æfenerfeweardas in heofonum sy.
becomes that they.NOM angels.NOM afterwards in heavens.DAT be
'He asked again about the name of the tribe from which they came.**Someone answered him** that they were called Angles. **He said**: It may well be so because they have angelic faces and it is only becoming that they will later be angels in heaven.' [Bede 2, [002000 (1.96.21-22)]]

Last but not least, the descriptive type shown in (10) clearly stands out from (7)–(9). In this case, the text presents the life and heroic death of Saint Martin. Before the story unfolds, St. Martin's background is presented and we get some additional information about his place of birth and family. Then, the story moves back to the main line of narration and the most important events of the saint's life are described. Apart from the fact that V1 declaratives shown in (10) play a different role in discourse, appearing in a sub-structure of the text, what is quite striking is the great accumulation of such clauses in one fragment, which is not an exceptional situation as such clusters of descriptive V1 declaratives appear in all of the "top four texts" identified in this study (cf. Table 1).

(10) *... þe Martinus wæs haten.* ***Wæs he Gode swiðe gecoren on***
... who Martin.NOM was called was he.NOM God.DAT very dear in
his þeawum. Wæs he in Pannana þære mægðe
his.DAT customs.DAT was he.NOM in Pannonia.DAT the.DAT diocese.DAT
in woruld cumen, in Arrea ðam tune. Wæs he
in world.ACC come in Arrea.DAT the.DAT city.DAT was he.NOM
hwæðre in Italia afeded, in <Ticinan> þære byrig. Wæs
nevertheless in Italy.DAT raised in Ticine.DAT the.DAT city.DAT was
he for worulde swiðe godre gebyrde. Wæron his
he.NOM for world.DAT very good.GEN birth.GEN were his.NOM
fæder & his modor buta hæðen. Wæs
father.NOM and his.NOM mother.NOM both.NOM heathen.NOM was

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

his fæder ærest cyninges þegn, *& þa æt nehstan*
his.NOM father.NOM first king.GEN thane.NOM and then at next
geþah þæt he wæs tribunus, þæt is ealdorman cyninges
happened that he.NOM was tribune.NOM that is alderman.NOM king.GEN
þegna. Þa sceolde he, sanctus Martinus, nyde beon
thanes.GEN then should he.NOM saint.NOM Martin.NOM necessary be
sona on his giogoðhade on geferræddenne cyninges þegna
soon on his.DAT youth.DAT on fellowship.DAT king.GEN thanes.GEN
'… who was called Martin. **He was very dear to God in his customs. He has come to the world in the diocese of Pannonia, in the city of Arrea. He was nevertheless brought up in Italy, in the city of Ticinum (Pavia). He was of very good birth to the world. His father and his mother were both heathen. His father was first the king's thane**, and then (he) was a tribune, that is the alderman of the king's thanes. Then he, St. Martin, in his youth had to belong to the fellowship of the king's thanes.'
[LS17.2, [000200 (1))]- [000800 (10))]]

After the aforementioned manual analysis of all the V1 clauses in context, it turned out that almost one third of all examples retrieved from YCOE corresponds to the descriptive function represented by (10), which is quite a high proportion, considering the generally low frequency of the structure. Table 2 illustrates the functional division in the texts with the highest number of V1 clauses (the presentational function, being closest to the narration-pushing function, is included under the label "narrative").

First of all, it should be noted that narrative V1 is very well represented in all the texts, though the proportions differ from 95% in Catholic Homilies I to 29% in Alexander's Letter. In three of the four texts with the highest number of V1 clauses, the proportions between narrative, quotative and descriptive uses are quite balanced. Alexander's Letter (a fictitious account of Alexander the Great's military campaign in India written to his master Aristotle, cf. Khalaf 2013) stands out from the other texts because of its rare use of quotative V1, but this must be related to the narrative structure of the travelogue since there are hardly any dialogues reported in the text. Most importantly, however, Table 2 shows that descriptive V1 is literally never used by Ælfric, even though he consistently uses V1 declaratives as a narrative device. This is a full confirmation of the tentative observation made in Cichosz (2022) on the basis of text samples. Interestingly, though, other texts are not that consistent, i.e. non-Ælfrician compositions, if they use V1 declaratives at all, would usually make use of both narrative and descriptive V1.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Table 2: Functions of V1 declaratives in YCOE texts.

text	narrative (incl. presentational)	quotative	descriptive	all
Bede's History	213 (46%)	93 (21%)	153 (33%)	459
Vercelli Homilies	27 (40%)	11 (17%)	29 (43%)	67
Blickling Homilies	14 (31%)	13 (29%)	18 (40%)	45
Alexander's Letter to Aristotle	12 (29%)	2 (4%)	28 (67%)	42
Ælfric's Catholic Homilies I	20 (95%)	1 (5%)	0 (0%)	21
Ælfric's Lives of Saints	28 (80%)	4 (20%)	0 (0%)	35
Ælfric's Catholic Homilies II	17 (85%)	2 (5%)	0 (0%)	20
Anglo-Saxon Chronicle D	11 (85%)	0 (0%)	2 (15%)	13
YCOE (total)	448 (53%)	136 (16%)	256 (31%)	840

When particular functions are presented as normalised frequencies (Figure 2) it becomes evident that the "top four" texts are exceptional not only in their particularly frequent use of V1 declaratives, but also in the employment of V1 clauses for the descriptive and quotative functions. Other texts, if V1 declaratives are attested in them at all, treat them as a narrative device.

When we zoom in on the longest YCOE files, where the probability of finding a reasonable number of V1 declaratives is the highest (see Table 3), it turns out that many texts clearly avoid the structure entirely, and there is no clear pattern connecting texts where V1 is in use. For example, Gregory's Dialogues C, Orosius and Cura Pastoralis, all being early translations just like Bede's History, show no V1 of any type. The only regularity visible in the data is that descriptive V1 is only used in early texts (though this is a one-way relation only, i.e. virtually all descriptive V1 clauses are early OE[4]but not all early OE texts have descriptive V1 clauses), while narrative V1 is scattered around the corpus and appears (mostly with very low frequencies) in various texts. Its use must have been stylistically conditioned since even Ælfric does not use it in all of his compositions (the highest frequency of V1 is visible in Catholic Homilies I and the structure is virtually absent from his Supplemental Homilies).

[4]The only exception is the Peterborough Chronicle (Anglo-Saxon Chronicle E), which is classified as a late text in YCOE, but it is partly based on earlier manuscripts so the possibility of retaining obsolete structures in this particular text is quite high.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Table 3: The use of V1 declaratives in the longest YCOE texts (normalised frequencies).

text	words	sub-period	genre	trans-lated	dialect	V1	narrative	descrip-tive
Ælfric's CH I	106,173	late	homilies	no	WS	330	292	0
Ælfric's LoS	100,193	late	biography	no	WS	210	200	0
Ælfric's CH II	98,583	late	homilies	no	WS	203	183	0
Gregory's Dialogues C	91,553	early	biography	yes	mixed	156	156	0
Bede's History	80,767	early	history	yes	mixed	5683	2637	1894
West-Saxon Gospels	71,104	late	bible	yes	WS	28	15	0
Cura Pastoralis	68,556	early	treatise	yes	WS	15	15	0
Ælfric's Sup. Homilies	62,669	late	homilies	no	WS	32	32	0
Heptateuch	59,524	late	bible	yes	WS	168	168	0
Orosius	51,020	early	history	yes	WS	78	78	0
Boethius	48,443	early	philosophy	yes	WS	227	227	0
Vercelli Homilies	45,674	early	homilies	no	WS(?)	1467	591	635
Blickling Homilies	42,506	early	homilies	no	mixed	1059	329	423
ASC E	40,641	late	history	no	mixed	221	197	25

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

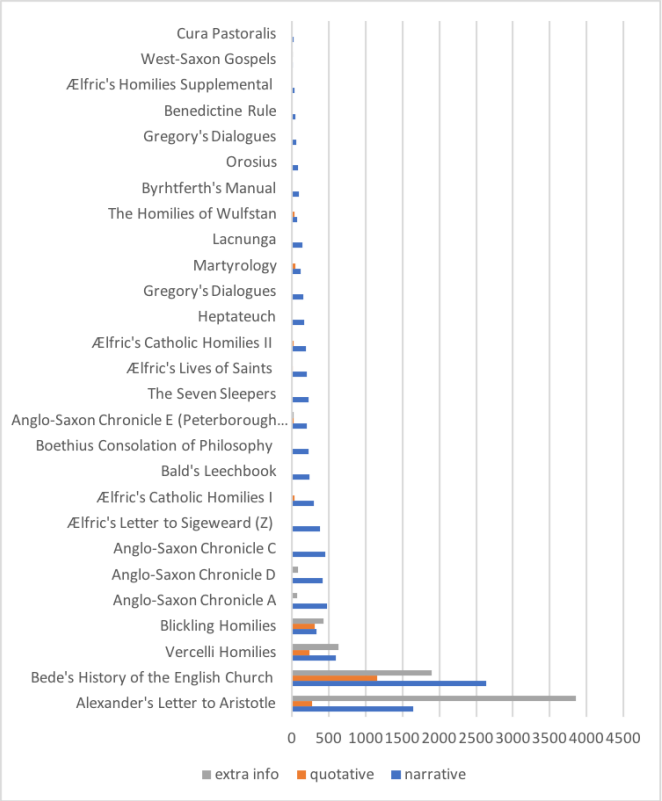


Figure 2: Normalised frequencies of V1 declaratives with particular functions per 1,000,000 words.

Alexander's Letter is not included in Table 3 because of its length (only 7,271 words), but it is also an early text, so the analysis shows that descriptive V1 clauses seem diachronically limited, but before any final conclusions are drawn, let us check whether the structure may safely be treated as native to OE.

4.2 Bede: influence of Latin

The discussion about the Latinate character of the syntax of OE Bede is long-lasting and so far inconclusive. On the one hand, Bede is not a gloss and it shows quite a lot of independence from the Latin source (according to Cichosz et al. 2016: 351, only 5% of all clauses in the text are translated phrase by phrase; for the West-Saxon Gospels the corresponding result is 15%). In addition, OE textual records, rich as they are, have their limitations so it would not be practical

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

to exclude a long text such as Bede from syntactic investigations only because sometimes "[t]he choices involved in translating Bede's Latin into OE [...] manifest themselves in somewhat artificial structures" (Rowley 2011: 9). It is a fact that the OE Bede contains unexpectedly high amounts of constructions which are considered foreign or at least inspired by Latin, e.g. the dative absolute (Stanton 2002: 58; Scheler 1961). On the other hand, the Bede translator did introduce numerous and sometimes substantial changes to the original composition, e.g. extensively modifying the preface and omitting certain events from the narration, letters relating to the English Church and some information concerning the physical world (St-Jacques 1983: 86). What is more, the overall dependence of V1 declaratives in Bede on the Latin source text suggested in earlier studies (Ohkado 2000) was shown to be only partial in more recent corpus-based investigations (Cichosz 2017b). What has not been checked so far, however, is the impact of Latin on V1 declaratives performing particular functions in the text.

Table 4 shows that the proportion of clauses inspired by the Latin source is not the same between the functions, but none of the functions is completely from the original.

Table 4: Relation to Latin in V1 declaratives from Bede depending on function.

	modelled	independent
narrative (with presentational)	133 (62%)	81 (38%)
quotative	48 (52%)	45 (48%)
descriptive	64(42%)	89 (58%)

(11) is an example of a V1 declarative with a narrative function which was classified as modelled on the Latin since both languages have the verb in the clause-initial position and the subject is expressed later in the clause. It turns out that as many as 62% of narrative V1 clauses in Bede may be treated as inspired by Latin, though the degree of closeness is not always the same (e.g. in many cases there is no subject in the Latin source).

(11) a. *Forðferde se arwyrða fæder Cuðbyrht in*
forth-went the.NOM honorable.NOM father.NOM Cuthbert.NOM in
Farne ðæm ealonde;
Farne.DAT the.DAT island.DAT

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

b. *Obiit autem pater reuerentissimus in insula*
died also father.NOM most honourable.NOM in
Farne
island.ABL Farne.ABL
'The honourable father Cuthbert died on the Isle of Farne.'
[cobede,Bede_4:30.374.1.3734]

Nonetheless, uses such as (12), which shows a substantially paraphrased fragment, are well attested in the text as well. This, together with the fact that narrative V1 declaratives appear in many original OE prose works and in poetry, may be treated as a clear indication of the native origin of the structure, but it seems that the actual frequency of the structure in some (translated) texts may be inflated by Latin influence.

(12) a. *Com se foresprecena hungur eac swylce hider on*
came the.NOM aforesaid.NOM famine.NOM also such hither on
Bryttas
Britons.ACC
'Likewise the aforesaid famine reached the Britons.'

b. *INTEREA Brettones fames sua praefata magis*
meanwhile Britons.NOM famine.ACC their.ACC aforesaid.ACC greatly
magis-que adficiens, ac famam suae malitiae
greatly-and suffering and reputation.ACC their.ACC wickedness.GEN
posteris diuturnam relinquens
posterity.DAT lasting.ACC leaving
'Meanwhile the Britons, suffering more and more from the aforesaid famine, and leaving the long-lasting reputation of their malice to the posterity....' [cobede,Bede_1:11.48.19.434]

In the case of quotative V1, the ratio of uses to modelled clauses is roughly 1 to 1, while in the case of descriptive V1, which has the most limited distribution in the corpus, examples which clearly do not follow the Latin source text are a visible majority (58%). This is a rather surprising result since the easiest explanation for the use of this untypical structure in OE prose records would be to treat it as an example of a foreign structure transferred from Latin into two translations (Bede and Alexander's Letter) and two collections of homilies, which – even though not translated directly from a particular Latin source – were most probably based on a number of different religious texts written in Latin, like most OE prose texts (Stanton 2002). This hypothesis, however, is nor corroborated by the data, since

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

most examples of descriptive V1 clause resemble (13), where no inspiration for the clause-initial placement of the verb may be identified in the Latin text.

(13) a. *Wæron hi begen on ciriclicum wisum ge on*
were they.NOM both.NOM on church.DAT ways.DAT and on
wisdome haligra gewrita genihtsumlice gelæred
wisdom.DAT holy.GEN scriptures.GEN enough educated

b. *ambo et in rebus ecclesiasticis, et in scientia*
both.NOM and in matters.ABL ecclesiastic-ABL and in science.ABL
scripturarum sufficienter instructi.
scriptures.GEN enough educated
'Both were adequately educated in matters related to the Church and in the science of the scripture.' [cobede,Bede_5:16.446.20.4489]

Naturally, the possibility of copying is unquestionable in cases such as (14), but the number of such instances in the text is too low to give proper justification to the hypothesis of foreign transfer.

(14) a. *Wæs þes ilca Æðelbehrt Eormanrices sunu;*
was this.NOM same.NOM Ethelbert.NOM Eormenric.GEN son.NOM

b. *Erat autem idem Aedilberct filius Irminrici,*
was also this.NOM Ethelbert.NOM son.NOM Eormenric.GEN
'This Ethelbert was the son of Eormenric.'
[cobede,Bede_2:5.110.16.1037]

In short, whatever the origin and motivation for the use of descriptive V1 is, Latin cannot be treated as its source and it should rather be interpreted as an additional factor, possibly increasing the frequency of the structure in the translated texts which use it.

4.3 Does the subject matter?

So far, the analysis has not identified a clear motivation for the extremely skewed distribution of V1 declaratives (especially those with the descriptive function) in OE prose records. The context-based analysis of V1 clauses, however, makes it possible to notice a certain regularity in the data, i.e. V1 declaratives with an overt subject tend to be followed by V1 clauses without a subject, just like in (15). Here, a V1 declarative with a narrative function is immediately followed by a subjectless V1 declarative with a descriptive function, providing additional information about the referent mentioned in the preceding clause.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(15) *Feng to his rice Osred his sunu;* ***wæs***
took to his.DAT kingdom.DAT Osred.NOM his.NOM son.NOM was
eahtawintre cniht
eight-winter.NOM boy.NOM
'Osred, his son, took over the throne; (he) **was an eight-year boy**.'
[Bede 5, [037800 (16.446.4)]]

In Bede, where V1 declaratives are so frequent, it is possible to come across several examples of such V1 clusters, though the subjectless V1 is not necessarily descriptive. In (16) a V1 narrative clause with a subject (*seo cwen* 'the queen') is followed by two subjectless V1 clauses which clearly operate on the highest level of the narrative structure, presenting subsequent events in the story.

(16) *Wæs heo seo cwen sona lustfulliende þære godan*
was she.NOM the.NOM queen.NOM soon glad the.DAT good.DAT
foresetenesse & willan þæs iungan. ***Sende him þa to***
purpose.DAT and will.DAT the.GEN young.GEN sent him.ACC then to
Cent to Ercenbyrhte þam cyninge*, se wæs hyre*
Kent.ACC to Ercenbyrht.DAT the.DAT king.DAT this.DAT was her.GEN
eames sunu; ***bæd þæt he hine arwurðlice to Rome***
uncle.GEN son.NOM asked that he.NOM him.ACC kindly to Rome.DAT
onsende
sent
'She, the queen, was soon glad about the good purpose and will of the young man. **(She) sent him then to Kent to Ercenbyrht the king**, he was her uncle's son; **(she) asked that he kindly sent him to Rome**.'
[Bede 5, [042100 (17.452.19-21)]]

Thus, it seems that V1 declaratives with a subject are not that specific in their functional range because, at least in Bede, the same functions are performed by subjectless V1 clauses. A similar phenomenon seems to operate in Lives of Saints. Here, subjectless V1 clauses are narrative in function as in (17), and they are rarely preceded by a V1 clause with a subject, because the latter are not that frequent in this text.

(17) *And he leop sona cunnigende his feðes, hwæðer he cuðe*
and he.NOM ran soon testing his.ACC feet.ACC whether he.NOM could
gan. ***Eode þa mid blisse binnan þam temple mid þam***
walk went then with glory.DAT into the.DAT temple.DAT with the.DAT

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

> ***halgum apostolum, þone hælend herigende***. *Þa oncneowan*
> holy.DAT apostles.DAT the.ACC Saviour.ACC praising then recognised
> *hine ealle þe hine cuðon ær, and micclum wundrodon*
> him.ACC all.NOM that him.ACC knew before and greatly wondered
> *þæs wædlan hæle.*
> the.GEN poor.man.GEN health.DAT
>
> 'And he soon ran, testing his feet, whether he could walk.**(He) went then with glory to the temple with the holy apostles, praising Jesus**. Then everyone who had known him recognised him and wondered greatly at the health of the poor man.' [ÆLS, Peter's Chair, 001400 (32-36)]

The number of non-conjunct main declarative clauses without a subject is not particularly high in OE prose. Table 5 shows that the whole YCOE contains only around a thousand examples of the structure, but the interesting thing is that in most texts V1 clauses without a subject clearly outnumber those with a subject. The only exceptions to this rule are (again) the "top four" texts: Bede, Vercelli Homilies, Blickling Homiles and Alexander's Letter, where the frequency of V1 declaratives with an overt subject is not only exceptionally high but also higher than the frequency of subjectless V1 clauses (it should also be noted that in one of Ælfric's texts, Catholic Homilies I, the two patterns are quite balanced, whereas in his other works the tendency is the same as in most YCOE texts).

In order to investigate the possible interrelations between V1 declaratives with and without a subject, I conducted a case study of two texts where both patterns are attested, but they seem to be on two extremes regarding the analysed structure: Bede (the whole text was taken into account) and Lives of Saints (only the books where V1 clauses with a subject are used at least once, i.e. Book of Kings (28 V1 clauses without a subject), Martin (25), Basil (25, Peter's Chair (18) and Maccabees (16), 112 (47%) clauses in total). The classification was based on the labels determined for V1 declaratives with a subject (narrative, with presentational as a subtype, quotative and descriptive). Table 6 presents the outcomes of this analysis (based on raw numbers).

It turns out that there may be some connection between the two patterns since Ælfric consistently avoids the descriptive function in both cases, employing V1 declaratives (with or without a subject, mostly the latter) to push the narration forward or introduce a quote. In Bede, on the other hand, all the functions are attested and even though the proportions are not identical, we cannot say that the presence of the subject makes a drastic difference. This result makes it possible to draw a tentative conclusion: perhaps it is the clause-initial position of the finite

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Table 5: V1 declaratives with and without an overt subject in YCOE texts.

text	with S	without S
Bede's History of the English Church	459	195
Vercelli Homilies	67	28
Blickling Homilies	45	26
Alexander's Letter to Aristotle	42	9
Ælfric's Catholic Homilies I	35	36
Ælfric's Lives of Saints	21	236
Ælfric's Catholic Homilies II	20	97
Anglo-Saxon Chronicle D	13	35
Boethius Consolation of Philosophy	11	38
Heptateuch	10	30
Anglo-Saxon Chronicle C	10	32
YCOE (total)	840	1,069

Table 6: A case study of the function of V1 declaratives with and without S in Bede and Lives of Saints.

	LoS		Bede	
	V1 with S	V1 no S	V1 with S	V1 no S
narrative	20 (95%)	89 (79%)	213 (46%)	123 (64%)
quotative	1 (5%)	23 (21%)	93 (20%)	35 (18%)
descriptive	0 (0%)	0 (0%)	153 (33%)	35 (18%)
all	21	112	459	193

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

verb which is responsible for the pragmatic function of the clause, and the subject is not a crucial element of the pattern at all.

Table 7: Subject types in V1 declaratives with an overt S in YCOE.

	personal pronoun	noun	other pronoun
narrative	140 (31%)	253 (56%)	58 (13%)
quotative	78 (57%)	44 (32%)	14 (10%)
descriptive	125 (49%)	109 (43%)	19 (7%)

When we compare subject types between all the functions of V1 declaratives identified in this study in YCOE (Table 7), it turns out that while the proportion of other (non-personal) pronouns is similar, personal pronouns are well-represented in the quotative and descriptive function, and less frequent in narrative V1, which is quite logical, considering the fact that narrative inversion often introduces new characters to the story. The tendency becomes completely clear when nominal subjects are divided into proper nouns, nouns with overt marking of definiteness (modification by a demonstrative pronoun, a possessive pronouns or a genitive) and without it (bare nouns, optionally modified by adjectives and quantifiers such as sum "some"). Table 8 shows shat in the case of descriptive V1 clauses, even if the subject is nominal, it is rarely discourse-new, while narrative V1 clauses co-occur with indefinite noun phrases in 25% of the cases (112 out of 451).

Table 8: Subtypes of nominal subjects in V1 declaratives in YCOE.

	definite marking	no marking	proper noun
narrative	105 (41%)	112 (44%)	36 (14%)
quotative	21 (48%)	17 (39%)	6 (14%)
descriptive	90 (83%)	2 (2%)	17 (16%)

In addition, it should be noted that 117 out of 140 cases of narrative V1 with personal pronoun subjects (84%) come from the "top four" texts, another 9 from one of the manuscripts of the Anglo-Saxon Chronicle and another 4 from Vercelli Homilies E (130 out of 140, 93% altogether) so the distribution of the structure is very limited.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Moreover, when we add clauses without an overt subject to the equation, an interesting pattern emerges. Table 9 shows the proportions for Bede, Table 10 – for the sample of Lives of Saints.

Table 9: Subject types in V1 declaratives in Bede.

	personal pro-noun	definite NP	indefi-nite NP	proper N	other pro-noun	no S
narrative	98 (29%)	42 (12%)	39 (11%)	21 (6%)	13 (4%)	123 (37%)
quotative	69 (54%)	4 (3%)	8 (6%)	2 (2%)	10 (8%)	35 (27%)
descriptive	75 (40%)	53 (28%)	1 (1%)	14 (7%)	75 (40%)	75 (40%)

Table 10: Subject types in V1 declaratives in LoS sample.

	personal pro-noun	definite NP	indefi-nite NP	proper N	other pro-noun	no S
narrative	0 (0%)	9 (8%)	8 (7%)	0 (0%)	3 (3%)	87 (81%)
quotative	0 (0%)	0 (0%)	1 (4%)	0 (0%)	0 (0%)	23 (96%)

It turns out that in Lives of Saints (LoS) the proportion of V1 declaratives without an overt subject is much higher in all of the attested functions than in Bede (81% vs. 37% for the narrative function, 96% vs. 27% for the quotative function; the descriptive function is not used in the former text), while personal pronouns are never used. Thus, it turns out that there are two important differences between the analysed texts: a complete lack of V1 declaratives with a descriptive function in LoS and consistent avoidance of V1 declaratives with personal pronoun subjects by Ælfric. Interestingly, when we add up the number of clauses with personal pronoun subjects and no overt subjects in Bede, the resulting proportions are much closer to the ones in LoS (66% vs. 81% for the narrative function, 81% vs. 96% for the quotative function). The difference between Bede and LoS in subject types in the narrative function proves significant when personal pronouns and clauses without overt subjects are counted separately (X^2 (3, N = 437) = 77.9917, $p < 0.00001$)[5], and even when the two categories are merged, the difference re-

[5]Subtypes of nominal subjects are merged for the purposes of the statistical analysis since some

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

mains significant (X^2 (2, N = 437) = 16.1207, p =.000316). It seems that narrative V1 declaratives with no overt subject such as (18) and those with a personal pronoun subject such as (19) are variants of the same structure whose purpose is to introduce subsequent events in the story when the main protagonist remains the same and there is no need to change the subject.[6]

(18) *Þa gelæhte Petrus hire liþian hand,* ***arærde hi***
then took Peter.NOM her.ACC supple.ACC hand.ACC raised her.ACC
upp hale of þam bedde.
up healthy.ACC of the.DAT bed.DAT
'Then Peter took her supple hand, (he) **raised her up healthy from the bed**.' [ÆLS (Peter's Chair) [002800 (73)]]

(19) *Þa geseah he swa þeostre dene ane under him in*
then saw he.NOM such dark.ACC valley.ACC one.ACC under him.DAT in
niþernesse gesette. ***Geseah he eac feower fyr onæled on***
bottom.DAT placed saw he.NOM also four.ACC fires.ACC lit on
þære lyfte noht micle fæce betweoh him
the.DAT air.DAT not great.DAT space.DAT between them.DAT
tosceaden.
separated
'Then he saw a deep valley beneath. **He saw also four fires lit in the air not far away from one another**.'
[Bede 3, [040400 (14.212.20)]-[040500 (14.212.22)]]

When the subject is a noun, the narrative V1 clause often uses a new subject as in the clause with *se cyning* 'the king' shown (20), or there is a subject shift. Descriptive V1 clauses may also use a nominal subject, but in such a case, the subject is never new but rather old or accessible (as in *his baan* 'his bones' in the same example).

(20) *Forðon þa he wæs mid wæpnum & mid feondum eall*
because then he.NOM was with weapons.DAT and with enemies.DAT all
utan beheped, & he seolfa onget þæt hine mon ofslean
outside cut.off and he.NOM self knew that him.ACC one.NOM kill
scolde, þa gebæd he for þam sawlum his weorodes.
shall then prayed he.NOM for the.DAT souls.DAT his.GEN army.GEN

of them are not represented in LoS.

[6]A similar observation was made with regard to Old Icelandic (Booth & Beck 2021).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Cwædon heo bi ðon þus in gydde: Drihten God miltsa
said they.NOM by this thus in song.DAT lord.NOM God.NOM pity
þu sawlum ussa leoda, cwæð se halga
you.NOM souls.DAT our.GEN people.GEN said the.NOM holy.NOM
Oswald, þa he on eorðan saag. ***Wæron his***
Oswald.NOM when he.NOM on earth.ACC sank were his.NOM
baan gelæded & gehealden in þæm mynstre, *þe*
bones.NOM taken and held in the.DAT monastery.DAT which
we ær cwædon æt Beardan ea. ***Heht se***
we.NOM before said at Bardney.DAT river.DAT ordered the.NOM
cyning, se ðe hine slog, his heafod on steng
king.NOM this.NOM who him.ACC killed his.ACC head.ACC on pole.ACC
asetton;
put
'Because he was surrounded on all sides by armed enemies, and he understood that he would be killed, he prayed for the souls of his army. Thus they said it in a song: 'Lord God, have mercy on the souls of our people, said the holy Oswald, when he fell to the ground.' **His bones were taken and preserved in the monastery at Bardney river**, which we mentioned before. **The king who slew him ordered that his head should be set on a pole**.' [Bede 3 [024700 (10.188.13)]-[025000 (10.188.20)]]

In LoS, nominal subjects used in narrative inversion (infrequent as they are) behave in a similar way, if they are not discourse-new, there is a change of protagonist in the story as in (21), when Avitianus is introduced, but then the story switches to the actions of St. Martin in order to get back to Avitianus a few clauses later.

(21) *Auitianus hatte sum hetol ealdorman,*
Avitianus.NOM was.called some.NOM cruel.NOM commander.NOM
wælhreow on his weorcum, se gewrað fela manna,
savage.NOM on his.DAT deeds.DAT who.NOM bound many.ACC men.GEN
and on racenteagum gebrohte to þære byrig Turonia, wolde
and on chains.DAT brought to the.DAT city.DAT Tour.DAT would
hi þæs on mergen mislice acwellan ætforan þære
they.ACC this.GEN on morning.DAT cruelly kill before the.DAT
burhware, þa wearð hit þam bisceope cuð. Þa
citizens.DAT then became it.NOM the.DAT bishop.DAT known then

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

smeade se halga wer hu he heora gehelpan
thought the.NOM holy.NOM man.NOM how he.NOM they.GEN help
mihte, and eode to middre nihte ana to his gatum,
might and went to middle.DAT night.DAT alone.NOM to his.DAT gates.DAT
and þa þa he inn ne mihte, he anbidode þærute. ***Wearð***
and then when he.NOM in not might he.NOM waited outside became
þa se ealdorman awreht færlice þurh Godes
then the.NOM commander.NOM awoken suddenly through God.GEN
engel*, and he him gramlice to cwæð, List ðu and rest*
angel.ACC and he.NOM he.DAT sternly to said lie you.NOM and rest
þe, and Godes þeowa lið æt þinum gatum?
REFL and God.GEN servant.NOM lies at your.DAT gates.DAT
'There was a cruel commander called Avitianus who bound many men and brought them in chains to the city of Tour, (he) wanted to kill them cruelly in front of the citizens the next morning. Then the bishop heard about it and considered how he could help them. He went in the middle of the night alone to the city gates and when he could not get in, he waited outside. **Then the commander was suddenly awoken by God's angel** and he sternly told him "You are lying and resting and God's servant is lying at your gates.' [ÆLS (Martin), [028200 (1143)]-[028400 (1151)]]

In short, it seems that in the fragments with a new topic or topic shift narrative inversion with a nominal subject is used, and in the cases with topic continuation, there is a text-dependent variation between V1 declaratives without an overt subject (attested in Ælfric and many other, mostly late, texts) and with a personal pronoun subject (attested in Bede, Vercelli Homilies, Blickling Homiles and Alexander's Letter, which are all early texts), and thus the most probably variable responsible for this variation is diachrony.

5 Summary and conclusions

The study confirms that V1 declaratives are a rare structure with a skewed distribution, but they are present in most OE prose texts, both early and late, translated and non-translated, written in West Saxon or showing mixed dialect features. While the frequency of V1 declaratives in Bede is inflated by Latin, uses of the Latin source correspond to all the functions of V1 identified in OE prose, so we cannot treat any of them exclusively as examples of foreign transfer. Since no clear variable connects the compositions where the pattern is used, it seems to

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

have been a stylistic choice of the (mostly) anonymous Anglo-Saxon writers, perhaps inspired by OE poetry where this discourse device is used most regularly (Mitchell 1985: §3932).

The basic function of V1 declaratives is narrative, i.e. they are used to introduce subsequent events in the story and represent clearly coordinating discourse relations. This function has the highest corpus frequency (53% of all examples) and the most balanced textual distribution (despite generally low frequency, single examples are attested in numerous texts). Next, there are descriptive V1 clauses, operating in the sub-structure of the text, i.e. providing additional information about known discourse referents and functioning as examples of elaboration and explanation. Such clauses, frequent as they are if YCOE is analysed as a whole (31% of the whole dataset), are largely restricted to 4 texts regularly mentioned in this study: Bede's Historia, Vercelli Homilies, Blickling Homilies and Alexander's Letter. Finally, the last function attested in the data is the V1 reporting construction, which constitutes only 16% of all examples and its distribution is also rather restricted.

The analysis suggests that V1 descriptive clauses are an early native structure (they are absent from all late texts, including Ælfric's works), and this function must have been taken over by other, probably more hypotactic structures, e.g. relative clauses. Narrative V1 was considerably more conventionalised in OE, since it appears in numerous texts, covering both early and late records. It must be noted that narrative V1 declaratives with pronominal subjects are used mostly by the "top four" early OE texts; they alternate with V1 without an overt subject in other YCOE texts. This has some important implications for the syntactic interpretation of the structure: OE is a V2 language but it has its specificity related to the behaviour of personal pronoun subjects: they usually fail to invert and there is a very specific group of closed-class elements which may trigger of a pronominal subject (this includes *þa* 'then' and *þonne* 'then', *ne* 'not', wh-words, cf. Ringe & Taylor 2014: 399-400). The fact that V1 declaratives feature personal pronoun subjects has been a topic of major discussions among OE syntacticians since it means that narrative inversion must be structurally equalled with other pronoun-inverting cases representing V-to-C movement (Pintzuk 1999; Ohkado 2004), but it is unclear what actually triggers the pattern (Ringe & Taylor 2014: 408). Considering the extremely limited distribution of narrative inversion with pronominal subjects, the question is whether the structure should be included in the same category, on a par with other, much better attested patterns. Perhaps the stronger presence of V1 declaratives in early OE data means that they are a remnant of an older, possibly Proto-Germanic, structure, which was already becoming obsolete. It is a fact that V1 declaratives disappeared from English soon

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

after the end of the OE period, and the constraint on personal pronoun inversion was a factor which could have contributed to its growing limitation and final loss.

All in all, the study reveals that an OE discourse device creating dynamics in the story is V1 (with or without an overt S), and it is quite certain that it was one of many forms with such a function available to Anglo-Saxon writers. This suggests a number of potential directions for further study. First of all, in order to have a full picture of the situation, it is necessary to take a closer look at both prose and poetry records. We know that narrative inversion is used quite regularly in OE verse, but there are no studies showing whether descriptive V1 declaratives are attested in poetry and discussing the discourse function of V1 clauses without an overt subject. What is more, a promising direction for further study would be a function-to-form mapping of various discourse devices used for specific functions in OE (opening a new story, summarising the discussion, transition to a new action, introducing a new character, quoting speech, describing a known referent, etc.). Such a study would be difficult to design and it would require quite a lot of close reading and contextual analysis, but this paper, among many other analyses of OE, shows quite clearly that it is impossible to understand OE discourse structure without a more qualitative approach to textual data.

References

Afros, Elena. 2022. The verb-initial independent declarative clauses on which the most complete witnesses to the Old English Bede disagree. *Studia Neophilologica* 96(1). 1–18. DOI: 10.1080/00393274.2022.2071762.

Allen, Cynthia L. 1995. *Case marking and reanalysis.* Oxford: Oxford University Press.

Asher, Nicolas & Alex Lascarides. 2003. *Logics of conversation.* New York: Cambridge University Press.

Asher, Nicolas & Laure Vieu. 2005. Subordinating and coordinating discourse relations. *Lingua* 115. 591–610. DOI: 10.1016/j.lingua.2003.09.017.

Axel, Katrin. 2007. *Studies on Old High German syntax: Left sentence periphery, verb placement and verb-second.* Amsterdam: Benjamins.

Bech, Kristin. 2012. Word order, information structure, and discourse relations: A study of Old and Middle English verb-final clauses. In María José López-Couso Meurman-Solin Anneli & Bettelou Los (eds.), *Information structure and syntactic change in the history of English*, 66–86. Oxford: Oxford University Press.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Booth, Hannah. 2018. *Clause structure and expletives: Syntactic change in Icelandic.* Manchester: University of Manchester. (Doctoral dissertation).

Booth, Hannah & Christin Beck. 2021. Verb-second and verb-first in the history of Icelandic. *Journal of Historical Syntax* 5(28). 1–53. DOI: 10.18148/hs/2021.v5i28.112.

Cichosz, Anna. 2017a. Old English verbs of saying and verb-initial order. *English Language and Linguistics* 21(3). 533–558. DOI: 10.1017/S1360674316000174.

Cichosz, Anna. 2017b. Verb-initial main clauses in Bede: A translation effect? In *Current developments in English historical linguistics. Studies in honour of Professor Rafał Molencki*, 68–89. Wydawnictwo Uniwersytetu Śląskiego.

Cichosz, Anna. 2020. Negation and verb-initial order in Old English main clauses. *Journal of English Linguistics* 48(4). 355–381. DOI: 10.1177/0075424220941911.

Cichosz, Anna. 2022. Old English V-initial and *þa*-VS main clauses: Independent constructions or allostructions? *Constructions and Frames* 14(2). 301–336. DOI: 10.1075/cf.21002.cic.

Cichosz, Anna, Jerzy Gaszewski & Piotr Pęzik. 2016. *Element order in Old English and Old High German translations.* Amsterdam: Benjamins.

Cichosz, Anna & Piotr Pęzik. 2021. Markers of Ælfric's style. *Style* 55(2). 223–252. DOI: 10.5325/style.55.2.0223.

Hinterhölzl, Roland & Svetlana Petrova. 2005. Rhetorical relations and verb placement in Early Germanic languages: Evidence from the Old High German tatian translation (9th century). In Michael Grabski Manfred Stede Christian Chiarcos & Luuk Lagerwerf (eds.), *Salience in discourse: Multidisciplinary approaches to discourse*, 71–79. Münster: Stichting/Nodus.

Hopper, Paul J. 1975. *The syntax of the simple sentence in Proto-Germanic.* The Hague: Mouton.

St-Jacques, Raymon C. 1983. *Hwilum word be worde, hwilum andgit of andgiete?* Bede's Ecclesiastical History and its Old English translator. *Florilegium* 5. 85–104.

Calle-Martín, Javier & Antonio Miranda-García. 2010. *Gehyrdon ge þæt gecweden wæs* – a corpus-based approach to verb-initial constructions in Old English. *Studia Neophilologica* 82. 49–57. DOI: 10.1080/00393271003795202.

van Kemenade, Ans. 1987. *Syntactic case and morphological case in the history of English.* Dordrecht: Foris.

van Kemenade, Ans & Bettelou Los. 2006. Discourse adverbs and clausal syntax in Old and Middle English. In Ans van Kemenade & Bettelou Los (eds.), *The handbook of the history of English*, 224–248. Oxford: Wiley Blackwell. DOI: 10.1002/9780470757048.ch10.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Khalaf, Omar. 2013. The Old English Letter of Alexander to Aristotle: Monsters and hybris in the service of exemplarity. *English Studies* 94(6). 659–667. DOI: 10.1080/0013838X.2013.814324.

Kiparsky, Paul. 1995. Indo-European origins of Germanic syntax. In Adrian Battye & Ian Roberts (eds.), *Clause structure and language change*, 140–169. Oxford: Oxford University Press.

Lehmann, Winfred P. 2005. *A grammar of Proto-Germanic*. Jonathan Slocum (ed.). Linguistics Research Center, University of Texas at Austin.

Lemke, Andreas. 2015. *The Old English translation of Bede's Historia Ecclesiastica Gentis Anglorum in its historical and cultural context* (Göttinger Schriften zur Englischen Philologie; 8). Göttingen: Georg August-Universität. DOI: 10.17875/gup2015-787.

Los, Bettelou. 2000. *Onginnan/beginnan* with bare and *to*-infinitive in Ælfric. In Olga Fischer, Anette Rosenbach & Dieter Stein (eds.), *Pathways of change: Grammaticalization in English*, 251–274. Amsterdam: Benjamins.

Mitchell, Bruce. 1985. *Old English syntax*. Oxford: Clarendon Press.

Ohkado, Masayuki. 2000. Verb first constructions in Old English, with special reference to Bede's ecclesiastical history of the english people. In Masachiyo Amano & Hirozo Nakano (eds.), *Synchronic and diachronic studies on language: A festschrift for Dr. Hirozo Nakano*. Nagoya: Nagoya University.

Ohkado, Masayuki. 2004. On the structure and function of V1 constructions in Old English. *English Studies* 85. 2–16. DOI: 10.1076/enst.85.1.2.29106.

Olga Fischer, Willem Koopman, Ans van Kemenade & Wim van der Wurff. 2000. *The syntax of early English*. Cambridge: Cambridge University Press.

Petrova, Svetlana. 2006. A discourse-based approach to verb placement in early West-Germanic. In Shinichiro Ishihar, Michaela Schmitz & Anne Schwarz (eds.), *Working papers of the SFB632, interdisciplinary studies on information structure 5*, 153–182. Potsdam: Universitätsverlag Potsdam.

Pintzuk, Susan. 1999. *Phrase structures in competition: Variation and change in Old English word order*. New York: Garland Publishing.

Plummer, Charles (ed.). 1896. *Venerabilis beadae historiam ecclesiasticam gentis anglorum*. Oxford: Oxford University Press. https://www.thelatinlibrary.com/bede.html.

Randall, Beth. 2005. CorpusSearch 2. Computer program. Designed by Beth Randall, Ann Taylor and Anthony Kroch. http://corpussearch.sourceforge.net/credits.html.

Ringe, Don & Ann Taylor. 2014. *The development of Old English*. Oxford: Oxford University Press.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Rowley, Sharon M. 2011. *The Old English version of Bede's Historia Ecclesiastica*. Cambridge: Brewer.

Scheler, Manfred. 1961. *Altenglische Lehnsyntax: Die syntaktischen Latinismen in Altenglischen*. Berlin: Ernst-Reuter.

Smith, Jesse. 1971. *Word order in the older Germanic dialects*. University of Illinois. (Doctoral dissertation).

Stanton, Robert. 2002. *The culture of translation in Anglo-Saxon England*. Brewer. Cambridge: D. S. Brewer.

Stockwell, Robert. 1977. Motivations for exbraciation in Old English. In Charles N. Li (ed.), *Mechanisms of syntactic change*, 291–314. New York: University of Texas Press. DOI: 10.7560/750357-009.

Taylor, Ann, Anthony Warner, Susan Pintzuk & Frank Beths. 2003. *The York-Toronto-Helsinki parsed corpus of Old English Prose (YCOE)*.

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

Wallis, Christine. 2013. *The Old English Bede: Transmission and textual history in Anglo-Saxon manuscripts*. University of Sheffield. (Doctoral dissertation). uk.bl.ethos.595184.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Chapter 9

V2-Relatives in Old English

Bettelou Los & Stefano Coretta

The verb-placement asymmetry between main (verb-second, "V2") and subclauses (verb-final) found in Dutch and German is less pronounced in Old English (OE) (Pintzuk 1999); the question is to what extent, as the distinction between main and subclauses cannot readily be determined on the basis of punctuation or other typographical conventions. It has long been noted that Old English shows constructions similar to the "V2-relatives" in modern Dutch and German (Mitchell 1985: 188), but they are much more difficult to identify. This paper proposes that the discourse functions of such clauses in Dutch and German might be used as a diagnostic for V2-relatives in OE, as these clauses tend to be used for "second mentions" of a major protagonist who has just been introduced. V2-relatives are less embedded than other relative clauses, and can be argued to represent a midway point in the grammaticalization of relatives, from a paratactic to a hypotactic structure. Key here is the difference between a demonstrative and the relative pronoun that developed from it.

1 Introduction

In the modern West Germanic languages, movement of the finite verb to second position ("verb-second" or "V2") is a main clause phenomenon, and hence leads to a word order asymmetry between main (V2) and subclauses (verb-final). Although the consensus is that Old English is also a V2-language, albeit with some modifications (van Kemenade 1987, see also §2.4 below), the asymmetry between main and subclause is less pronounced (Pintzuk 1999). Main and subclauses, and sentences in general, are not distinguished in the mss by a consistent punctuation system, which adds to the challenges of determining the extent of the asymmetry. There is one type of clause in particular that has long been known to be

Bettelou Los & Stefano Coretta. 2025. V2-Relatives in Old English. In Ulrike Demske & Barthe Bloom (eds.), *Discourse structure and narration: A diachronic view from Germanic*, 263–294. Berlin: Language Science Press. DOI: ??

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

analytically ambiguous (Allen 1977: 86, Mitchell 1985: 88), and may well be a contributing factor to the difficulty of establishing to what extent OE exhibits the word order asymmetry found in the modern West Germanic languages:[1]

(1) *Maurus gemette ænne man eft* ***se*** *wæs yfele getawod*
Maurus met a man afterwards SE.NOM.M was evilly stricken
'Maurus met a man afterwards who was evilly stricken.'
[ÆLS[Maur]:283.1661]

(2) *And Maurus ða gemette ær he to mynstre come ænne dumbne*
and Maurus then met before he to monastery came a dumb
cnapan ⁊ ***se*** *wæs creopere eac*
boy and SE.NOM.M was cripple too
'and Maurus then met before he came to the monastery a dumb boy and he was a cripple as well.' [ÆLS[Maur]:19.1507-8]

For OE, syntactic investigations have been able to rely on the York-Toronto-Helsinki Parsed Corpus of Old English Prose (YCOE, Taylor et al. 2003), a 1.5 million word corpus of Old English prose based on the 1981-release of the Toronto *Dictionary of Old English Corpus* containing all the major Old English prose works. Each word of the corpus is tagged for part of speech, and syntactically annotated by labelled bracketing in the Penn Treebank format, which has since become the industry standard for syntactic annotation. Data can be retrieved by CorpusSearch 2 (Randall 2005–2007).

In YCOE, (1) has been coded as a relative clause, and hence as a subclause (IP-SUB), whereas (2) has been coded as a main clause (IP-MAT), in spite of their very similar structures including their main-clause-like word order, with V2. It may well be the presence of the Tironian sign '⁊' for "and" in (2) that guided the annotator to an analysis of a sentence consisting of two main clauses rather than a main clause and an embedded clause, which is the analysis suggested by the coding of (1). In the PDE translation in Skeat's (1966 [1881–1900]) edition of Ælfric's *Saints Lives* from which (1)-(2) have been taken, cases like (1) are routinely translated by relative clauses, and the *se*-strategy is recogised as one of a range of relativizing strategies available in OE, so from that perspective the interpretation of clauses like (1) as relatives is unproblematic. What is a problem is that

[1]The demonstrative form is in bold, and the finite verb is underlined; they will be marked like this throughout the paper; also note that the demonstrative, the source of the ambiguity, will be glossed as *se*, and will be variously translated as a relative pronoun, a demonstrative pronoun or a personal pronoun.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

relative clauses are not necessarily subordinate clauses – they may also be independent main clauses, and hence have main clause word order. Coding these as IP-SUB will boost the number of subordinate clauses in the data that have the same word order as main clauses, and hence obscure the word order asymmetry between main and subclauses in OE. That OE is an asymmetrical V2 language is already clear from van Kemenade's (1987, 1997) work; see also Taylor (2014: 403-5), and particularly Meklenborg Salvesen & Walkden (2017), who conclude from their investigation of OE *þæt*-complement clauses that OE does not allow embedded V2. This paper adds some further data to these insights as they pertain to relative clauses, and shows that the specific characteristics of V2-relatives echo those found in modern West Germanic, and lead to their specific discourse function of marking a newly introduced protagonist.

This paper will look at all relative clauses in two OE texts: Ælfric's *Lives of Saints* (Skeat (1966) [1881–1900]) and the H ms of *Gregory's Dialogues* (Hecht 1965 [1900–1907]), and compare the use of *se*-relatives with the other types of relative in terms of position of the verb, discourse functions, and antecedents. The paper is structured as follows: §2 will present the types of relative clauses that have been identified for OE: *þe*-relatives, which are marked only by the invariant complementizer *þe*, the "universal embedder"; *se þe*-relatives, which are marked by the universal embedder as well as by a demonstrative pronoun; and *se*-relatives, which are marked only by the pronoun, and which are the focus of the investigation. §3 discusses what has been said about V2-relatives in the modern West Germanic languages. §4 and §5 present the data and the methodology; as expected, it is the *se*-relative that shows the highest frequencies of V2, but what is new is that they serve a particular discourse function in narrative, in that they instantiate the second mention of a newly-introduced protagonist. §6 will place these findings in the wider context of the semantics of main and subclauses, and the development of hypotaxis. §7 concludes.

2 Typology of OE relatives

2.1 Introduction

This section will provide a very brief overview of the main types of relative clause in OE: *þe*-relatives, *se þe*-relatives and, finally, the *se*-relatives, the type that most concerns our investigation into V2-relatives. For more detailed accounts, the reader is referred to the vast literature on the history of relative clauses in English that cannot be done justice here, e.g. Allen (1977), Mitchell (1985: §§2130ff), Traugott (1992: 232), Fischer et al. (2000: 659ff) , Suárez-Gómez (2012).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

2.2 *þe*-relatives

Most frequent of all types is the *þe*-relative. In YCOE, these are marked as having C occupied by *þe* or *ðe* while the Spec,CP position remains empty. *Þe*-relatives are usually verb-final, and restrictive. Unlike *se*-relatives, they allow resumptive pronouns and preposition stranding – evidence of a gap, and hence of movement/copying (Allen 1977; van Kemenade 1987). *Þe* is a worn-down form of *þæt* 'that' which has become an invariant *universal embedder* in OE, marking not only relative clauses but also other clauses that we assume to be embedded (cf. Gothic *ei*). The consensus in the literature is that *þe*-relative clauses are embedded, and subordinate. Example (3) contains two *þe*-relatives (*þe* and *ðe* in bold) as well as a *se*-relative (*ðone*); all three are coded as IP-SUB in YCOE. *Se*-relatives will be discussed in §2.4.

(3) *Sum æþel-boren þægn wæs philippus gehaten · **ðone** asende se*
Some nobly-born thane was Philippus called SE.ACC.M sent the
*casere Commodus **þe** on ðam dagum · rixode fram rome byrig to*
emperor Commodus ÞE on those days ruled from Rome city to
*ðære byrig **ðe** is gehaten alexandria*
the city ÞE is called Alexandria
'A certain nobly-born thane was named Philip, who the emperor Commodus, who was in charge in those days, sent from the city of Rome to the city which is named Alexandria.' [ÆLS_[Eugenia]:5.198]

The second *þe*-relative in (3) is restrictive, and postmodifies an NP (*to ðære byrig* 'to the city') that also contains a demonstrative form, the singular feminine *ðære*, which is grammaticalizing into a definite article in OE; *ðære* is used here to signal to the reader that they should be able to indentify the referent, not because it has been mentioned in the earlier discourse (there is no earlier mention of a city which could be the antecedent) or because readers are expected to know it (unlike Rome), but because there is going to be a postmodifier (in the form of a *þe*-relative) which will clarify what city is meant. Sánchez Argüelles & Los (2022: 172) identify this use of the definite article as marking an important stage in the grammaticalization process, i.e. the stage at which definite articles appear with a referent that is new in the discourse; the use of the definite article is "licensed" by a postmodifier that "activates the general knowledge necessary to locate the referents in a contextual set" (Wagener 2017: 74). Breban (2012) refers to them as "definite first mentions". Typically, such a use of the definite article signifying identifiability is no longer compatible with a demonstrative reading.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

The first instance of the *þe*-relative, however, is not restrictive – it postmodifies a named referent, as does the *se*-relative with *ðone*. The reason for this will be explored in §3.4.

2.3 *Se þe*-relatives

Se þe relatives have a form of the distal demonstrative *se* as well as the embedder *þe*. The distal demonstrative doubles as the definite article in OE. Its paradigm is given in Table 1:

Table 1: The paradigm of the distal demonstrative in OE

	singular			plural
	masculine	feminine	neuter	all genders
nom	se	seo	þæt	þa
gen	þæs	þære	þæs	þara
dat	þæm	þære	þæm	þæm
acc	þone	þa	þæt	þa
instr		þy, þon		

Se þe relatives have an inherent degree of ambiguity, as they may represent two or perhaps three separate types: (i) a *þe*-relative that has a demonstrative antecedent as an appositive NP with a restrictive relative clause, comparable to PDE "i.e." (*hiera mægas him mid wæron* ***þa*** **þe** *him from noldon* 'their kinsmen were with him, i.e. those who did not want to leave him' in [ChronA 755.1–29]); (ii) a *þe*-relative whose demonstrative antecedent has a syntactic function in the higher clause, comparable to the generic "free relative" use of PDE *he that lives in a glass house should not throw stones* (Allen 1977; see also Allen 2020 for more detail and analysis of this type); (iii) there is also the possibility that *se þe* is a "compound" relativizer.

In YCOE, types (i) and (ii) are distinguished from each other by type (ii) having an empty Spec,CP position, as the demonstrative is best analysed as a constituent of the higher clause, outside the relative clause. We have followed YCOE coding in our searches, which means that type (ii) will have been included under the *þe*-relative.

The *se þe*-relative has been described as "mostly restrictive" (Mitchell 1959). *Se þe*-relatives are not found with preposition stranding, and this is the only type that exhibits "case attraction", a fairly rare phenomenon in the case of type

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(i), in which the demonstrative appears with the case it would have if it was a constituent of the higher clause (van Kemenade 1987: 150; Allen 2020); this might be a kind of appositional use, cf. *him* in the translation:

(4) *ge seceað þone hælynd* ***þone*** *þe on rode ahangen wæs*
you seek the.ACC Saviour SE.ACC.M ÞE on cross hanged was
'You seek the Saviour, him, who was hanged on the cross.'
[cowsgosp,Mt_[WSCp]:28.5.2139]; (Allen 1977: 87)

2.4 *Se*-relatives

Se-relatives are the type that is the focus of this investigation. An example is (5):

(5) *Se feorða godspellere is Iohannes Cristes moddrian sunu,* ***se***
the fourth gospelwriter is John Christ.GEN aunt.GEN son SE.NOM.M
wæs Criste swa leof þæt he hlynode uppan his breoste.
was Christ.DAT so dear that he leaned upon his breast
'The fourth gospelwriter is John, Christ's aunt's son, who was so dear to Christ that he leaned upon his breast.' [ÆLS_[Mark]:159.3304]

Se-relatives often have V2 word order (Allen 1977; van Kemenade 1987: 149). *Se*-relatives do not show case attraction, and do not occur with resumptives or preposition stranding, i.e. without a clear structural gap (Wende 1915, Allen 1977). *Se*-relatives in OE are, as rule, non-restrictive (Los & van Kemenade 2018: 131-132, 138; Roma 2007: 266). Non-restrictive relatives are known to be less embedded than restrictive relatives: they do not define their antecedent but add extra information, and are less integrated into the structure of the higher clause from a syntactic, prosodic, and informational point of view (Huddleston & Pullum 2002: 1034–1035). Non-restrictive relative clauses in PDE are well-known to exhibit main clause behaviour, such as verb phrase preposing, negative constituent preposing and topicalization (Emonds 1970, Hooper & Thompson 1973, Green 1976, de Vries 2012, Heycock 2017). In terms of discourse function, they are often used for expressing the next event on a plotline as in (6) (foregrounding) but also for information in terms of explanations, reasons and causes (backgrounding), as in (7):

(6) *She gave the letter to the clerk, who copied it.* (Depraetere 1996: 699)

(7) *If Peter called the dean, who hates me, I would be in trouble.*
(Jasinskaja & Poschmann 2020: 142)

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Another interesting use of relative clauses in Dutch is noted by de Vries (2012: 183) (in bold):

(8) *Joop,* ***die nu zijn kamer gaat opruimen!****, komt straks buiten spelen.*
Joop who now his room goes clean comes later outside play
'Joop, who is now going to clean his room!, will come play outside later.'

A sentence like (8) might be uttered by Joop's mother to a friend of Joop's who has presented himself at the door, importantly within Joop's hearing; the non-restrictive relative clause is in fact an order to Joop.

The most natural way to translate *se*-relatives into PDE is by a relative clause, particularly if they have human antecedents, as in (5). Although singular animate antecedents can be relativized by *who* or *that*, they no longer allow a translation by an independent demonstrative, unless they are in the plural, witness the contrast between (9a) and (9b) below. Example (9c) shows "a contrast in animacy, *he* referring to a person or animal, *that* to an inanimate" (Huddleston & Pullum 2002: 1504–1505). Compare also the anaphoric uses in (9d–9e).

(9) a. ***Those*** *who obtain a score of 90% will win a prize.*
b. ******That*** *who obtains the highest score will win a prize.*
c. ***He****/****That*** *saved my life.*
d. *The population of Victoria far exceeds* ***that*** *of Queensland.*
e. **The premier of Victoria will be meeting with* ***that*** *of Queensland.*
(Huddleston & Pullum 2002: 1504, their (3) i–v)

In Dutch and German, main clauses headed by demonstrative pronouns coerce topic shift (Comrie 1997): the implication of using a demonstrative rather than a personal pronoun coerces interpreting the referent as a newly-introduced one rather than as a continuing topic. This is no longer the case in PDE, so that the referent of *she* underlined in (10) is ambiguous – this is what adds interest to the exchange:

(10) Columbo: *No, my wife is not here. She had to go to Chicago to look after her mother. <u>She</u> had a fall and broke her hip.*
Woman at party: *Oh, your wife broke her hip? How terrible*
Columbo: *No, her mother.*
(Columbo, series 10.1, episode *No time to die*;
(Los & van Kemenade 2018: 134)

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

The natural choice to disambiguate the two possible referents in Dutch and German would be the personal pronoun for the continued topic (the wife) and the demonstrative for the new topic (the mother). In PDE, a relative clause would have disambiguated the referent; "who had a fall and broke her hip" could only have referred to the mother. This suggests that (non-restrictive) relative clauses may have stepped in as a compensatory strategy once the topic shifting function of demonstratives was lost; it also explains why a main clause headed by a "standalone" demonstrative referring to a singular human antecedent is much more likely to be annotated as a relative clause than as a main clause by annotators who, as in the case of YCOE, tended to be native speakers of English.

This loss in functionality of the singular demonstrative is already evident in Early Middle English, where we find innovative periphrases – which did not catch on – *he this, she this* (Mustanoja 1960: 137; Miura 2016) as translations for topic-shifting Latin *iste*, indicating a need to restore the earlier functionality. The account provided in Allen (2022) shows that *se*–relatives – reduced to invariant *þe* after the loss of the paradigm of table 1– are no longer found after 1150 with antecedents in the previous discourse, only with generic referents in free relatives, corresponding to Present-Day English *he who*; the loss of gender is tentatively identified as being implicated, supporting a suggestion made in Los & van Kemenade (2018), who point to the fact that the contrast in acceptability between singular and plural demonstratives in PDE as in (9a) versus (9b) makes sense in the context of the OE demonstrative paradigm, which was gendered in the singular but not in the plural.

Cases like (5), and the *ðone*-clause in (3), ambiguous between main clause and subclause, are the focus of the remainder of this paper. This ambiguity extends to standalone demonstratives as complements of a preposition in a clause-initial PP, as in (11):

(11) *His wif wæs gecyged claudia be* ***þære*** *he gestrynde twægen suna ·*
his wife was called Claudia by SE.DAT.F he had two sons
auitum · 7 særgium · 7 ane dohtor eugenian þe we embe spræcaþ.
Avitus and Sergius and one daughter Eugenia ÞE we about talked.
'His wife was called Claudia, by whom he had two sons, Avitus and Sergius, and one daughter Eugenia who we mentioned earlier.
[ÆLS_[Eugenia]: 15.202]

The fact that they do not allow preposition stranding, so always show "pied piping" as in (11), has been argued to be evidence that *se*-relatives involve movement or copying (Allen 1977: 91). Note that the absence of stranding could equally

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

well serve as evidence of main clause "Topicalization" (movement to the prefield); PP-initial adjuncts linking to previous discourse are much more frequent in OE than in later periods, particularly PPs that contain standalone demonstratives as in (11); see Los et al. (2023) for the decline of such "local anchors" in the history of English. PDE favours subjects as the default expression of given information, so that a translation like *by/with her he had two sons* is less felicitous in PDE than a relative clause *by/with whom he had two sons*. This is why such PPs, too, are much more likely to be annotated as relative clauses than as a PP-initial main clauses.

Note that the surface V3 order of (11) is part of the manner in which "V2" plays out in OE: the finite verb will move to C in "focus" contexts (interrogatives, negative initial constituents as well as the "narrative operator" *þa* 'then' as the first element) but to a lower head if some other initial phrase, such as a PP, is occupying the prefield, which leads to surface V3 if the subject is a pronoun, as in (11) (van Kemenade 1987). This means that the sequence *be þære he gestrynde* is a possible main clause order.

In YCOE, *se*-relative clauses receive a syntactic annotation in which the demonstrative is in Spec,CP and C is empty (marked by "0"); as a relative, the clause-label indicates a subclause (IP-SUB). As an example, (12) shows the annotation of the *ðone*-clause of (3); (5) and (11) are annotated along similar lines.

(12) (CP-REL-1 (WNP-ACC-4 (D^A +done))
(C 0)
(IP-SUB (NP-ACC *T*-4)
(VBD asende)
(NP-NOM (D^N se) (N^N casere)

An example of a standalone demonstrative coded as heading a main clause is (13), with the demonstrative a complement of a preposition:

(13) *& þa sealde he heom mid his agenre handa ane trywene flascan*
and then gave he them with his own hands a wooden bottle
wines fulle, to þon þæt hi mihton heom ***þa*** *on*
of-wine full to that that they might for-themselves SE-ACC.SG.F in
heora færelde to underngeweorce habban. Of ***þære*** *hi*
their journey for breakfast have from SE-DAT.SG.F they
druncon, oð þæt hi to rauennam becomon
drank, until that they to Ravenna came
'and then, with his own hands, he gave them a wooden bottle full of wine, so that they might have that for breakfast during their journey. They

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

drank from it, until they arrived in Ravenna.' (i.e. it lasted them all the way to Ravenna) [GD_1_[C]:9.66.12.742-743]

In (13), a referent (a feminine noun meaning 'bottle', underlined) is introduced and then referred back to by two instances of *se* (in bold). The second instance, *of þære*, receives the following syntactic annotation in YCOE, with the clause-label indicating a main clause (IP-MAT):

(14) (IP-MAT (PP (P Of)
(NP-DAT (D^D +t+are)))
(NP-NOM (PRO^N hi))
(VBDI druncon)

What makes an analysis like (14) more likely for an example like (13) is the distance between the antecedent and the demonstrative. Faced with a similar dilemma – demonstrative-headed main clause or embedded relative? – in her Old High German data, Schlachter (2012) also uses distance between antecedent and demonstrative to support a main clause analysis of analytically ambiguous demonstrative-headed V2 constructions.

The question is whether (3) (the ðone-clause), (5) and (11) should receive a syntactic annotation along the lines of (14), with the demonstrative (whether as part of a PP or as a subject or object) heading an independent main clause, rather than (12), with the demonstrative heading a subclause.

3 V2-relatives in Dutch and German

3.1 Introduction

The dilemma posed by *se*-headed structures in OE is presented in the literature as the choice of a main clause headed by a standalone demonstrative versus an embedded relative clause. The emphasis of the literature on V2-relatives in Dutch and German, however, is not so much on the status of main versus relative clause but on the difference between embedded and less-embedded relative clauses. The analysis of main clauses headed by demonstratives as separate sentences never seems to be in doubt, presumably given the clear markings of sentences in written texts by means of punctuation, and their intonational phrasing in speech. The discussions in e.g. Gärtner (2001) and Zwart (2003) focus on V2-clauses that are part of a sentence that includes a previous clause where the referent, typically, is

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

newly-introduced, of the type of (15b–15c). The contrast, then, is between (15b) and the embedded, verb-final relative clause as in (15c).

(15) a. *iemand* *had* *twee zonen.*
someone had two sons
[main clause with V2]

b. *Er* *was* *iemand* ***die*** *had twee zonen.*
there was someone that had two sons
[main clause + V2-relative]

c. *Er* *was* *iemand* ***die*** *twee* *zonen had.*
there was someone that two sons had
[main clause + embedded relative]
'There was someone who had two sons'.

Both Gärtner (2001) and Den Dikken (2003) discuss a third structure as in (16), on the basis of an observation made by Zwart (2003: 6) about responses to a nineteenth-century dialect questionnaire. This questionnaire presented the sentence in (15a) as the start of a story, its function being to introduce the father. The respondents subsequently gave (15b), the V2-relative, but also, unexpectedly, (16):

(16) *Er* *was* *iemand* *en* ***die*** *had twee zonen.*
there was someone and that had two sons
[main clause + main clause]
'There was someone and he had two sons'.

Example (16) contains two main clauses, the second headed by a standalone demonstrative. These were apparently felt, by the repondents, to be equivalent, an indication that they both serve the same function. Gärtner (2001: 109), discussing a similar set of sentences (*Es gibt Philosophen (und) die stammen aus Griechenland* 'there are philosophers who come from Greece') notes that the main difference the version with and without *und* 'and' is one of informational units – a relative clause, whether V2 or embedded, will form an informational unit, as also defined by their intonation/prosody (a rising tone before the relative clause). In the version with *und*, a "falling tone, drop, pause, or similar features" (p. 109) would be expected. This does not sufficiently explain the Dutch case of (16), however, where the first main clause is just as uninformative as in (15), and would not have a falling tone after *iemand*.

The different emphases in the various discussions can be visualized by the Venn-diagram in (1):

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

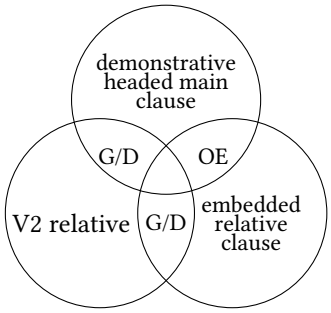


Figure 1: Focus of relevant investigations in the literature on OE and on German (G) and Dutch (D)

The existence of the V2-relative as a less-embedded relative clause has not been mooted in the literature on OE – the dilemma is, as far as we are aware, always presented as being between a relative clause (assumed to be embedded by definition) and a main clause with a standalone, independent demonstrative. We will see in the following sections that the tests put forward in the literature about Dutch and German V2-relatives to untangle the three possibilities, i.e. (15b)–(15c)) and (16), will be too fine-grained for OE; but we will be able to say something about the embedded option versus the rest, at least in some OE texts, and hence identify verb position as a fairly clear marker for less embedded relative clauses.

3.2 Embedded relatives versus V2-relatives

Various syntactic tests demonstrate that V2-relatives behave like main clauses, unlike more embedded relatives. Some tests require manipulation of the data and native speaker judgements about the results that are not available to investigators of a dead language, particularly if these judgements "are sometimes subtle" (Gärtner 2001: 104), and marked by "??" even by native speakers. The item that stands out as a potential feature to check in the OE instances is the fact that the V2-relative needs to be clause-final (p. 100) – an indication, perhaps, like the main-clause characteristics found in the German examples, that the V2-relative started out as a paratactic construction.

Note that Gärtner's (2001) paper provides a single PDE translation for his examples comparing V2-relatives with their more embedded counterparts, like (17a) and (17b), as their meanings do not differ; Gärtner's notation (/) denotes that the V2-relative in (17a) is intonationally integrated into the previous clause, just like its more embedded relative clause counterpart in (17b); it forms a single informational unit with the matrix clause, just like an embedded relative (p. 109). V2-

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

relatives like (17a) are in fact presented as variants that are felicitous only when certain conditions are met, unlike the embedded versions – in a V2 relative, the antecedent (*eine Seite* in (17)) must be presentational, specific and indefinite:

(17) a. *Das Blatt hat eine Seite, (/)* ***die*** *ist ganz schwarz.*
b. *Das Blatt hat eine Seite, (/)* ***die*** *ganz schwarz ist.*
'The sheet has a side that is completely black.' (Gärtner 2001: 98)

The specificity-requirement is best illustrated by an example from Coniglio & Hinterhölzl (2020):

(18) a. *Hans sucht eine Frau,* ***die*** *hat blaue Augen* (de dicto)
b. *Hans sucht eine Frau,* ***die*** *blaue Augen hat* (de re, de dicto)
'Hans is looking for a woman who has blue eyes.'

In (18a), with the V2-relative, Hans is looking for a specific individual (*de dicto*), who is known to him – either directly or by hearsay – and who happens to have blue eyes; (18b), the embedded relative, can also refer to a generic entity (*de re*) – Hans might be looking for a future wife who he hasn't met yet, and who he wants to be blue-eyed. See also Gärtner (2001: 119-120) for examples showing a similar contrast. The upshot of the discussion is that V2-relatives are main clause-like, but show a degree of integration into a higher clause that is reminiscent of an embedded relative. The pivot is the demonstrative itself – and how it differs from a relative pronoun. We will return to this below.

3.3 Demonstrative-headed main clauses versus V2-relatives

A comparison between V2-relatives as in (19a) and demonstrative-headed main clauses as in (19b) (examples from (Gärtner 2001: 133) reveals subtle semantic differences in terms of the interpretation of quantification (see also Coniglio & Hinterhölzl 2020: 205):

(19) a. *Apfeldorf hat viele Häuser,* ***die*** *stehen leer.*
'Apfeldorf has many houses that are empty.'
b. *Apfeldorf hat viele Häuser.* ***Die*** *stehen leer.*[2]
'Apfeldorf has many houses. Those are empty.'

[2] Coniglio & Hinterhölzl (2020) have a proximal demonstrative (*Diese*) for Gärtner's distal demonstrative. Note that proximal demonstratives, in German as well as in OE, do not have relative pronoun counterparts, so do not present the same analytic dilemma.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

The interpretational difference in German seems to carry over in PDE; in (19a), with its V2-relative, many of the houses in Apfeldorf are empty, but there is no implication that Apfeldorf has many houses. The V2-relative is like the embedded relative in this respect. In the demonstrative-headed main clause in (19b), there is an assertion that Apfeldorf has many houses and that all of them are empty. Such judgements cannot be replicated for a dead language like OE, however, although it may be relevant that (19b) can also be expressed by *and* (..., *und die stehen leer*).

Interesting effects appear if we consider Gärtner (2001)'s other famous example, (17), of which de Vries (2012: 195) notes that *eine Seite* has specific reference, and that this leads to a pragmatically odd reading if the second clause is not integrated into the first one (if it was, e.g., separated from the first clause by *and*), as sheets tend to have two sides; de Vries concludes that the existential quantifier set up by *eine Seite* has high scope, which includes the second clause in a V2-relative; see also Gärtner (2001: 129: "one function of [V2-relatives] would be to signal wide scope of the indefinite it modifies"). The first clause ("the sheet has a side") can stand on its own if the indefinite has a non-specific reading, but this reading is unavailable because the demonstrative that heads the second clause requires a specific referent (de Vries 2012: 195). Although this may not be helpful as a diagnostic for our OE examples, what we can take away here is that what causes the close relationship between the V2-relative and the demonstrative-headed main clause has to do with the demonstrative and its interaction with the Spec,CP position, which is its most likely position in a V2 language (see also the YCOE-coding of the OE relatives presented in §2.4). This also explains why the V2-relatives do not show main clause-like behaviour with respect to some of de Vries' tests, which topicalize various constituents from the clause (de Vries 2012: 182–183): the pre-field position is already filled (by the demonstrative).

We can also recall Gärtner's (2001: 98) remark about intonation (see §3.2) – we would expect a demonstrative-headed main clause like (19b) to be immediately preceded by a final boundary marking (Ladd 1986, cited in Gärtner 2001: 98). Final boundaries are marked in modern texts by punctuation, following the idea that a sentence, from full stop to full stop, is "[a]n utterance or complete rhetorical structure which expresses a single idea or *sententia*" (Parkes 1992: 306); the modern concept of "sentence" can of course contain several clauses, both main and embedded, demarcated by commas, colons or semi-colons as well as adverbs and complementizers, but we cannot rely on such meta-signs in earlier texts.

Key to the selection of V2-relatives over demonstrative-headed main clauses appears to be the amount of information provided by the previous clause. The fact that, in the case of V2-relatives, the previous clause (which is the higher clause in the same sentence) appears to be presentational in nature may be the

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

reason why it is felt to be incomplete in terms of how informative it is, so that the subsequent elaboration pinpointing the referent triggers a more integrated structure.[3] Note that observations that the referent needs to be specific, but that the antecedent is often indefinite, seem to point to the denotational link being one of restriction; this is also noted by Gärtner (2001: 113): "in terms of interpretation [V2-relatives] should be treated on a par with restrictive modification." Note that many of the V2-relatives seem to do more than restriction, and convey, in addition, a high degree of salience. In (15b), the expectation is that the sons will be the next topic in the discourse – this is part of the "topic shifting" property of the standalone demonstrative (see again example (10) in §2.4). This conforms to Brandt's (1990: 49) (cited in Gärtner 2001: 131) observation that V2-relatives are restricted to presentational constructions that introduce an individual into the discourse. This is testable in a historical corpus.

3.4 Conclusions

Although the focus of the literature about V2-relatives in Dutch and German is not quite the same as that on the long-standing ambiguity of these structures in OE, and much of the diagnostic apparatus involved requires native speaker judgements which are not available to us, there is one aspect that seems testable: the fact that the V2-relative takes as its antecedent a referent that is newly-presented to the discourse. It follows from this function that the antecedent is likely to be specific, and indefinite. In what follows, we are working from the hypothesis that these demonstrative-headed structures are typically used for what we will call "second mentions": a protagonist has been introduced, and the *se*-clause "takes up" the story – either by providing relevant background information or a foregrounded event. This may not lead to an answer to the question of the exact clausal status – main or embedded? – but it might go some way to explain why we find V2-like word orders in a set of clauses that have received the label "IP-SUB" in YCOE, what function V2 serves, and what a paratactic construction with a standalone demonstrative can tell us about the genesis of embedded relative clauses.

[3] Gärtner (2001: 113) refers to one of his examples of V2-relatives "modifying a preposterously vacuous matrix clause".

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

4 Methodology

4.1 Texts

Based on the discussion of §3, we will focus on the presentational function of the referent of V2-relatives. If the V2-relative picks up a newly introduced, specific entity in the modern West Germanic languages, this might also be its function in OE. As introducing and maintaining protagonists is very much a feature of narrative texts, we will be investigating these "second mentions" in Ælfric's *Lives of Saints* (Skeat 1966, 1881–1900), as our primary text. The *Lives of Saints* is a collection of 37 (mostly) narrative texts in Old English, the only narrative collection of some length (100,193 words) in YCOE that is not a direct translation from Latin. Although Ælfric must have relied for their contents on Latin works, his views on the translation process, explicitly stated in his preface to the translation of Genesis, make it unlikely that he would opt to stay close to the text of his Latin or Greek *Vorlage*, and follow it slavishly (Koopman 1992: 77–78). The narratives in the *Lives of Saints* are not a random collection but "the product of systematic thought" (Hurst 1972: 43). Ælfric's linguistic awareness is evident from the prefaces to his works, where he emphasizes that he intends to write in plain language, and it is clear that he thought of his work as a "consciously 'literary' act" (Clemoes 1956: 182).

Although the main focus of the investigation is on Ælfric's *Lives of Saints*, one other text in YCOE was investigated in order to see whether the findings from *Lives of Saints* could be replicated elsewhere. This second text was a translation from Latin (unavoidable as Ælfric's *Lives of Saints* is the only untranslated narrative text in OE): *Gregory's Dialogues* (ed. Hecht, 1900–1907). This is a collection of legends about early Italian saints, which makes its subject matter comparable to Ælfric's *Lives of Saints*. Bishop Wærferth's original translation of the text was undertaken some time between 870–890, so at least a century before Ælfric's *Lives of Saints*, but we used the H ms, which, at 25,593 words, is shorter than the earlier C ms, and also, very unusually for the period, not a straightforward copy of the C text but a thorough re-working which alters the syntax; clauses are grouped together into *sententiae* in a more coherent fashion (Yerkes 1982), and the position of finite verbs is often changed. This makes it likely that the H ms is less dependent on its Latin *Vorlage* than C. This revision is also more likely to be contemporary with Ælfric's work, as it was probably produced a century or a century and a half after Wærferth's translation (Yerkes 1982: 9-12).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

4.2 Analysis

The various types of relative clauses (see §2) were extracted by means of CorpusSearch 2 (Randall 2005); cases where the relative clause provides an explanation of a term or phrase (usually in Latin), somewhat akin to the equivalent of *id est*, like (20), were noted but excluded from the analysis, as the relative pronoun in such cases is invariably *þæt* 'that' and is more like a meta-comment than a relative clause:

(20) *sum wæs arwurðes lifes wer, se wæs mid Godes gife*
some was honourable.GEN life.GEN man SE was with God.GEN gift
gebledsod & be naman genemned Benedictus, ***þæt*** *is on Englisc se*
blessed and by name called Benedictus SE.NOM.N is in English the
gebletsoda.
blessed
'There was a man of honourable life who was blessed with God's gift and called by the name of Benedictus, which is in English "the blessed one".'
[cogregdH,GDPref_2_[H]:94.23.971]

Also excluded from the analysis were demonstratives as complements of prepositions, like (11) above.[4] The remaining relative clauses were annotated for the presence of a D-head in Spec,CP with an empty ("zero") C, following the coding in YCOE. When we discovered that no human personages were referred to by the neuter pronoun *þæt*, in spite of what we might expect of a language like OE with grammatical rather than natural gender where *wīf* 'woman' and *mægden* 'girl' are neuter nouns, the D-head was split out according to masculine/feminine forms (denoted by "SE" in the analysis) versus neuter ("THAT").[5]

We also coded for Second mention, early verbs (i.e. V2, main clause word order), and the presence/absence of *þe*. The definition of "early verb" was a fairly

[4]This was mainly done in an effort to keep numbers manageable; our impression of the ca. 40 instances in Ælfric's *Lives of Saints* is that these PPs conform to the *se*-relatives we investigated in every respect.

[5]This is not due to a lack of female protagonists, as there are many. The main reason appears to be the way protagonists are introduced – the *se*-relative often follows a name:

(i) *Sum wif hatte Sintice, seo wæs six gear blind*
Some woman was.called Sintice SE was six year blind
'There was a woman called Sintice, who had been blind for six years.'
[ÆLS_[Thomas]:263.7705]

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

restricted one, to exclude cases where there was too little material in the clause to make a judgement, or where the early position of the verb could be due to extraposition of a heavy constituent like a clause. Some other variables were also tracked, i.e. whether the relative clause was part of a set of relative clauses in a single sentence, which might have influenced their word order, or whether the antecedent was a divine being that would not require an introduction.

5 Findings

5.1 *Se*-relatives in in Ælfric's Lives of Saints and *Gregory's Dialogues*

The numbers of Ælfric's *Lives of Saints* and the OE translation of *Gregory's Dialogues* (Ms H) are provided in Table 2. It is clear that the proportion of *se*-relatives is much larger (28% of the total) in *Gregory's Dialogues* than in Ælfric's *Lives of Saints* (17%); we don't want to make too much of this difference, although an appeal to chronology might fit a narrative in which *Gregory's Dialogues*, the earlier text, relies more on *se*-relatives not marked by the universal embedder *þe* than the later *Lives of Saints* because the *þe*-construction is more grammaticalized and a more recent development; to nail this point, it would require a much more extensive investigation than the present one. The purpose of this paper is to compare the *se*-relatives with "early verbs" to what has been said about the V2-relatives in §3.

Table 2: *Se*-relatives in Ælfric's *Lives of Saints* and the OE translation of *Gregory's Dialogues*, ms H

	Ælfric's Lives	Gregory's Dialogues, H
all relatives	1261	373
of which *se*-relatives (excl. *þæt*)	215	106
of which no *þe*	108	80
of which 2nd mention	100	60
of which also early V	94	39

As an illustration of the "second mention" function, Table 3 presents the first 10 instances in the data where the conditions "D=SE", "second mention" and "early verb" are all met; for reasons of accessibility, the data are presented in Skeat's PDE translation, while following the YCOE sentence boundaries. The first sentence

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

has the *who* or *which* in bold that translates an OE *se*-form heading a V2-relative. Note that they all follow the introduction of a new referent, which is why they were coded as "second mentions". The second sentence illustrates the referent's persistence in the discourse. The only referent who does not "persist" is Hilarion (#7); Paschasius in #5 remains activated and persists in the discourse, although he happens to be absent in the second sentence.

Table 3: A random sample of the data coded as "D=SE", "second mention" and "early verb" in Skeat's PDE translation. The *se*-form (expressed by *who*, *which*, *he* and *this* in this translation from Skeat, in bold) with V2 encodes a second mention of a newly introduced referent in the first sentence; their degree of persistence can be gleaned from the referent re-appearing in the second sentence (also in bold).

#	first sentence	next sentence
1	A certain nobly-born thane was named Philip, **whom** the emperor Commodus sent [...] to the city which is named Alexandria [Eugenia 5.198]	**This thane Philip** was not baptized unto God
2	Another man was also blind for seven full years; **he** had a guide who led him everywhere. [Swithun 202.4349]	One day **he** went out as he often did, and the guide became angry, and left the blind man, and ran away, and the other knew not how he should come home
3	There was a certain woman called Syntyche, **who** had been six years blind, and was then healed by the holy apostle, and came, seeing clearly, unto her kinswoman named Migdonia, who had left her blind. [Thomas 263.7705]	Then said Migdonia: 'This man is God Himself, or God's angel, who hath enlightened **thine** eyes thus without leechcraft.'
4	Likewise Athelwold, the holy bishop, who now worketh miracles through God, often told us, that he knew a man with bishop Ælfheah, **who** would drink in Lent whenever it pleased him. [Ash_Wed 65.2738]	Then one day **he** prayed the bishop Aelfheah to bless his cup; he would not, and the fool drank without blessing, and went out.
5	This came to the ears of the nobly-born youth who was wooing Lucy, who was named Paschasius, an impious idolater, **who** enticed the holy maid to make offerings to devils; [Lucy 57.2202]	but the Lord's virgin said, 'A pure offering is this, and acceptable to God, that one should visit widows, and comfort exiles, and help orphan children in their affliction.
6	How there are eight Chief Sins, which sorely fight against us: one is called Gula, that is Gluttony in English, **which** maketh a man eat and drink before the time, or again to take too much in food or in drink. [Memory_of_Saints 268.3471]	**This** destroyeth both soul and body, because it bringeth upon the man much sickness, and bringeth him to death through immoderate drinking;
7	and turned him thence to a certain holy man [**who** was] called Hilarion, with some of his men who would not leave him. [Agnes 381.1985]	Four estates he gave up entirely, together with himself, for the reception of strangers and for alms-deeds.
8	and the Saviour provided, after He had ascended to Heaven, that He should send to the king, as He had before spoken, one of the seventy whom He had chosen to preach, **who** was called Thaddeus, that he might heal the king. [Abdon_and_Sennes 124.4795]	**He** came then, by God's commission, to the aforesaid city, and healed the afflicted king in the Saviour's might, so that the citizens greatly wondered thereat.
9	Again there was a leader, named Seron, in the land of Syria, **who** quoth to his people, 'I will get me a name and overcome Judas, and them that are with him, who despised the king.' [Maccabees 298.5031]	**He** gathered then his host, and went with great array to Judea-land, and many people with him.
10	Again, his son Jove, whom ye worship as a god, **who** desired to kill his unclean father that devoured his brothers as soon as they were born, **this** Jove was filled with foul lust, [Sebastian 172.1310]	and [Ø] took his own sister to his unclean wedlock, even as ye read in your histories.

It became very clear early on that there is a difference between divine referents (God, Christ) and other referents – the divine referents are rarely "presented" but assumed to be known, and hence rarely qualify as "second mentions". When referred to by a *se*-relative rather than a *þe*-relative, they do show higher rates of "early verbs", however, so that in this they follow the trend of *se* being associated

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

with V2-like word order; these instances have been marked "Divinity" in Figure 2 below. Figure 2 clearly shows that the triad "*se*-relative/second mention/early verb" is a coherent package in Ælfric's *Lives of Saints* (dark purple) if we keep "divine referents" separate.

A second look at the items that are "second mention" but do not have an "early verb" (light purple in Figure 2) reveals that all six of these in Ælfric's *Lives of Saints* were not coded as such because of the restricted definition of what we coded as "early verb" (see §4.2); they all have the finite verb in second place, but either do not contain enough constituents to rule out a different analysis, or have an object of such length that it can be argued to have been extraposed, making the position of the finite verb analytically ambiguous.

The situation with respect to verb-placement is not as clearcut in *Gregory's Dialogues*. There are 21 cases of *se*-relatives that did not meet our criteria for "early verb"; three of them have pronominal subjects and surface V3, so can be claimed to have main clause order (see §2.4). Of the remainder, there are only three cases which have been excluded because of the possibility of extraposition;

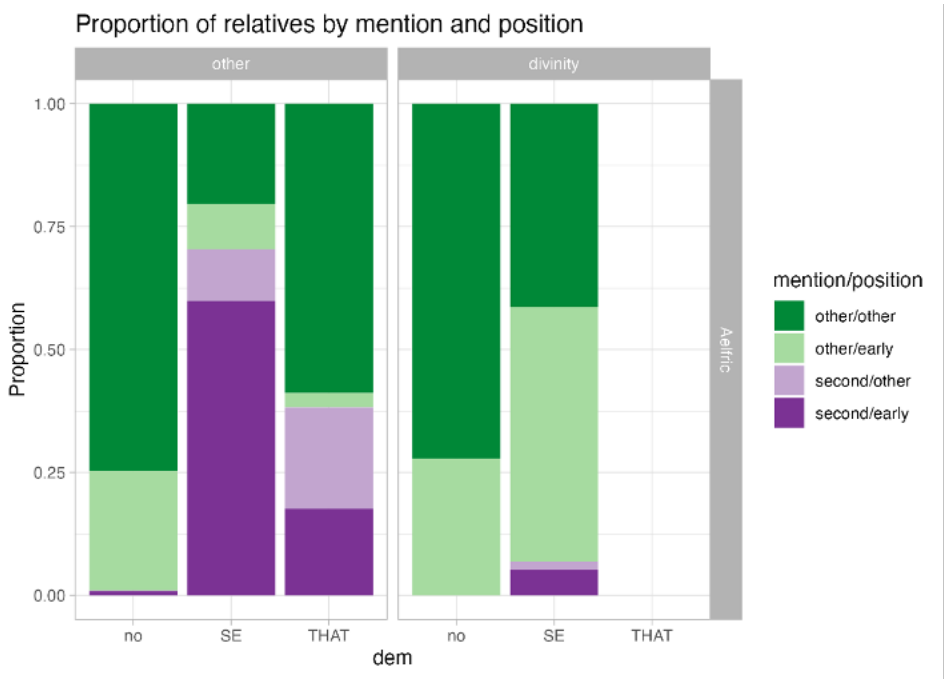


Figure 2: Proportion of demonstratives by mention and by position in Ælfric.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

the rest all have material intervening between the D-pronoun and the finite verb, one of them an anacoluthon, where a *se-* is started but not finished, so that there is a subject *se* but no finite verb:

(21) *On þam þinge ic oncnawe, þæt Benedictus hæfde Paulus gewrixle,*
In that thing I understand that Benedictus had Paul to.example,
se *þa þa he þolode scipes forwyrd & lyre eallra*
SE.NOM.M then then he suffered ship.GEN wreck and loss all.GEN
þara þinga, þe þæar on wæron, þa onfeng he sylf him to
the.GEN things.GEN that there on were then received he self him as
frofre eallra þara lif, þe him mid ferdon.
comfort all.GEN SE.GEN.PL lives that him with travelled
'In which thing I see that Benedict imitated St. Paul: who, when he suffered the wreck of a ship and the loss of all the goods that were on it, then he himself received, for his comfort, the lives of all that were in his company.' [cogregdH,GD_2_[H]:17.141.10.1370]

In terms of the observation that a German V2-relative needs to be clause-final (Gärtner 2001: 100), this was also a consistent finding in Ælfric's *Lives of Saints* – all V2-*se*-relatives were clause-final. In the OE translation of *Gregory's Dialogues* there were occasional cases like (22), however:

(22) *& he þa on þære ilcan stowe, þe ic bufan ymbe spræc,* ***seo*** *is*
and he then in the same place that I above about spoke SE.NOM.F is
geciged Subpentoma, manega gear his lif adreah on halgum dædum
called Suppentonia many years his life conducted in holy deeds
'and he then in the same place that I mentioned above, that-one is called Suppentonia, for many years conducted his life by holy deeds.'
[cogregdH,GD_1_[H]:8.52.9.494]

The *se*-clause is not clause-final; its appears to resemble a PDE deaccented parenthesis (cf. Gärtner 2001: 102), and it is quite likely that these *X is called*-phrases are formulaic chunks. Note that this case was not coded as a "second mention", as the place in question has already been presented in the previous discourse; the most we can say is that it is reactivated here. There is another case in the same text of a *se*-relative that is not clause-final. This one is a "second mention", but note that it is also anacoluthic; the subject of the higher clause, *Exhilaratus ure gefera* 'Exhilaratus our companion' is the antecedent of the *se*-relative, which explains why the *se*-relative immediately follows it; but this apparently causes such a disruption to the flow of the discourse that this subject

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

is re-started after the *se*-relative by means of a resumptive pronoun *he* (also in bold):

(23) *Eac hit gelamp on sume tid, þæt Exhilaratus ure gefera,* ***þone***
also it happened at one time that Exhilaratus our companion SE.ACC.M
þu canst gecyrredne to rihtre drohtnunge, ***he*** *wæs fram his hlaforde*
you know converted to right conduc he was by his master
asended to þam mynstre, þæt he sceolde þam Godes were
sent to the monastery that he should the.DAT God.GEN man.DAT
Benedicte beran twa treowene fatu wines fulle, þa syndon on
Benedict bear two wooden bottles wine.GEN full SE are in
folkisc flaxan gehatene.
vernacular flagons called
'Upon a certain time, Exhilaratus our companion, whom you know as having been converted to the right conduct, was sent by his master to the monastery of the man of God, to carry him two wooden bottles, commonly called flagons, full of wine.'
[cogregdH,GD_2_[H]:18.141.19.1372]

It is clear from Table 2 that a minority of *se*-relatives in these texts did not "qualify" as second mentions. One of them is (24), which is a reactivation of a referent who has been mentioned earlier, *se ylca broðor* 'that same brother':

(24) *Þa ær þæs Anastasies seofeðan dæge, wæs eac forðfered se*
Then before that.GEN Anastasius' seventh day was also departed that
ylca broðor, ***se*** *swaþeah <u>næs</u> na geciged on þære nihte*
same brother SE.NOM.M however not-was not named in that night
betweoh þa oðre broðru
among the other brothers
'Then before that Anastasius's seventh day was also departed that same brother, that one, however, was not named that night among the other brothers.' [cogregdH,GD_1_[H]:8.53.28.510]

Example (24) has not been coded as an "early verb" because *swaþeah* 'however, nevertheless' intervenes. Reactivation after a diversion is an established function of *se* (also for names: *se Cynewulf* 'This Cynewulf, i.e. the one we were talking about earlier' in the *Cynewulf and Cyneheard* interpolation in the Anglo-Saxon Chronicle [ChronA, Plummer 755.1-38]) and of demonstratives in general. Another example of a referent already having been mentioned – in this case, many

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

times before, as the referent is Martin, the protagonist of this "life" – is (26). This example does not have an "early verb" either:

(25) *and Godes sacerdas synd gewuldrode mid þære onwrigennysse*
and God.GEN priests are glorified by the revelation
Martines forðsiðes, ***þonne*** *se halga Michahel mid englum*
Martin.GEN departure.GEN SE.ACC.M the holy Michael with angels
underfeng and Maria seo eadiga mid mædenlicum werodum, and
received and Maria the blessed with virginal companies and
neorxnewang gehylt bliðne mid halgum.
paradise holds happy among saints
'and the priests of God are glorified by the revelation of Martin's departure, whom the holy Michael with his angels and blessed Mary with companies of virgins received; whom paradise holdeth, happy among saints.' [ÆLS_[Martin]:1435.6918]

Of the five *se*-relatives in Ælfric's *Lives of Saints* that have "early verbs" but do not qualify as second mentions, four form a coherent group as a favourite ending of a saint's life, referring back to a divine antecedent with variations on the phrase *þam is wuldor and wurðmynt on ealra worulda woruld* AMEN 'to him is glory and honour in world without end, amen' (e.g. [ÆLS [Peter's_Chair]: 292.2467]; these formulae have a Latin model.[6] The fifth is a case of reactivation, with *þone foresædan Tatheum* 'the aforementioned Tatheus' as its referent ([ÆLS [Abdon and Sennes]:132.4799]); cf. (21) above.

Others do not qualify because although they are part of a follow-up to a presentational clause, they are "third" rather than "second" mentions, as the "second mention proper" is the preceding *se*-relative that gives the priest's name:

(26) *Þa wæs þær sumre neah cyricean mæssepreost,* ***þam*** *wæs*
then was there some.GEN nearby church masspriest SE.DAT.M was
nama Florentius, ***se*** *wæs Florenties ures subdiacones yldra*
name Florentius SE.NOM.M was Florentius' our subdeacon.GEN older
fæder.
father
'Then was there of a nearby church a masspriest who was by name Florentius. He was the father of our subdeacon Florentius.'
[cogregdH,GD_2_[H]: 8.117.2.1151]

[6]We thank one of the reviewers of this paper to point this out to us.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

For the statistical analysis, we fitted a Bayesian linear model using a Poisson distribution for the outcome variable (number of relatives), with demonstrative, second mention and early verb as predictors (including interactions thereof); visualizations of the model results are provided for Ælfric's *Lives of Saints* (Figure 3) and *Gregory's Dialogues* (Figure 4). The figures show the predicted percentage of occurrence of different demonstratives depending on mention (other vs second) and verb position (early vs other). Different Credible Intervals (CrIs) are indicated in the figures by means of different shades of blue: from lighter to darker, 50, 70, 80 and 95% CrIs. A CrI of, say, 90% indicates that the predicted percentage is within that interval at 90% certainty, based on the data and model. Larger intervals have greater certainty than smaller intervals (with greater precision in the predicted percentage there is greater uncertainty).

In general, *se* with second mention and early verbs tends to occur at the same rate in *Lives of Saints* and *Gregory's Dialogues* but while this stands out in Ælfric, this is not so much the case in *Gregory's Dialogues* – *se* also occurs more or less as frequently with verbs in non-early positions, as we saw above, and in cases

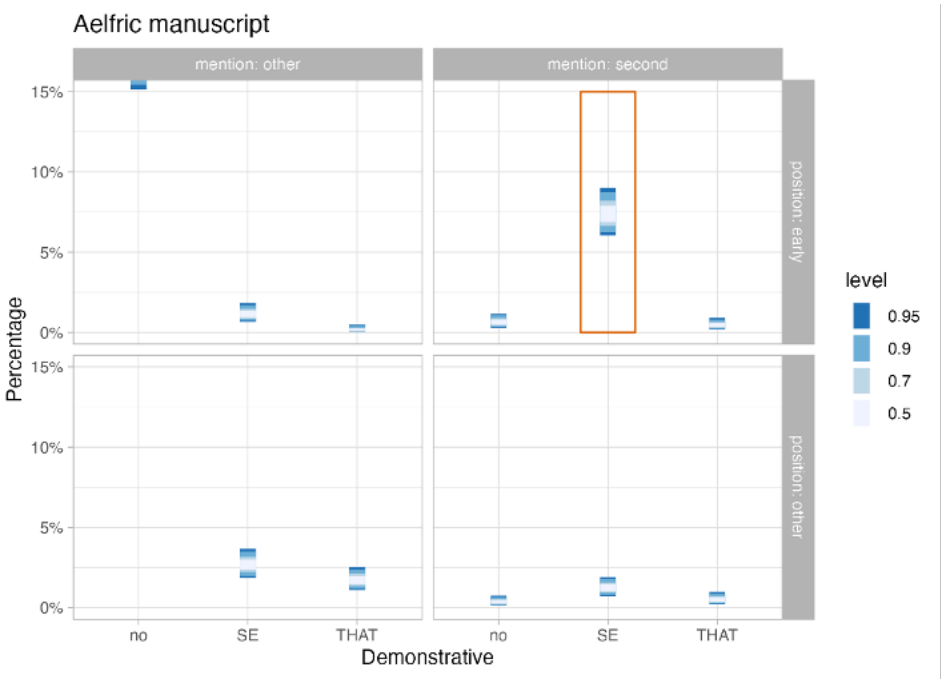


Figure 3: Ælfric's Lives of Saints, Predicted Credible Intervals (50, 70, 90, 95%) of the posterior probability distribution from a Bayesian model fitted to number of occurrences of demonstratives.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

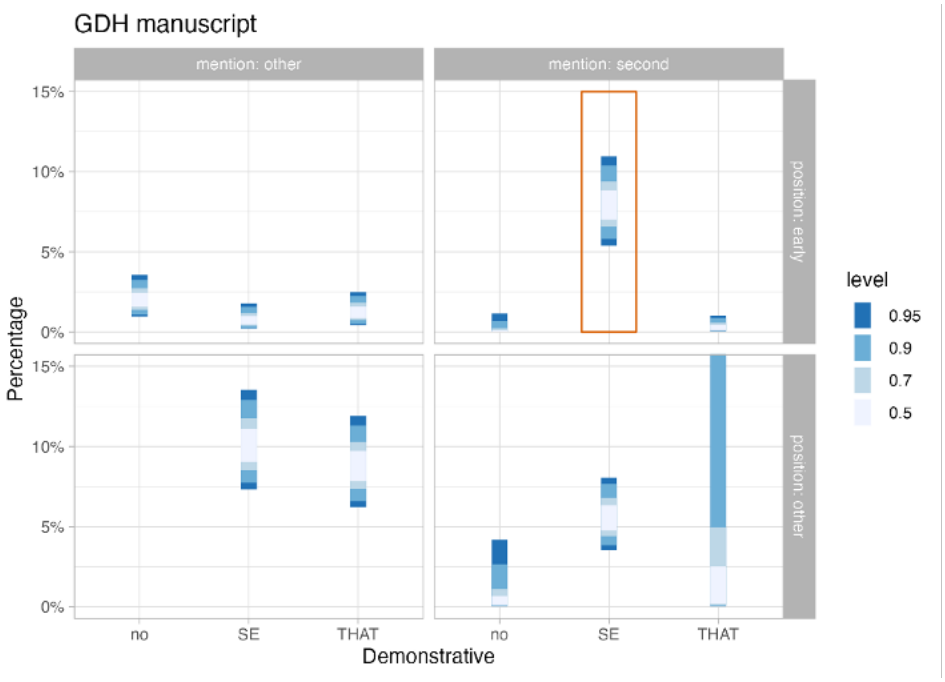


Figure 4: Gregory's Dialogues ms H, Predicted Credible Intervals (50, 70, 90 and 95%) of the posterior probability distribution from a Bayesian model fitted to number of occurrences of demonstratives.

where they do not qualify, strictly speaking, as second mentions. The difference between Ælfric's *Lives of Saints* and the H ms of *Gregory's Dialogues* in how consistently they use *se*-relatives may be due to the fact that the former reflects a more sophisticated style by a very careful author (Ælfric is labelled a "conscious stylist" in Hurst 1972) while the latter is a fairly pedestrian translation of a Latin text, which shows its origins in spite of the fairly substantial editing of the H ms.

5.2 Demonstrative-headed main clauses in Ælfric's *Lives of Saints*

To put these findings in perspective, we looked at *se*-clauses in Ælfric's *Lives of Saints* that are coded as demonstrative-headed main clauses rather than relative clauses in YCOE.[7] There are 43 cases of these demonstrative-headed main clauses that might possibly qualify as meeting most of the conditions for relative-clause-status in that they are not connected to the preceding clause by *and*, and the

[7]Note that here, too, *se*-clauses refer to masculine, feminine and plural demonstratives, and exclude neuter forms of *this* and *that*, and demonstratives as complements of prepositions.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

demonstrative is the first element in the clause. Following Schlachter (2012), who found that capital letters were systematically used in Old High German mss to signal sentence beginnings, including demonstrative-headed main clauses, we investigated whether capitalization might provide any clues to clausal status. Comparing images of the ms made available on the website of the British Library to the edition of Ælfric's *Lives of Saints*, on which the text in YCOE is based (Skeat), it appears that Skeat follows the capitalization of the ms quite faithfully, using it to demarcate the text into what are at times quite large stretches of discourse, paragraphs rather than sentences. The YCOE sentential coding follows this system. An example is (27), where both *se*s are capitalized, and follow a sentence-code; note that the second *se* would not qualify as a "second mention"; instead, it is reactivating the referent of the first *se* after the narrative is interrupted by background information – reminiscent of Schlachter's (2012) finding that antecedents of these clauses can be quite distant, and that the demonstrative-headed clauses can apparently select and reactivate topics, which cannot be achieved by ordinary relative clauses.

(27) *Saul hatte se forma cyning þe ofer Godes folc rixode.*
Saul was.called the first king who over God.GEN people ruled
Se *wæs to cynincge ahafen swyðor for folces gecorennysse*
SE.NOM.M was to king raised rather for people.GEN choice
þonne ðurh Godes ræd. Fela oðre cynincgas rixodon ær
than by God.GEN counsel many other kings ruled earlier
geond ealne middaneard ofer hæðenum leodum, ac ofer
throughout all middle-earth over heathen peoples but over
Israhela folc þe on God belyfde næs nan eorðlic cynincg
Israel.GEN people who in God believed not.was no earthly king
ærðan þe Saul, swa swa hi sylfe gecuron, ofer hi cynerice
before that Saul as as they self chose over them kingdom
underfencg. ***Se*** *beah hrædlice fram þæs ælmihtigan Godes*
received SE.NOM.M bent soon from the almighty God.GEN
willan and nolde be his wissunge and be his witegan lare
will and not.wanted by his guidance and by his prophets.GEN teaching
faran, and se yfela gast hine drehte mid deofollicum sticelsum, and
go and the evil spirit him troubled with diabolic instigations and
on ungewitte his mod awende.
into madness his mind turned
'Saul was the name of the first king who reigned over God's people. He was raised to be king rather by the people's choice than by God's counsel.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

> Many other kings had reigned before throughout the whole world over heathen nations; but over the people of Israel, who believed in God, there was no earthly king before that Saul (as they had themselves chosen) assumed the dominion over them. He turned quickly aside from the will of Almighty God, and would not walk by His instruction and by the teaching of His prophets, and the evil Spirit troubled him with diabolic instigations, and turned his reason into madness.' (tr. Skeat).
>
> [ÆLS (Book of Kings) 1–8]

All of the 23 cases of capitalized *se* in *Lives of Saints* have early verbs; all but two are the first word of a "sententia", which is often close to the beginning of the chapter, as in (27). This makes sense as the protagonist is likely to be introduced there, and once introduced, will be referred to by a form of *se*. This is completely parallel to the "second mention" *se*-relatives with early verbs; the difference is that the preceding chunk in the case of these demonstrative-headed main clauses has been informative enough to warrant "sentence-hood" (cf. Gärtner's observation that the choice between a V2-relative and a demonstrative-headed main clause appears to depend on "information chunking" (see §3.3).

Note that of the two cases where capitalized *se* appears inside a "sententia", one is the first word of direct speech, coded as a standalone demonstrative, like the other capitalized standalone demonstratives, while the other one, (28) below, is coded as a relative clause. Example (28) is part of the following story: A widow realizes that she is "destroying herself by deadly sins" and writes all her wicked deeds on a piece of paper, which she takes to the holy man Basil, imploring him to blot her sins out. He keeps vigil that night to pray for her:

(28) *Git þa Basiliuus gebæd for þæt wif. Waciende þa niht, and þæt*
Yet then Basilius prayed for the woman waking the night and that
gewryt ageaf þam foresædan wife, and þa wæron
writing gave.back the.DAT aforementioned woman.DAT and then were
þa synna ealle adilegode butan anre synna, ***Seo*** *wæs seo mæste.*
the sins all blotted.out except one sin SE.NOM.F was the greatest
'Still Basil prayed for the woman, keeping vigil that night, and gave back the writing to the aforesaid woman, and then were the sins all blotted out, save one of the sins, which was the greatest.'
[ÆLS [Basil]:549.848]

Note that we have arrived at the "central reportable event" (Labov 1972) of the story. In a PDE translation, we would have marked this important turn by

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

"However", to mark the contrast that at first, the saint seems to have succeeded: the slate of the woman's sins has been wiped clean – but it turns out that there is still one sin left. The start of a new sentence rather than just a new clause inserts a pause before the devastating turn that the one sin that is not blotted out also happens to be the most reprehensible one. A demonstrative-headed main clause interpretation seems warranted here, even though it is not coded as such in YCOE.

Two of the remaining 19 clauses are preceded by left-dislocations, as part of a correlative generic expression: *þa ðe fram gode bugon to bysmor-fullum hæðenscype* ***þa*** *wurdon gescynde* 'as for those who turn from God to shameful idolatry, those were put to shame' [ÆLS [Book of Kings] 43.3687]. The remaining 17 are all "second mentions" of referents that have just been introduced – three are expressions like *sum þara* 'some of those' or *an þara* 'one of those' which find their referents in the immediately preceding discourse; eight follow names, four follow indefinite expressions (*sum earm ceorl* 'some poor man' [ÆLS [Swithun]:97.4274]; *ænne blindne mann* 'a blind man' [ÆLS [Denis]:51.5817]; *micel folc manna* 'a great multitude of people' [ÆLS [Martin]:1015:6626] and two follow what we would label definite expressions, although they refer to entities newly brought onto the scene: Christ's holy angel in *Asende ure Hælend Crist his halgan engel mid þe* 'Our Saviour will send his holy angel with you' [ÆLS [Apollinaris]:25.4544] and *an God* 'one God', who is introduced in direct speech in a successful attempt to convert an unbeliever [ÆLS [Cecilia]:163.7213]. These new referents all persist in the subsequent discourse.

These *se*-clauses, then, are clearly of the same form (distal demonstrative as initial element in a V2-clause) and the same function (second mentions, as well as, in a few cases, reactivations) as the *se*-clauses that were coded as relatives. Half of them appear in separate "sententiae" and hence offered clear signs to the annotator of their main clause status, while the other half were preceded by clauses that were apparently informative enough to prompt an analysis of "complete sentence".

6 Discussion

One of the interesting outcomes of the comparisons of V2 relatives with embedded relative clauses is the possible light they cast on the diachrony of how the latter developed out of demonstrative-headed main clauses: the embedded relative shows signs of grammaticalization in that they are more versatile than V2-relatives and demonstrative-headed main clauses in terms of the *de re/de*

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

dicto contrast reported in §3.2 (examples 18a–18b). The pivot here is the change from demonstrative to relative pronoun: the demonstrative has to refer to an entity that is specific, unlike the relative pronoun. If we follow the observation in Roberts & Roussou (2003: 221) about the kind of semantic bleaching which accompanies the grammaticalization process, we can argue that the more general use of the relative pronoun indicates that it is more bleached than the demonstrative. Roberts & Roussou argue that bleaching, often described as the loss of lexical content, would be more accurately described as the loss of non-logical content. Logical content is "insensitive to specific facts about the world" (von Fintel 1995, in Roberts & Roussou 2003: 221); the meaning of a functional element is predictable and stable, i.e. independent of the (lexical) items it scopes over. It follows that functional elements have more general applications than lexical elements; cf. also Lehmann (1995 [1982])'s parameter of "paradigmatic variability" as a metric for quantifying the degree of grammaticalization. The value of the measure increases when restrictions imposed on the contexts in which the functional element can appear are dropped (Lehmann 1995 [1982]: 141), or if there is an increase in its distribution (p. 142). The relative pronoun of a properly embedded relative clause, then, can be argued to be more grammaticalized than the demonstrative.

Another difference between embedded relative clauses and V2 relatives indicating that the former are more grammaticalized is their ability to express backgrounded material, i.e. information that has already been expressed explicitly in the previous discourse (from (Gärtner 2001: 130):

(29) a. A: *Kennst du jemanden, [**der** ein Fahrrad <u>besitzt</u>]?*
b. ??B: *Ja, ich KENNE jemanden, (/) [**der** <u>besitzt</u> ein Fahrrad].*

(29b) is fine for embedded relative clauses, with the answer just echoing the question, but degraded for V2-relatives (Gärtner 2001: 130). This observation also points to the greater versatility of embedded versus V2-relatives (and demonstrative-headed main clauses).

The embedded relative clause is part of a general development along a cline of clause linking, with two maximally elaborated paratactic clauses at one end and a single clause a compressed, embedded complement at the other end (Lehmann 1988). Movement along this cline is not necessarily unidirectional: although Lehmann's examples show movement towards compression (involving his parameters of hierarchical downgrading, the main clause syntactic level, desentialization and explicitness of linking), nonfinite/infinitival clauses, which are positioned towards the compressed end of the cline on all parameters, are much more

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

likely to represent "upgraded" nominalizations (Los 2007). In the case of relative clauses, there is a consensus in the literature that we are dealing with "downgrading". For OE, Mitchell (1985: 261) specifically points to the V2-position of the verb as a main-clause feature inherited from the paratactic origin of the construction. Parataxis also accounts for the requirement that the V2-relative must be in clause-final position (see §3.2). The frequency of such "extraposed" relatives shows a diachronic decline in a number of European languages (Wallenberg 2016), suggesting that relative clauses, earlier paratactically linked to the preceding clause, become more integrated into that clause over time; the V2-relatives in Dutch, German and OE might reflect an intermediary stage in the process of parataxis to hypotaxis, poised between restriction (an embedded function) and foregrounding (a main clause function). It is typical for grammaticalization that both the original construction as well as any intermediary stages may coexist with the maximally grammaticalized construction ("layering", Hopper & Traugott 2003: 106).

Cases like (29a-b) above are discussed by Gärtner as part of a discussion "assertionality" (Gärtner 2001: 125-131) – as is well-known, main and subclauses in language generally tend to align along the assertion/foregrounding versus the presupposition/backgrounding axis (see e.g. Cristofaro 2003). Whether, and how, the alignment is marked overtly (by word order like V2, by indicative versus subjunctive, or by other means) will depend on diachrony, and may be the result of competition between which function wins out as the first item in the construction of utterances – assertion (marking realis versus irrealis) or foregrounding (marking the main event of the plot line out from any embedded events). Assertion-first may provide reasons or explanations for a foregrounded event embedded under the assertion, even though reasons and explanations are background in terms of discourse structure; and the natural rapid churn of narrative devices to signal peak marking, episode boundary marking, or the creation of suspense are likely to lead to the development and subsequent decline of a multitude of expressions that prioritise one alignment over the other, so that either foregrounding or assertion ends up being relegated to syntactically embedded positions (Los 2015: 216-230).

If main clause assertion is used as a frame for a particular narrative device, the foregrounded event follows in an embedded clause. Examples are episode boundary marking in Middle English by means of the frame *It befell that* (Brinton 1996), cf. also *hit gelamp* 'it happened' in (23) above; or the use of the *it*-cleft, specifically in the frame *It was then that*, to mark "the central reportable event" in a story (Los 2015: 219, 242). A third example is sentential relatives being used as foregrounding devices (Huddleston & Pullum 2002: 1148). A preponderance of

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

one alignment over the other might in time lead to a particular kind of marking (see e.g. Gärtner & Eythórsson 2020, reported in Coniglio & Hinterhölzl 2020: 227).

Syntactic developments may obscure established patterns of marking main and subclause asymmetries from one period to the next – in the history of English, for example, finite subjunctive clauses expressing "dependent desires" (the complements of verbs with meanings of fear, promise, order, hope, expect and the like) were ousted by non-finite (*to*-infinitival) clauses (Los 2005), obscuring any alignments that may have been established between assertions and indicatives or V2, versus irrealis and subjunctives or verb-final. An example of a pocket of misalignment in PDE are conditional clauses where C can be expressed not only by a "bespoke" element like *if* but also by finite verb movement to C, now restricted to a dwindling set of lexemes (the auxiliaries *had* or *should*; Roberts & Biberauer 2017) – a case where finite verb movement to C marks an embedded clause.

It is also a common finding that subclauses tend to preserve older orders, whereas main clauses tend to innovate; the reason appears to be that main clauses have to satisfy various communicative requirements, the positioning of focus and discourse-old or discourse-new material, and they therefore tend to develop special constructions not found in the subclause (see Bybee 2001). The finite verb may possibly have functioned as a focus marker first, as still in Hungarian (Comrie 1997: 63), and may later have become entrenched as a clause-typing device. Any asymmetry developing between main and subclause in terms of whether or how the clause-typing head, say C, is expressed is probably a secondary or epiphenomenal development.

7 Conclusion

This paper has investigated whether more can be made of the observation in the literature that a certain type of relative clause, i.e. the *se*-relative, is found more often with main clause word order (in terms of finite verb movement to C or to a lower head) than other types, particularly the *þe*-relative. The latter also shows other signs of being more integrated into the previous clause than the *se*-relative. We have argued that the *se*-relative in OE exhibits similarities with the "V2-relatives" in modern Dutch and German, although some of the clauses coded as "relative" in the OE corpus can also be analysed as demonstrative-headed main clauses. As verb position is not a clear diagnostic of clausal status in OE, we looked at a discourse function, i.e. whether or not the *se*-clause expresses a "second mention" of a main protagonist. Both texts showed the same configuration

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

of *se*-clause, second mention and V2-order as a coherent package, although the use of the V2 *se*-relatives as a "second mention" device is most consistent in Ælfric's *Lives of Saints*, rather than in the OE translation of *Gregory's Dialogues* (ms H)).

There are many issues that could not be addressed in the scope of this paper; as avenues for further research, the analysis of V2-*se*-relatives in other OE texts would be a next step. Another issue that is outstanding is how the function of restriction operates in *se*-relatives and in *se*-headed main clauses, and how it interacts with the foregrounding function of the *se*-clause. In both *se*-relative and in *se*-headed main clause, the topic-shifting function of *se* coerces identification with a new referent who has been introduced in the previous clause; although indefinite, the referent is already specific when first introduced, unlike the type of antecedent usually associated with restrictive relative clauses; there is no sense of the *se*-clause identifying a subset from a larger set. In §3.3, we saw that Gärtner (2001) identifies the function of V2-relatives as restriction, although the *se*-relative in OE is said to be mainly non-restrictive in the literature (see §2.4). This observation might be explained by the fact that they often follow proper names in OE (as in (i) in footnote 5), even if those names cannot be said to be the antecedent (which in (i) would be *sum wif* 'a woman'), although the antecedent in all of these cases is a specific entity. Cf. also the discussion in Den Dikken (2003) about a Dutch example where the first clause translates as *There were two boys on the beach*, where the question is whether this entails that there were only two boys on the beach, or more, and the same question might be asked about the two sons in (15). The PDE counterparts of the "second mention" relatives capable of performing required the topic-shifting function would be non-restrictive relative clauses (cf. for instance the discussion of the Columbo-example (10) above), which are in fact typical foregrounding devices in PDE. A provisional conclusion might be that restriction is just not relevant here (see also Denison & Hundt 2013), and that the primary function of V2-*se*-relatives is presentational.

Acknowledgements

References

Allen, Cynthia L. 1977. *Topics in diachronic English syntax*. Amherst: University of Massachusetts. (Doctoral dissertation).

Allen, Cynthia L. 2020. Case and preposition stranding in Old English free relatives. *NOWELE* 73(2). 193–200. DOI: 10.1075/nowele.00039.all.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Allen, Cynthia L. 2022. Pronominally headed relative clauses in early English. *English Language and Linguistics* 26(1). 105–132. DOI: 10 . 1017 / S136067432100006X.

Brandt, Margareta. 1990. *Weiterführende Nebensätze*. Stockholm: Almqvist & Wiksell.

Breban, Tine. 2012. Functional shifts and the development of English determiners. In Anneli Meurman-Solin, María José López-Couso & Bettelou Los (eds.), *Information structure and syntactic change in the history of English*, 271–300. Oxford: Oxford University Press.

Brinton, Laurel J. 1996. *Pragmatic markers in English: Grammaticalization and discourse functions*, vol. 19 (Topics in English Linguistics; 19). Berlin: De Gruyter. DOI: 10.1515/9783110907582.

Bybee, Joan. 2001. Main clauses are innovative, subordinate clauses are conservative: Consequences for the nature of constructions. In Joan Bybee & Michael Noonan (eds.), *Complex sentences in grammar and discourse: Essays in honor of Sandra A. Thompson*, 1–17. Amsterdam: Benjamins.

Clemoes, Peter. 1956. *Ælfric's Catholic Homilies*. University of Cambridge. (Doctoral dissertation).

Comrie, Bernard. 1997. Pragmatic binding: Demonstratives as anaphors in Dutch. In Matthew L. Juge & Jeri L Moxley (eds.), *Proceedings of the Twenty-Third Annual Meeting of the Berkeley Linguistics Society*, 50–61. Berkeley Linguistics Society.

Coniglio, Marco & Roland Hinterhölzl. 2020. Discourse conditions on relative clauses: A crosslinguistic and diachronic study on the interaction between mood, verb position and information structure. In Anke Holler, Katja Suckow & Israel de la Fuente (eds.), *Information structuring in discourse*, 201–233. Leiden: Brill.

Cristofaro, Sonia. 2003. *Subordination* (Oxford Studies in Typology and Linguistic Theory). Oxford: Oxford University Press.

Den Dikken, Marcel. 2003. A comment on the topic of topic-comment. *Lingua* 115. 691–710. DOI: 10.1016/j.lingua.2003.11.005.

Denison, David & Marianne Hundt. 2013. Defining relatives. *Journal of English Linguistics* 41. 135–67. DOI: /10.5167/uzh-78238.

Depraetere, Ilse. 1996. Foregrounding in English relative clauses. *Linguistics* 34. 699–731. DOI: 10.1515/ling.1996.34.4.699.

de Vries, Mark. 2012. Parenthetical main clauses – or not? On appositives and quasi-relatives. In Lobke Aelbrecht, Liliane Haegeman & Rachel Nye (eds.), *Main clause phenomena: New horizons* (Linguistics Today 190), 177–202. Amsterdam: Benjamins.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Emonds, Joe E. 1970. *Root and structure-preserving transformations.* Massachusetts Institute of Technology. (Doctoral dissertation).

von Fintel, Kai. 1995. The formal semantics of grammaticalization. In *North East Linguistics Society*, vol. 25(2), 175–189.

Fischer, Olga, Ans van Kemenade, Willem Koopman & Wim van der Wurff. 2000. *The syntax of Early English.* Cambridge: Cambridge University Press.

Gärtner, Hans-Martin. 2001. Are there V2 relative clauses in German? *Journal of Comparative Germanic Linguistics* 3. 97–141. DOI: 10.1023/A:1011432819119.

Gärtner, Hans-Martin & Thórhallur Eythórsson. 2020. Varieties of dependent Verb Second and verbal mood: A View from Icelandic. In Rebecca Woods & Sam Wolfe (eds.), *Rethinking verb second*, 208–239. Oxford: Oxford Scholarship Online.

Green, Georgia M. 1976. Main clause phenomena in subordinate clauses. *Language* 52. 382–397. DOI: 10.2307/412566.

Hecht, Hans. 1965. *Bischof Wærferth von Worcester Übersetzung der Dialoge Gregors des Grossen.* Darmstadt: Wissenschaftliche Buchgesellschaft.

Heycock, Caroline. 2017. Embedded root phenomena. In Martin Everaert & Henk C. van Riemsdijk (eds.), *The Wiley Blackwell companion to syntax*, 1–37. Hoboken: Wiley & Sons.

Hooper, Joan B. & Sandra A. Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4(4). 465–497.

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

Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language.* Cambridge: Cambridge University Press.

Hurst, James. 1972. *Ælfric.* New York: Twayne.

Jasinskaja, Katja & Claudia Poschmann. 2020. Projection to the speaker: Nonrestrictive relatives meet coherence relations. In Anke Holler, Katja Suckow & Israel de la Fuente (eds.), *Information structuring in discourse*, 141–162. Leiden: Brill.

van Kemenade, Ans. 1987. *Syntactic case and morphological case in the history of English.* Dordrecht: Foris.

Koopman, Willem. 1992. Old English clitic pronouns: Some remarks. In Frances Colman (ed.), *Evidence for Old English* (Edinburgh Studies in the English language; 2). Edinburgh: John Donald.

Labov, William. 1972. *Language in the inner city.* Philadelphia: University of Pennsylvania Press.

Ladd, D. Robert. 1986. Intonational phrasing: The case for recursive prosodic structure. *Phonology Yearbook* 3. 311–340. DOI: 10.1017/S0952675700000671.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Lehmann, Christian. 1988. Towards a typology of clause linkage. In John Haiman & Sandra A. Thompson (eds.), *Clause combining in grammar and discourse*, 181–225. Amsterdam: Benjamins.

Lehmann, Christian. 1995. *Thoughts on grammaticalization (LINCOM studies in theoretical linguistics 1)[1982])*. München: LINCOM Europa.

Los, Bettelou. 2005. *The rise of the to-infinitive*. Oxford: Oxford University Press.

Los, Bettelou. 2007. *To* as a connective in the history of english. In Ursula Lenker & Anneli Meurman-Solin (eds.), *Connectives in the history of English* (Current Issues in Linguistic Theory; 283), 31–60. Amsterdam: Benjamins.

Los, Bettelou. 2015. *A historical syntax of English*. Edinburgh: Edinburgh University Press.

Los, Bettelou, Gea Dreschler, Ans. van Kemenade, Erwin Komen & Stefano Coretta. 2023. The decline of local anchoring: A quantitative investigation. *English Language & Linguistics* 27(2). 345–372. DOI: 10.1017/s1360674323000047.

Los, Bettelou & Ans van Kemenade. 2018. Syntax and the morphology of deixis: The loss of demonstratives and paratactic clause linking. In Marco Coniglio, Andrew Murphy, Eva Schlachter & Tonjes Veenstra (eds.), *Atypical demonstratives: Syntax, semantics and pragmatics*, 127–60. Berlin: De Gruyter. DOI: 10.1515/9783110560299-005.

Meklenborg Salvesen, Christine & George Walkden. 2017. Diagnosing embedded V2 in Old English and Old French. In Eric Mathieu & Robert Truswell (eds.), *Micro-change and macro-change in diachronic syntax*, 168–181. Oxford: Oxford University Press.

Mitchell, Bruce. 1959. *Subordinate clauses in Old English poetry*. Merton College, University of Oxford. (Doctoral dissertation).

Mitchell, Bruce. 1985. *Old English syntax*. Oxford: Clarendon Press.

Miura, Ayumi. 2016. Reassessing Latin influence on he/she this in Middle English. *English Studies* 97(4). 420–441. DOI: 10.1080/0013838X.2016.1138694.

Mustanoja, Tauno F. 1960. *A Middle English syntax*. Part I: Parts of Speech. Helsinki: Société Néophilologique.

Parkes, Malcolm B. 1992. *Pause and effect: An introduction to the history of punctuation in the west*. Aldershot: Scolar Press.

Pintzuk, Susan. 1999. *Phrase structures in competition: Variation and change in Old English word order*. New York: Garland Publishing.

Randall, Beth. 2005. CorpusSearch 2. Computer program. Designed by Beth Randall, Ann Taylor and Anthony Kroch. http://corpussearch.sourceforge.net/credits.html.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Roberts, Ian & Theresa Biberauer. 2017. Conditional inversion and types of parametric change. In Bettelou Los & Pieter de Haan (eds.), *Word order change in acquisition and language contact: Essays in honour of Ans van Kemenade*, 57–78. Amsterdam: Benjamins.

Roberts, Ian G. & Anna Roussou. 2003. *Syntactic change: A minimalist approach to grammaticalization*. Cambridge: Cambridge University Press.

Roma, Elise. 2007. Relativisation strategies in insular Celtic languages. In Paola Ramat & Elise Roma (eds.), *Europe and the mediterranean as linguistic areas: Convergencies from a historical and typological perspective*, 245–288. Amsterdam: Benjamins.

Sánchez Argüelles, Gerardo & Bettelou Los. 2022. Incipient articles in Old East Scandinavian varieties. *NOWELE* 75(2). 160–193. DOI: 10.1075/nowele.00067.san.

Schlachter, Eva. 2012. *Syntax und Informationsstruktur im Althochdeutschen: Untersuchungen am Beispiel der Isidor-Gruppe*. Heidelberg: Winter.

Skeat, Walter William. 1966. *Ælfric's Lives of Saints*. Oxford: Oxford University Press.

Suárez-Gómez, Cristina. 2012. Clause linkage across time and genres in early english: A preliminary approach to relative clauses. *Studia Neophilologica* 84(2). 138–50. DOI: 10.1080/00393274.2012.733200.

Taylor, Ann. 2014. Syntax. In Don Ringe & Ann Taylor (eds.), *The development of Old English*, 1–2. Oxford: Oxford University Press.

Taylor, Ann, Anthony Warner, Susan Pintzuk & Frank Beths. 2003. *The York-Toronto-Helsinki parsed corpus of Old English Prose (YCOE)*.

Traugott, Elizabeth C. 1992. Syntax. In Richard M. Hogg (ed.), *The Cambridge history of the English language*, vol. I, 207–408. Cambridge: Cambridge University Press.

van Kemenade, Ans. 1997. V2 and embedded topicalization in Old and Middle English. In Ans van Kemenade & Nigel Vincent (eds.), *Parameters of morphosyntactic change*, 326–352. Cambridge: Cambridge University Press.

Wagener, Terje. 2017. *The history of Nordic relative clauses*, vol. 304 (Trends in Linguistics. Studies and Monographs [TiLSM]). Berlin: De Gruyter. DOI: 10.1515/9783110496536.

Wallenberg, Joel. 2016. Extraposition is disappearing. *Language* 92(4). 237–56. DOI: 10.18148/hs/2016.v0i1.15.

Wende, Fritz. 1915. *Über die nachgestellten Präpositionen im Angelsächsischen* (Palaestra 70–73). Berlin: Mayer & Müller.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Yerkes, David. 1982. *Syntax and style in Old English: A comparison of the two versions of Wærferth's translation of Gregory's Dialogues*. Binghamton: Center for Medieval & Early Renaissance Studies.

Zwart, Jan-Wouter. 2003. *Iets over zgn. V2 relatieven in het Nederlands*. Ms University of Groningen.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Discourse structure and narration

The volume Discourse structure and narration: A diachronic view from Germanic deals with questions of information structuring at discourse level, focusing on narrative discourses. More precisely, it is about the contribution of grammatical devices to the organization of texts as well as their diagnostic potential for the narrative text type. Although it is well-known that information packaging had a much greater impact on the distribution of grammatical patterns in historical stages of a language than it does today, so far many studies on the relationship between information structure and grammatical patterns do not go beyond the sentence level, in other words, they do not take into account the possible influence of the text type on the manifestation of certain grammatical patterns. How and to which degree changes in grammatical patterns correlate or are affected by changes in either discourse and/or narrative structure, how the two layers interact with each other and affect each other, and how such issues can be operationalized are still understudied. This volume aims to shed more light on these issues by presenting eight papers, which address these questions more or less explicitly. As the research questions imply, the papers all take a historical or diachronic perspective. Another commonality between the studies is that they all focus on data from Germanic languages, as we assume that by comparing closely related languages, the relationships in question become more pronounced. Specifically, the languages in question are German, Dutch, English and Icelandic. Understandably, the contributions in this volume can only highlight some aspects of the complex relationship between grammar and narration(s). Addressing among others questions of narrative progression, temporal structure, reference tracking and discourse functions, the contributions discuss phenomena such as temporal adverbials at the left periphery as well as later in the clause, left dislocation structures, fronting of the finite verb in dependent and independent clauses, linguistic means to express aspectual and tense information, and the distribution of nominalization patterns across text types.