

Chapter 6

In consideration of age and genre: A diachronic corpus analysis of deverbal nouns in German verbonominal constructions and complex prepositions

Fabian Fleissner

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



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, Möslin 1981, Admoni 1973, 1990, von Polenz 1987, 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

¹This 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).

²In 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.

³Throughout the paper, examples from the DTA-corpus are cited according to the following pattern: [author, year].

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 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

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 the forerunners of this protracted and laborious transition to German

as the language of science. Leibniz is the one 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 internationally renowned French or adherence to Latin. In Italy, the transition from Latin to the “vernacular” as a language of science was already gradually taking place in the 16th century, and in France and England in the 17th century, academic writers in Germany predominantly retained 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 the 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.

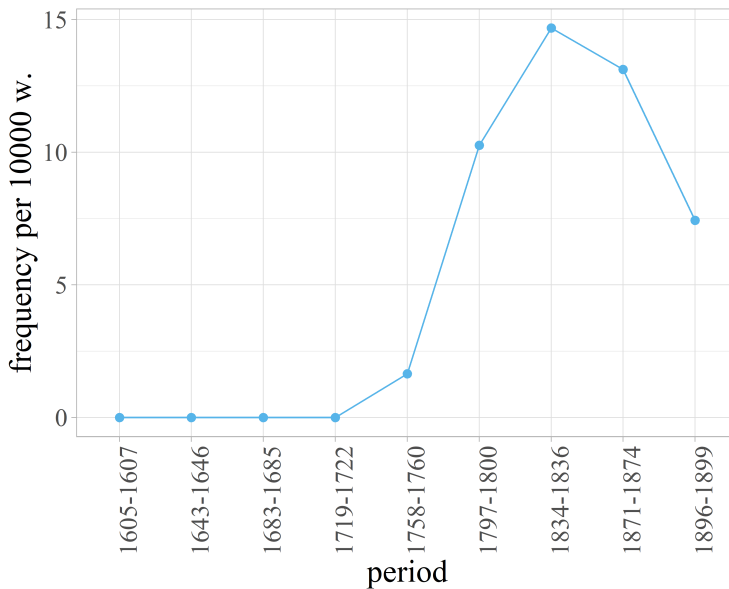


Figure 1: CPs of the pattern [in N P/GEN]; N = *Hilfe, Beihilfe, Bezug, Beziehung, Hinsicht, Rücksicht*

In the present study, I will show that the history of the two structures under investigation is related to the addressed extralinguistic developments. I argue

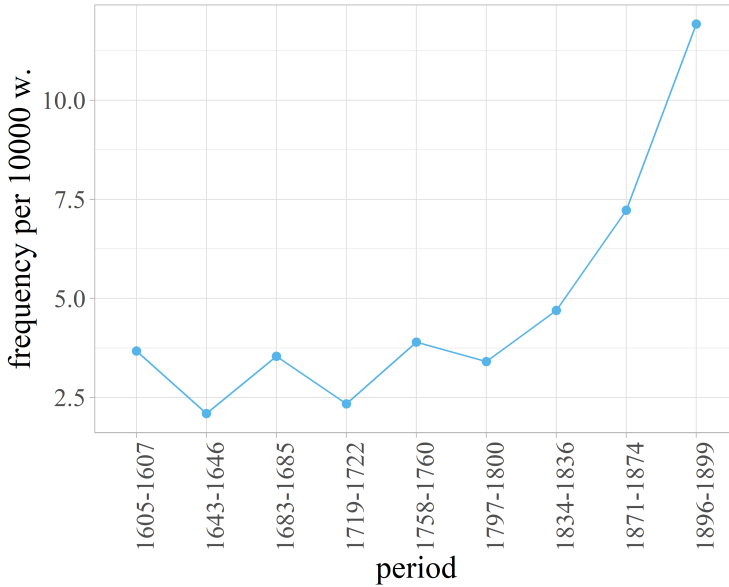


Figure 2: VNCs of the pattern [*in N kommen*]

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 Section 2, I will introduce my research objects and theoretical background. Section 3 presents data and methods. Section 4 presents the results of the corpus analysis which are discussed in Section 5. Section 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 various 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. 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 VNCs in German are usually conceived of as an expanding open class. As mentioned in Section 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, 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)

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 has 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).

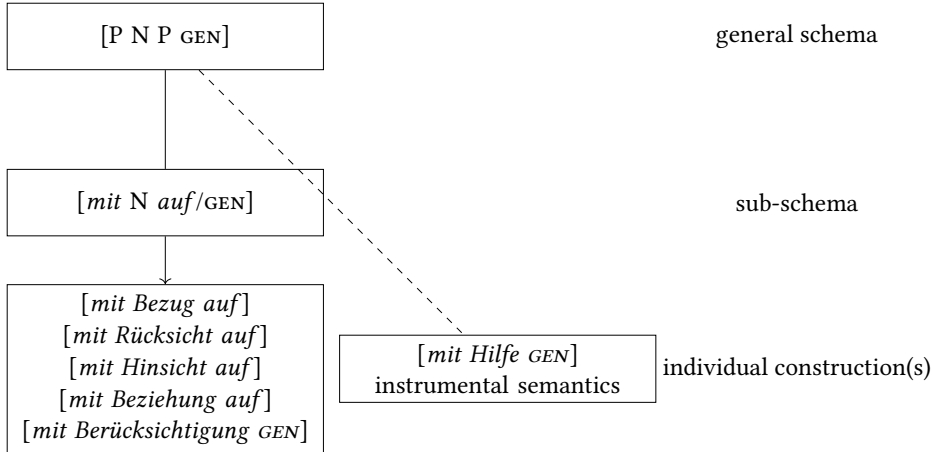


Figure 3: Partial network of German complex prepositions (Ruf & Smirnova forthcoming)

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* ‘assistance’ or *Unterstützung* ‘support’, 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 for example (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 this research project, I 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 the 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, I 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 not only makes up the largest proportion of the corpus but 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 AGE and GENRE are concluded and discernible in the

⁴I restrict myself 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].

⁵A complete list of the nouns as well as the data used for this study can be found at <https://osf.io/yancd>.

data. A comprehensive overview across all time periods would obscure the focus on the discursive functional characteristics of individual constructions, as written language patterns also undergo change. 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 a verb significantly, which is not surprising, 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 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 Section 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 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

^aThese are mainly nominalized infinitives whose age are difficult to determine.

The mere comparison of absolute frequencies may not be 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 of 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 *Wissenschaft* ‘science’.

The absolute distribution already suggests that, as anticipated, both constructions predominantly occur 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.

⁶Reference 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.

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

4 Analysis

In this section, I will present and discuss the results of the corpus study based on the data outlined in Section 3. I will initially provide the findings for the parameter AGE (Section 4.1) and subsequently for GENRE (Section 4.2), before proceeding to the synthesis in Section 4.3. Initially, the focus lies on conducting statistical tests for the respective sub data sets to elucidate the distributional patterns hinted at. For this purpose, Pearson Residuals⁷ 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.

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 independence 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.

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,

⁷Pearson Residuals are included in the results obtained from the `chisq.test()` function in R (R Core Team 2023).

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

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 expectation. Conversely, blue tiles represent negative residuals, signaling scenarios where observed frequencies fall below expectation. 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.

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

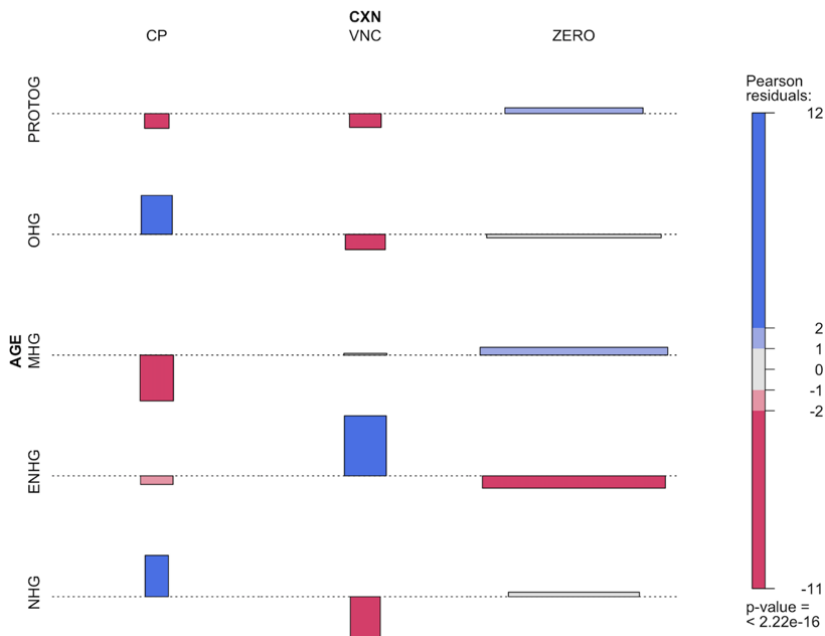


Figure 4: CXN by AGE in an extended association plot

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 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.

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

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

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 of the constructions CP and VNC exhibit a slight tendency towards narrative texts. Figure 5 shows the association plot for GENRE.

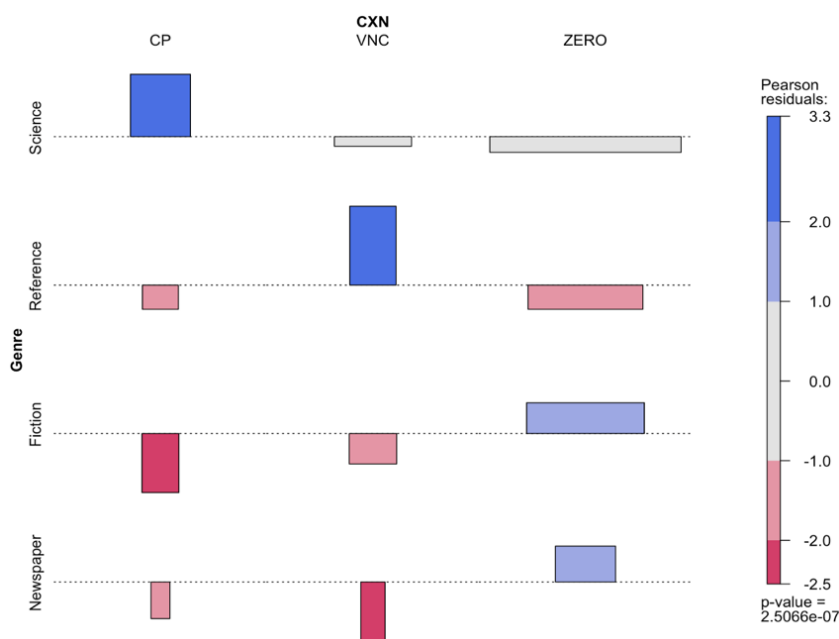


Figure 5: CXN by GENRE in an extended association plot

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 is 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 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 shed light on the role of ENHG and NHG nouns 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. This method 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 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 are ENHG).

For CPs, all the nouns that appeared only once in the data set were 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.

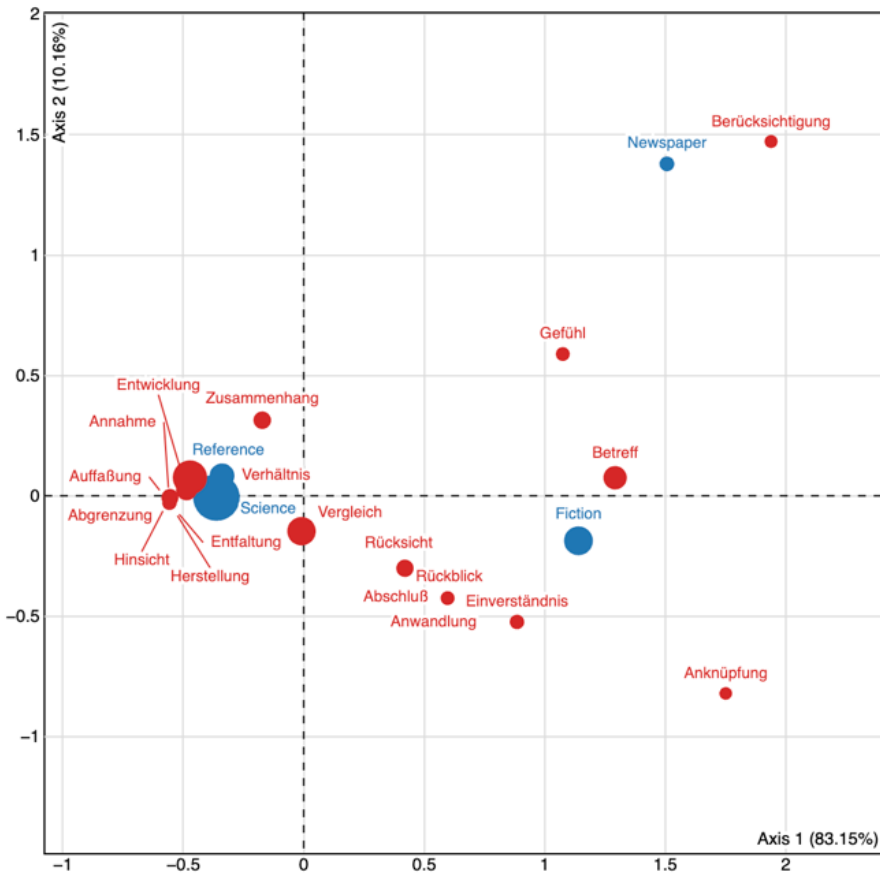


Figure 6: NHG nouns by GENRE in a CA plot (CP), see Table 9

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 con-

nections between different elements or concepts. They are employed to provide context, clarify relationships between ideas under discussion, or emphasize their relevance. 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 (6–8).

- (6) *Ich bezweifle, ob man die Vergangenheit **in ihrem Verhältnis zur Gegenwart** überhaupt richtig zu beurteilen vermag.*
'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)

This subgroup of nouns behaves similarly with regard to the concept of RELATION and suggest a thematic connection associated with analyzing relationships of different subjects.

A second major subgroup can be distinguished 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

⁸The 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 position in the middle of the plot. This illustrates the limitations of exploratory statistical methods, as they cannot account for the individual habits of specific authors. Unsurprisingly, Otto von Bismarck leans towards a formal style even in his personal writings, reflecting his accustomed political demeanor.

‘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 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.

It can be concluded 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.

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,

Table 7: Semantic groups within the general pattern [in N_{dev} P/GEN]

group subgroup	relation	reference
		consideration
NOUNS	<i>Verhältnis, Vergleich, Entwicklung, Zusammenhang, Abschluss, Entfaltung</i>	<i>Betreff, Rücksicht, Hinsicht, Auffassung, Einverständnis, Rückblick, Berücksichtigung, Anknüpfung</i>
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 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 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 7 shows the distribution of the 20 most frequent OHG, MHG, and ENHG nouns in relation to their genre affinity.

This distribution pattern presents a less structured image compared to NHG nouns. A clear distinction between narrative and non-narrative genres, as seen 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

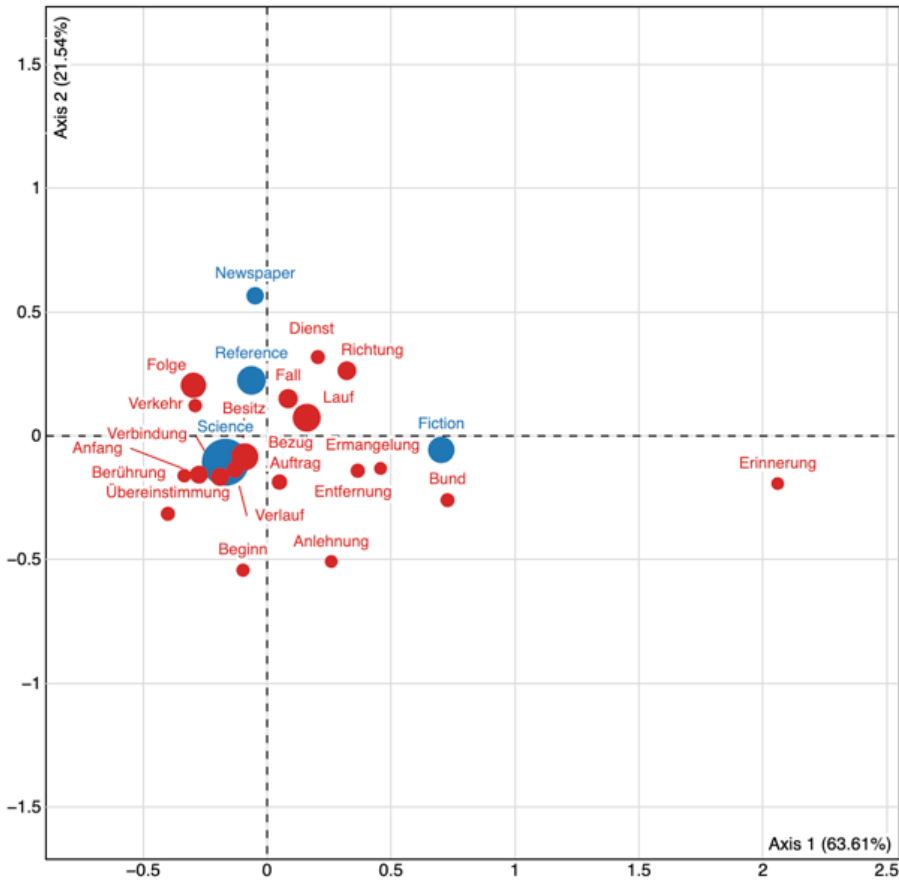


Figure 7: PROTOG, OHG, MHG and ENHG nouns by GENRE in a CA plot (CP), see Table 10

the structure [*in* N_{dev} _] only from the 18th century on, see Figure 8. It seems to be a general trend in German that *ung*-derivations are increasingly 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 the younger word *Beziehung* ‘relation’, but also why *Lauf* ‘course’ prevails over *Verlauf* ‘course’. These two frequent lexemes are responsible for the striking association of CPs with OHG nouns, which is shown in Figure 4. 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 along *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.

"in Bezug" – Verlaufskurve

Basis: DTA-Gesamt+DWDS-Kernkorpus

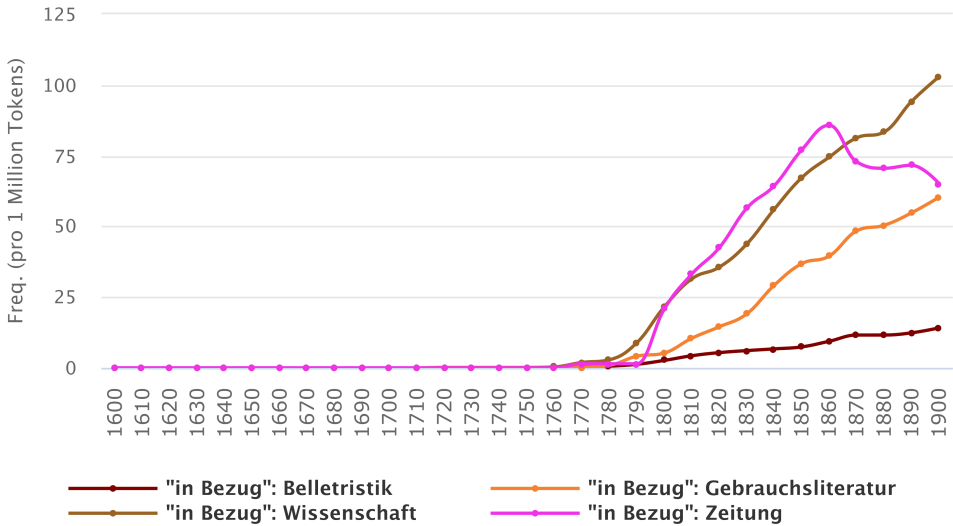


Figure 8: Token frequency of [*in* Bezug _] in the DTA Kernkorpus

When examining VNCs and their favored ENHG nouns with the highest frequency, two observations can be made. Firstly, the token count significantly exceeds that of CPs, as shown in Figure 4. 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.

The most significant difference 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 tendency 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 Verfuche, denselben mit der modernen Entwicklungslehre in Einklang zu bringen, sind völlig fehlgeschlagen.*

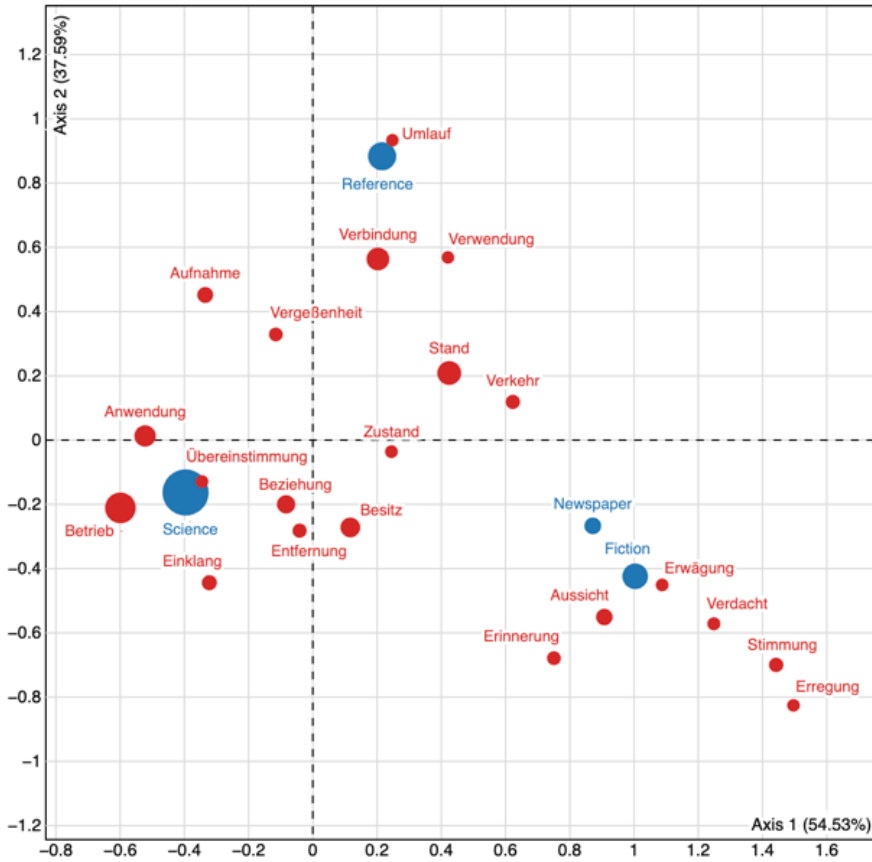


Figure 9: ENHG nouns by GENRE in a CA plot (VNC), see Table 11

‘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 flammt.*

‘In fact, **he has considered** the place from where this weathervane comes.’ (Fontane, 1899)

As can be seen in 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 typically lies 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).

- (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 in 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 of 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*

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	<i>Verbindung, Beziehung, Übereinstim- mung</i>	<i>Aussicht, Einklang, Erinnerung, Vergessenheit, Verdacht, Erwägung</i>	<i>Betrieb, Anwendung, Drehung, Verkehr, Umlauf, Verwendung, Verbindung</i>	<i>Stand, Besitz, Aufnahme, Entfernung, Zustand</i>		<i>Stimmung, Zustand, Erregung, Beziehung, Verbindung, Stand</i>
CHARACTERISTICS	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	
CONSTRUCTION	[in N _{dev} V]					

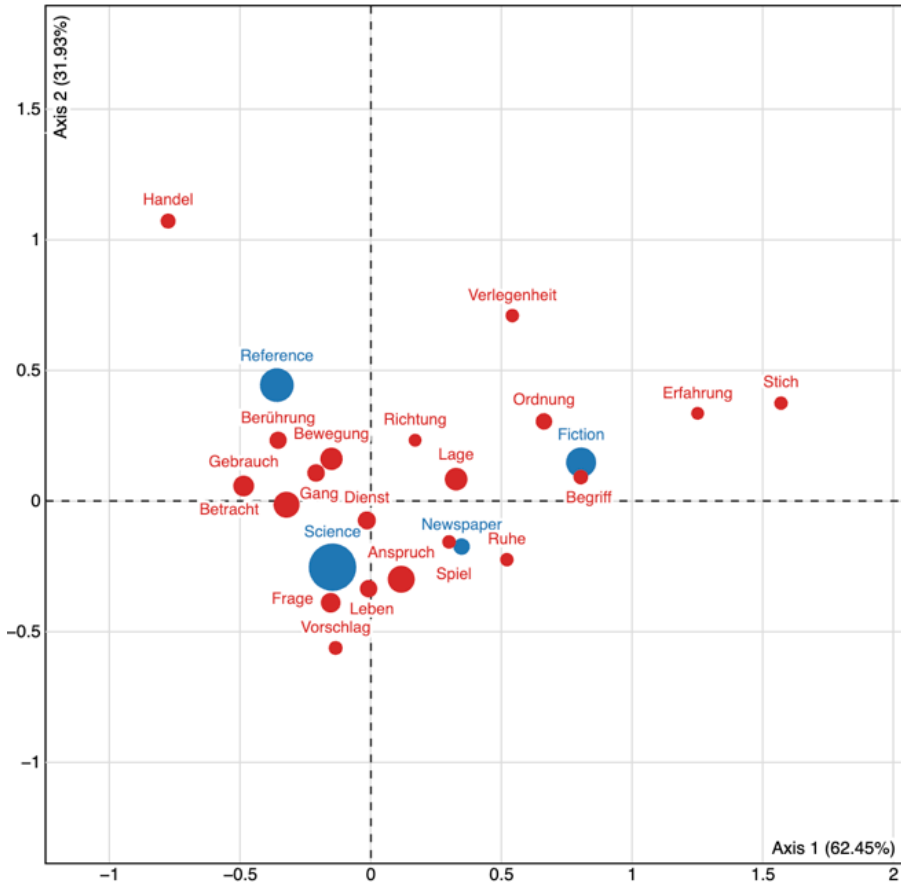
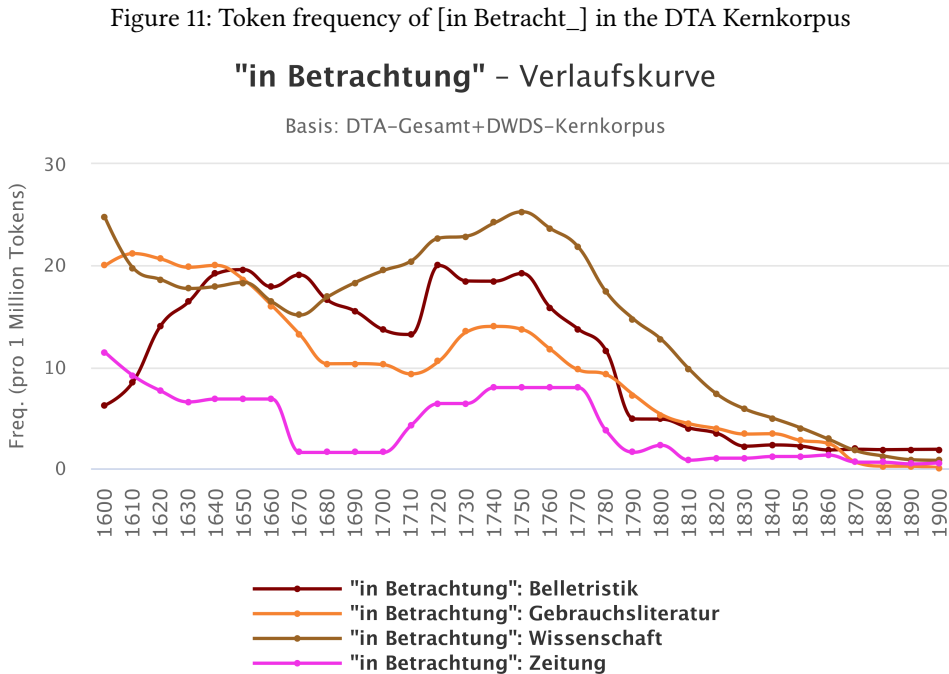
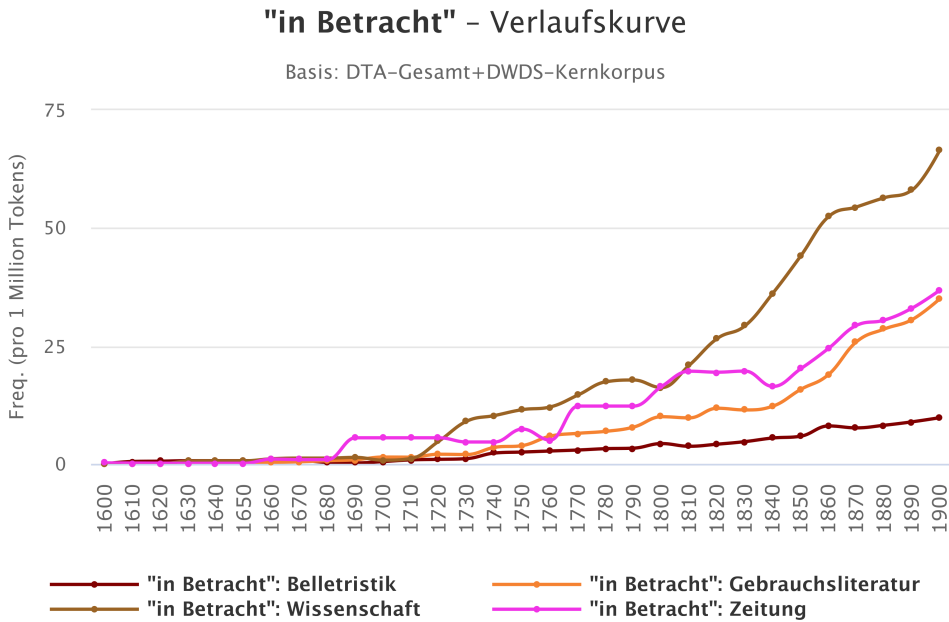


Figure 10: PROTOG, OHG, MHG and NHG nouns by GENRE in a CA plot (VNC), see Table 12

‘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 established 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, the aforementioned tendency to shorten is found. 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 there was already an established structure that was additionally fueled by the increased availability of nominal material in Early New High German.

5 Discussion

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 backgrounding 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 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

⁹In this case, [*in Betrachtung* _] and [*in Betrachtung* V] can be considered almost identical, as the structure is formed almost exclusively with a verb.

to continue operating within this constructional framework, one 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 13 shows the partial network of such a REFERENCE discourse schema.

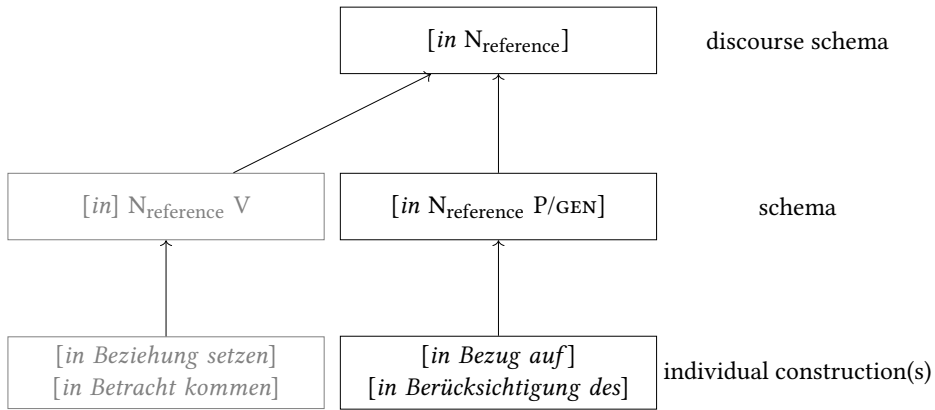


Figure 13: Partial network of $[in N_{dev} \dots]$

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 much more diverse and freer syntactic patterns are at play 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 the established structure to draw from. Figure 14 shows the partial network for VNCs.

It is evident that VNCs do not exhibit uniform discourse-specific behavior. The two identified and frequency dominant sub-schemas USAGE and REFERENCE re-

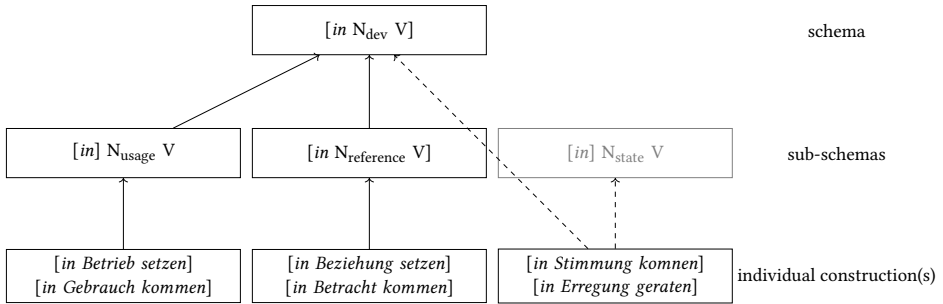


Figure 14: Partial network of German verbonominal constructions [in N_{dev} V]

flect 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 occurring outside of the structures under study. This is shown by a final look at the distribution of the 20 highest frequency ZERO nouns, see Figure 15.

The triangulated system, already encountered with the VNCs in Figure 9, appears again. 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 the typical null model in this genre, but CPs of the pattern [in N_{dev} P/GEN] deviate significantly.

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 initial observation mentioned that German has developed a strong tendency towards nominal style. This phenomenon

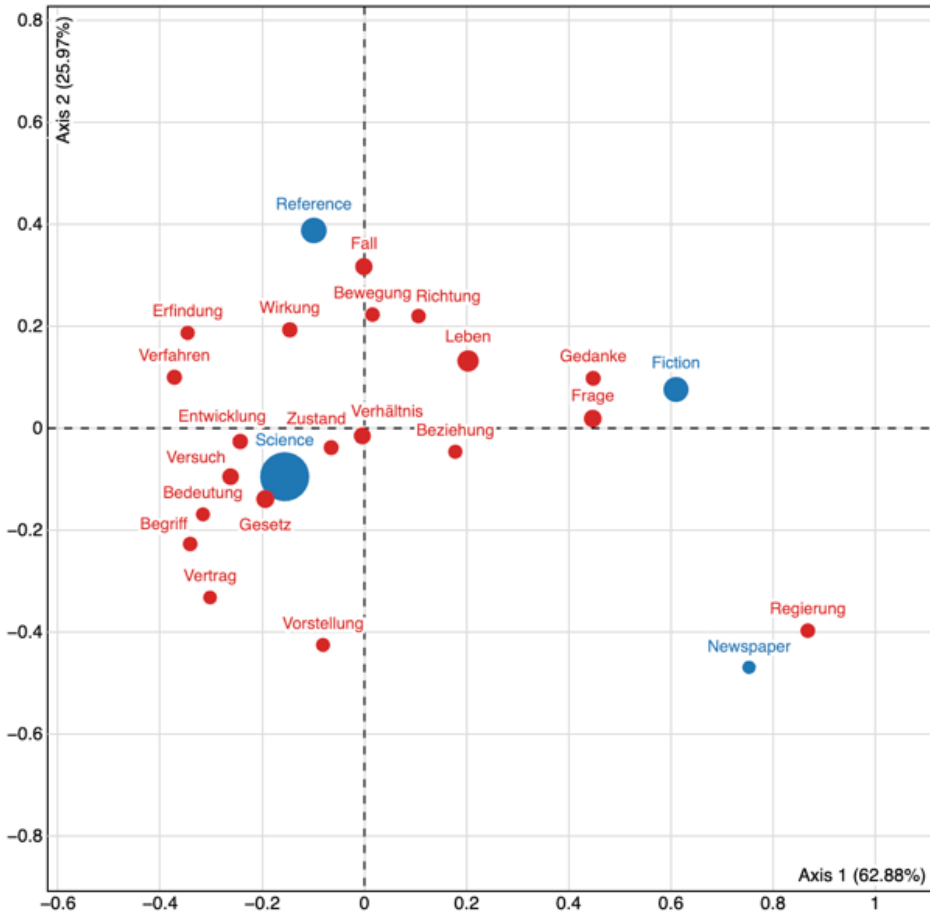


Figure 15: Most frequent nouns by GENRE in a CA plot (ZERO), see Table 13

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 the 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. 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 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 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

The research reported in this paper stems 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.
- Admoni, Vladimir G. 1990. *Historische Syntax des Deutschen*. Tübingen: Niemeyer.
- 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.
- 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.
- Meyer, David, Achim Zeileis & Kurt Hornik. 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/>.
- Ruf, Regina & Elena Smirnova. Forthcoming. Mit Hilfe von Kreativität? How German complex prepositions are shaped. In Filatkina Natalia & Sabine Arndt-Lappe (eds.), *Dynamics at the lexicon-syntax interface: Creativity and routine in word-formation and multi-word-expressions*. Berlin: De Gruyter.
- 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.
- 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

Table 9: NHG nouns by GENRE in a CA plot (CP), see Figure 8

noun	n	noun	n
<i>Verhältnis</i> ‘relation’	125	<i>Gefühl</i> ‘feeling’	6
<i>Vergleich</i> ‘comparison’	81	<i>Berücksichtigung</i> ‘consideration’	3
<i>Betreff</i> ‘regard’	49	<i>Herstellung</i> ‘production’	3
<i>Entwicklung</i> ‘development’	36	<i>Abschluss</i> ‘conclusion’	2
<i>Zusammenhang</i> ‘connection’	20	<i>Anknüpfung</i> ‘connection’	2
<i>Rücksicht</i> ‘consideration’	19	<i>Entfaltung</i> ‘unfolding’	2
<i>Hinsicht</i> ‘regard’	18	<i>Anwandlung</i> ‘mood’	2
<i>Auffassung</i> ‘conception’	8	<i>Annahme</i> ‘assumption’	2
<i>Einverständnis</i> ‘agreement’	8	<i>Anpassung</i> ‘adaption’	2
<i>Rückblick</i> ‘regard’	6	<i>Abgrenzung</i> ‘limitation’	2

Table 10: PROTOG, OHG, MHG and ENHG nouns by GENRE in a CA plot (CP), see Figure 9

noun	n	noun	n
<i>Lauf</i> ‘course’	253	<i>Übereinstimmung</i> ‘agreement’	28
<i>Bezug</i> ‘relation’	235	<i>Dienst</i> ‘service’	28
<i>Folge</i> ‘consequence’	200	<i>Bund</i> ‘tie’	27
<i>Fall</i> ‘case’	94	<i>Besitz</i> ‘property’	23
<i>Richtung</i> ‘direction’	87	<i>Berührung</i> ‘contact’	22
<i>Verbindung</i> ‘connection’	79	<i>Verkehr</i> ‘traffic’	21
<i>Anfang</i> ‘begin’	75	<i>Beginn</i> ‘begin’	19
<i>Verlauf</i> ‘course’	57	<i>Ermangelung</i> ‘lack’	15
<i>Auftrag</i> ‘order’	42	<i>Anlehnung</i> ‘backing’	14
<i>Entfernung</i> ‘distance’	30	<i>Erinnerung</i> ‘memory’	13

Table 11: ENHG nouns by GENRE in a CA plot (VNC), see Figure 11

noun	n	noun	n
<i>Betrieb</i> ‘usage’	277	<i>Verkehr</i> ‘traffic’	23
<i>Stand</i> ‘condition’	149	<i>Entfernung</i> ‘distance’	22
<i>Verbindung</i> ‘connection’	131	<i>Erinnerung</i> ‘memory’	19
<i>Anwendung</i> ‘usage’	110	<i>Vergessenheit</i> ‘oblivion’	19
<i>Besitz</i> ‘property’	87	<i>Verdacht</i> ‘suspicion’	11
<i>Beziehung</i> ‘relation’	69	<i>Zustand</i> ‘state’	10
<i>Aussicht</i> ‘prospect’	47	<i>Erwägung</i> ‘contemplation’	9
<i>Aufnahme</i> ‘uptake’	40	<i>Übereinstimmung</i> ‘agreement’	9
<i>Einklang</i> ‘harmony’	31	<i>Erregung</i> ‘excitement’	8
<i>Drehung</i> ‘turning’	25	<i>Umlauf</i> ‘circulation’	7
<i>Stimmung</i> ‘mood’	25	<i>Verwendung</i> ‘usage’	7

Table 12: PROTOG, OHG, MHG and NHG nouns by GENRE in a CA plot (VNC), see Figure 12

noun	n	noun	n
<i>Anspruch</i> ‘claim’	222	<i>Ordnung</i> ‘order’	56
<i>Betracht</i> ‘consideration’	203	<i>Handel</i> ‘trade’	39
<i>Lage</i> ‘situation’	144	<i>Begriff</i> ‘concept’	35
<i>Bewegung</i> ‘motion’	139	<i>Vorschlag</i> ‘suggestion’	27
<i>Gebrauch</i> ‘usage’	115	<i>Spiel</i> ‘game’	24
<i>Frage</i> ‘question’	98	<i>Stich</i> ‘stab’	23
<i>Dienst</i> ‘service’	76	<i>Ruhe</i> ‘calm’	23
<i>Gang</i> ‘motion’	69	<i>Verlegenheit</i> ‘embarrassment’	23
<i>Leben</i> ‘life’	67	<i>Erfahrung</i> ‘experience’	18
<i>Berührung</i> ‘touch’	65	<i>Richtung</i> ‘direction’	17

Table 13: Most frequent nouns by GENRE in a CA plot (ZERO), see Figure 15

noun	n	noun	n
<i>Leben</i> ‘life’	2146	<i>Zustand</i> ‘state’	860
<i>Gesetz</i> ‘law’	1404	<i>Regierung</i> ‘government’	808
<i>Frage</i> ‘question’	1350	<i>Begriff</i> ‘concept’	802
<i>Fall</i> ‘case’	1282	<i>Bewegung</i> ‘motion’	796
<i>Verhältnis</i> ‘relation’	1230	<i>Richtung</i> ‘direction’	788
<i>Versuch</i> ‘attempt’	1164	<i>Erfindung</i> ‘invention’	753
<i>Entwicklung</i> ‘development’	944	<i>Beziehung</i> ‘relation’	732
<i>Wirkung</i> ‘effect’	927	<i>Vorstellung</i> ‘concept’	721
<i>Verfahren</i> ‘procedure’	906	<i>Bedeutung</i> ‘meaning’	706
<i>Gedanke</i> ‘thought’	869	<i>Vertrag</i> ‘contract’	655