Determiner omission in German prepositional phrases

TIBOR KISS

Sprachwissenschaftliches Institut – Ruhr-Universität Bochum


Abstract

In this paper, we present an analysis of so-called determinerless PPs in German, i.e. prepositional phrases that allow singular count nouns to occur without an accompanying determiner, despite other rules in the grammar requiring the presence of the determiner. The analysis is based on annotated corpus data, which are fed into a statistical classifier (applying logistic regression). Superficially, the syntax of bare prepositional phrase is erratic, and intuitions cannot be easily elicited. The analysis is based on data sets for two couples of German prepositions, viz. mit ‘with’ and ohne ‘without’, and über ‘over, above’ and unter ‘under, below’. The results of the classifiers applied to annotated data indicate which syntactic, morphological and semantic features are responsible for determiner omission. We are able to detect common properties of all four prepositions, as well as preposition-specific, and idiosyncratic properties. The apparently erratic conditions for determiner omission can be discerned by tracing the interaction of these properties.
1. **Introduction: Bare Prepositional Phrases and Preposition-Noun Combinations**

There is a condition governing noun phrases that holds for many languages, but seems to be weakened (or even reversed) if the NP occurs as complement of a preposition. This condition says that a determiner is required if an NP is headed by a singular count noun. A count noun like *bus* must not appear without a determiner if it is realized as the object of a verb, cf. (1a). But it can – in fact **must** (cf. Himmelmann 1998: 316; Baldwin et al. 2006: 172) – be used without a determiner, when embedded under the preposition *by*:

(1) (a) She took *(the) bus.

(b) She came by *(the) bus.

The construction in (1b) is sometimes called **determinerless PP** or **preposition+noun combination** (PNC, cf. Stvan 2009), reflecting that the peculiarity depends on the presence of a preposition.

The ostensibly offending combination in (1b) does not necessarily establish an irregularity. It could be described by stating that the determiner **must** be dropped if the noun is the complement of *by* and its semantics corresponds to *means of transportation* (cf. Baldwin et al. 2006).

In this paper, we would like to single out a specific sub-class of determinerless PPs, which we call **Bare PPs** (BPP). BPPs show an irregular behaviour that does not lend itself easily to rules of the type suggested above to deal with (1b). BPPs are restricted by the following properties:

P I. (Countability) The complement of the preposition must be headed by a singular count noun.

P II. (Restriction to P) The omission of the determiner inside the PP must **not** lead to a semantic shift of the nominal complement, nor to ungrammaticality. Determiner omission *salva veritate* (without affecting the truth, i.e. without changing the meaning) and *salva congruitate* (without affecting the grammaticality) is thus **only** possible inside PP (cf. Himmelmann 1998: 323–324).

P III. (Optionality) The omission of the determiner must not be obligatory, it should thus always be possible to replace a BPP by a PP. Determiner omission is **optional** inside PP.

P IV. (Phrasality) The internal structure of the nominal complement should be extendable in principle. The complement of a preposition inside a BPP is a phrase.

The construction in (1b) is not a BPP: the omission of the determiner is obligatory. The properties P I to P IV are language independent and thus constitute BPPs cross-linguistically.

For the purpose of the present analysis, we will focus on BPPs in German, starting with the examples in (2).

(2) (a) Von einem Peso **mit (einem) Che-Guevara-Bildnis** über authentische Zigarren

from a peso with a Che-Guevara-effigy over authentic cigars
bis zu Medikamenten wird vieles feilgehalten.
through to pharmaceuticals PASS.AUX a.lot kept.for.sale

‘A variety of things are kept for sale, beginning with a peso with an effigy of Che Guevara, including authentic cigars, up to medical drugs.’

(b) Mit {einer besonderen}, besonderer Genhmigung können sie bei der Messe mit a special special permission can they at the mass vergorenen Traubensaft verwenden. fermented grape.juice use.INF

‘They can use fermented grape juice during the mass by special permission.’

(c) Wer mit {einer} Kreditkarte zahlt, sollte sein Konto im Auge behalten.

with a credit.card pays, should his account in.the eye keep.INF

‘Those who pay with a credit card should keep an eye on their bank account.’

The bracketing of the determiner indicates its optionality in (2a, c); similarly, (2b) illustrates the optionality of the determiner, which in this case is reflected in the declension class of the adjective. The examples thus satisfy P III (Optionality). The pertinent nouns in (2) may also be pluralized, as is indicated in (3), corroborating their status as count nouns.

(3) mehrere Bildnisse/ Genehmigungen/ Kreditkarten

several effigies/ permissions/ credit.cards

‘several effigies/permissions/credit cards’

The omission of the determiner in (2) neither leads to ungrammaticality nor to a semantic shift of the nominal complement. With regard to the latter, the situation should be compared to determiner omission outside of PP. Here, either the semantics of the nominal complement changes, so that the pertinent expressions cannot be substituted salva veritate, or the whole phrase receives a non-compositional interpretation. The first case is discussed with count/mass alternations in Huddleston & Pullum (2002: 335–337):

(4) (a) Considerable injustice was revealed during the enquiry.

(b) Two injustices were revealed during the enquiry.

The presence of a numeral – two – in (4b) leads to an interpretation that differs from the one without a determiner in (4a): in (4a), the sentence refers to the abstract concept of ‘injustice’, while in (4b), the pluralized phrase refers to EVENT INSTANTIATIONS.

The second case occurs in conventionalized constructions like the X-ist-X-und-Y-ist-Y construction in German: 4

(5) Bildnis ist Bildnis und Foto ist Foto.

effigy is effigy and picture is picture
‘Effigies should be kept apart from pictures.’

This example does not convey the tautological truth that an effigy is an effigy and a picture is a picture, but instead that effigies and pictures must be kept apart. The examples in (4) and (5) thus do not provide evidence against P II.

The examples in (6) show that the nouns presented in (2) must appear together with a determiner when marked singular and realized as an object of a verb. Thus, P II (Restriction to P) is satisfied in addition to P I, and P III.

(6) (a) Auf dieser Münze findet sich *(ein) Che-Guevara-Bildnis. 
    on this coin finds REFL a Che-Guevara-effigy
    ‘An effigy of Che Guevara is found on this coin.’

(b) Zur Verwendung von Traubensaft benötige ich *(eine) Genehmigung.
    for use of grape juice need I a permission
    ‘I need a permission to use grape juice.’

(c) Gestern beantragte ich *(eine) Kreditkarte.
    yesterday applied for I a credit card
    ‘I applied for a credit card yesterday.’

BPPs are possible if the preposition mit shows mereological (2a), conditional (2b), and modal (instrumental) interpretations (2c). The mereological interpretation covers the presence (in the case of mit) or absence (in the case of ohne (‘without’)) of an object, property or feature (cf. Schröder 1986: 162; Kiss et al. 2016: 121–122). The sense CONDITIONAL relates a prerequisite to a dependent state of affairs that is expressed in the sentence in which the PP is realized – modulo the PP (cf. also Helbig & Buscha 2007: 379; Schröder 1986: 151–152).  

5 Modal interpretations in general describe the modification of events or actions; instrumental interpretations in particular cover instruments, which are used in actions and events (cf. Schröder 1986: 146–148; Kiss et al. 2016: 279–280).

Yet, BPPs are not allowed under every circumstance. Optionality (P III) is to be interpreted unidirectional: BPPs can always be transformed into PPs by adding an appropriate determiner, and – in the case of German – changing the declension class of the noun and adjective(s) accordingly. The reverse, however, is not true: not each and every PP can be turned into a BPP, as can be illustrated with the examples in (7).

    Abuhalima has on the 6. March with an airplane of the South African Airways the USA left
    ‘Abuhalima left the USA on 03/06 with an airplane of the South African
Airways.’

(b) Das Ministerium erhielt die Kopie eines Telegramms des damaligen sowjetischen stellvertretenden Verteidigungsministers mit *(dem) Befehl, Soviet deputy defence secretary with the order, Wallenberg sei festzunehmen.

Wallenberg is SBJV to be apprehended

‘The ministry received a copy of a telegram of the former soviet deputy defence secretary, containing the order that Wallenberg should be apprehended.’

(c) Die Räuber konnten zusammen mit *(einem) dritten Komplizen fliehen.8

The robbers were able together with a third accomplice flee.

‘The robbers were able to flee together with a third accomplice.’

In (7a), the preposition assumes the sense modal (instrumental again). But here, the determiner cannot be omitted.9 The same holds for the mereological interpretation in (7b). One could argue that the major difference between the examples in (2) and (7) is that the latter examples show postnominal extensions, a postnominal genitive NP in (7a), and a finite clause in (7b). But this does not hold for (7c), and yet, the determiner cannot be omitted here. Also, it is not that simple that a postnominal extension always blocks determiner omission, as can be witnessed from (8).

(8) Es muss energisch der Meinung widersprochen werden, wonach den Unternehmen mit statutarischer Verankerung des Opting out Nachteile erwachsen werden. disadvantages arise.

‘It is necessary to strongly object against the opinion, according to which companies with a statutory anchoring of an opting out will accrue a disadvantage from it.’

The grammatical examples show adjectival modification, and postnominal extensions, thus establishing P IV. The apparently unsystematic pattern observed in (2), (7) and (8) is reflected in the grammaticality judgments of German speakers, who are often unable to judge the grammaticality of (constructed) BPPs. They are also reluctant to coin new BPPs. According to the German Duden grammar (Duden 2005), all BPPs in (2), (7) and (8) should be ungrammatical.10

The set of German simple prepositions allowing BPPs is quite large: it comprises at least 22 prepositions. While it would be worthwhile to analyse all 22 prepositions, considerations of feasibility dictate that we presently deal with a smaller subset. In this paper, we will discuss BPPs headed by
mit ‘with’, ohne ‘without’, über ‘over, above’, and unter ‘under/below’. The choice of prepositions is governed by commonalities and differences of the prepositions: The prepositions mit and ohne realise a unique cluster of senses, which contains antonymous sense pairs corresponding to Baldwin et al. (2006: 171–172) distinction of positive vs. negative senses. The preposition ohne is the only preposition in German that occurs more often in BPPs than in PPs, just as the prepositions without, zonder, and sin, in English, Dutch, and Spanish, respectively. The variation available with ohne is illustrated in (9).

(9) (a) Die Sozialdemokraten auf der anderen Seite bekundeten Mühe mit ihrer Rolle als Regierungspartei ohne (eine) Parlamentsmehrheit.

‘The Social Democrats struggled with their role as being a governing party without a parliamentary majority.’

(b) Ohne (den) Eingriff in die Entwicklung müsste 1998 mit einem Gesamtaufwand von 1362 Millionen Franken für den Betrieb der Akutspitäler gerechnet werden.

‘Without an intervention into the development, the total costs for running the emergency hospitals will amount to 1,362 million SFR.’

(c) Man kann ein so kleines Kind nicht schlagen, ohne *(den)Vorsatz, es entsprechend zu treffen.

‘One cannot beat such a small child without the intention to actually hit it.’

The prepositions unter and über have been chosen because they antonymically share a spatial sense, and also because their respective behaviour with respect to BPPs requires rather different explanations. The examples in (10) illustrate BPPs found with unter.

(10) (a) Als Folge dieser Ostdrift wurde die westeuropäische Polarluft selbst unter {einer gleichzeitigen), gleichzeitiger} Teilung ostwärts verschoben.

‘As consequence this GEN eastern.drift PASS.AUX the western.European polar.air even under a concurrent concurrent division ostwärts verschoben.’
eastward shifted

‘The Western European polar air moved eastwards, and was simultaneously divided as a consequence of this eastern wind current.’

(b) Es ist denkbar, dass man solche Veränderungen unter *(einem) Mikroskop

it is conceivable that one such transformations under a microscope wahrnehmen könnte.

perceive.INF could

‘It is possible that such transformations could be observed under a microscope.’

(c) Israelische Sicherheitsleute vermuten, dass sich ein Selbstmordattentäter

Israel security.officers assume that REFL a suicide.attacker unbemerkt unter *(eine) Ansammlung religiöser Juden

unnoticed under a crowd religious Jews mischen würde.

‘Israeli security officers assume that a suicide attacker tried to blend in with a crowd of religious Jews.’

The examples in (11) provide evidence for the existence of BPPs with über.

(11) (a) In den vergangenen Wochen haben fast alle Münchner Gegner

in the past weeks have nearly all Munich opponents über(eine) ungerechte Behandlung geklagt.

over an unfair treatment complained

‘Almost all opponents from Munich have complained about an unfair treatment during the past weeks.’

(b) Das Programm soll über(das) Kabelnetz in der

the channel shall over the cable.net in the Deutschschweiz verbreitet werden.

German.speaking.Switzerland distributed PASS.AUX

‘The channel is to be distributed through the cable net in German-Speaking Switzerland.’

(c) Die Behandlung des Bundesrats hatte die Diskussion über

the treatment of the federal.councillor had the discussion over (eine) Rationierung in der Medizin entfacht.

a rationing in the medicine sparked

‘The treatment of the federal councillor had sparked the discussion about scaling down in medicine.’

(d) Die Aufregung über *(die) Tunnelfinanzierung haben den
the commotion over the financing of the tunnel have the
aussenpolitischen Erfolg der Ministerin völlig überschattet.
foreign political success the minister completely overshadowed
‘The minister’s success in foreign policy has been completely overshadowed
by the commotion caused by the financing of the tunnel.’

To identify the syntactic and semantic factors of determiner omission, and to distinguish them from purely idiosyncratic properties related to individual combinations, we apply a methodology termed ANNOTATION MINING (Chiarcos et al. 2008; Kiss et al. (2010)), in combination with GENERALIZED LINEAR MIXED MODELLING (GLMM, Zuur et al. 2009). The data were extracted from a Swiss-German newspaper corpus (Neue Zürcher Zeitung). Annotations comprise all kinds of linguistic levels, including lexical, morphological, syntactic, and semantic information in particular. GLMM provides an extension of logistic regression that does not only allow to determine general factors that are responsible for determiner realisation and omission, but also allows us to address the question of the productivity of the construction itself, by looking at the influence of individual nouns.

The main results of the present paper can be summarised as follows: The formation of BPPs is a productive process with three of the four prepositions (ohne, mit, unter), but is restricted to specific nouns for über.

The analysis will establish general structural conditions that interact with preposition-specific conditions, and possibly also with lexical idiosyncrasy. The structural complexity of the nominal complement of the PP has been identified as a general factor inhibiting determiner omission: genitive NP complements, prepositional complements and modifiers of N, as well as relative clauses and clausal complements appearing to the right-hand side of N thus inhibit determiner omission. This condition even leaves a mark in the analysis for über, where determiner omission is generally blocked.

For the prepositions mit and ohne, a specific sense supports determiner omission, and a number of other senses inhibit it to a stronger or lesser degree. The converse situation applies to unter: a specific sense inhibits determiner omission, while other senses facilitate it. Lexical idiosyncrasy plays a role in the analysis of mit, unter, and über, but can be neglected in the analysis of ohne. In contrast to prior analyses, we have started with analysing individual prepositions. As will be discussed in Section 5, the results justify this step. In particular, it does no longer come as a surprise that speakers are reluctant to judge or coin BPPs if the behaviour of BPPs is only partially governed by general conditions, which interact with preposition-specific, and even idiosyncratic properties. The interaction would hardly be detected if the preposition-specific differences in the analyses would be neglected.
The remaining paper is structured as follows: In Section 2, we discuss previous proposals to the analysis of determinerless PPs. Section 3 is concerned with the methodology employed in the present analysis. Section 4 presents the analysis, starting with the rule-based components of the models in Section 4.1. The relevance of the random component, which consists of idiosyncrasy introduced by individual nouns, will be dealt with in Section 4.2. Section 4.3 will discuss the random component in light of previous proposals dealing with ‘N-based’ PNCs (Le Bruyn et al. 2012; Baldwin et al. 2006; Stvan 2009), and points out the differences. Section 5 summarises the results and concludes the paper.

2. THE TREATMENT OF DETERMINERLESS PPs IN PREVIOUS PROPOSALS

BPPs have not been narrowed down by conditions like P I to IV in previous research. In the following, we will discuss pertinent aspects of several recent proposals to PNCs, beginning with formal aspects in Section 2.1, and addressing questions related to the interpretation in Section 2.2.

2.1 Structural aspects of determiner omission

Any syntactic analysis of a given phenomenon must address whether the phenomenon should be treated as syntactic, or in more general terms, as rule-based in the first place. Several proposals argue explicitly against a rule-based account of PNCs, and suggest an approach in terms of fixed expressions.12

Huddleston & Pullum 2002 and Duden 2005 assume that PNCs in English and German, respectively, form a finite set of listed constructions with high cardinality. Confronted with a possibly very large set of fixed expressions, speakers will not be able to judge their acceptability. Instead, tests could only reveal memory capabilities, or even the level of the speaker’s erudition. Duden 2005 tries to accommodate for a subclass of PNCs and assumes that determiner omission becomes possible in PPs in the genre of law and naval language. The examples given above, however, do not fall under either. Huddleston & Pullum (2002: 409–410) extend this idea from FIXED EXPRESSIONS to FIXED FRAMES for English PNCs. Kiss (2007) and Dömges et al. (2007) have shown, however, that the construction is indeed productive, using a quantitative measure to gauge whether new instances of a construction emerge over time.

A crucial property of BPPs is P I (Countability), so the treatment of countability may help to understand BPPs. Analyses of countability, however, often assume that the count/mass distinction is expressed in the lexicon (be it at the level of the lemma or the individual sense). If such a lexical analysis is assumed, it remains to state that BPPs do not obey a condition otherwise obligatory for count nouns, as already provided in P I. Borer 2005, however, provides an explicit syntactic account
of the count/mass distinction. Borer 2005 assumes that nouns are interpreted as countable if they appear in the necessary syntactic context of a Quantity Phrase (QP). The plural-marking on a noun or the presence of a singular indefinite determiner provide such a context equally well. If a noun that otherwise would occur in the context of a QP shows up in a context lacking such a phrase, then the noun receives a mass interpretation (similar to grinding interpretations of count nouns in other contexts, cf. Pelletier 1975). Borer’s proposal is indeed well-suited to account for cases of count/non-count polysemy with food terms, as have also been discussed in Huddleston & Pullum (2002: 336–337). However, we would thus expect that BPPs do never occur under Borer’s analysis. They lack a plural marking and a determiner by adherence to P I, and counterfactually, they would always violate P II, because the omission of the determiner leads to a meaning shift. Instead, we would expect a completely regular phenomenon: nouns that otherwise occur as count nouns would regularly receive a mass interpretation when occurring as a determinerless complement of a preposition. But the nouns in (2), (7)–(11) do not receive a mass interpretation. An alternative would be to assume that the examples contain a covert QP. But under such an assumption, we would need some indication why a covert QP could not be realized inside NPs that are not embedded by prepositions, cf. (1a) and (6), and why not all realisations of a covert QP equally lead to grammaticality in PPs.13

The property of Phrasality (P IV) is related to the issue of productivity. For BPPs, we have claimed that they show a syntactic structure that does not differ from the syntactic structure of nominal projections in other contexts – of course modulo the presence of the determiner. There are various options to implement such a condition, e.g. by assuming that the preposition (optionally) selects a nominal projection instead of a DP. A particular implementation of this idea is provided in Baldwin et al. (2006: 175). They assume that the preposition by – if heading a PNC as in (1b) – selects for a nominal projection the head of which still has an open SPR-dependency in terms of HPSG (Pollard and Sag 1994), and hence lacks a determiner. The analysis allows a full-fledged N’ projection in all other respects, including complementation and modification of the noun. Such an analysis seems appropriate for BPPs obeying Phrasality (P IV), but it contradicts the assumption stated in Baldwin et al. (2006: 167) that constructions of type (1b) are restricted with regard to modification. In addition, such a proposal would have to meet the challenge imposed by the contrasts between (2) and (7). With respect to P IV, Himmelmann (1998: 317–319) mentions that Rumanian and Albanian restrict determiner omission to UNMODIFIED NOUNS. Yet, Himmelmann provides examples where modification and a lack of a definite determiner go hand in hand (Himmelmann 1998: 328). When Himmelmann (1998: 332) address the ‘overall complexity of the phrase’, it is not the syntactic structure of the nominal complement but the internal complexity of the preposition.
Trawiński et al. 2006 deal with the syntax and semantics of subclass of PNCs, which violates P III and P IV. The construction also differs from BPPs in that a postnominal PP complement is obligatory. Trawiński et al. 2006 assume a raising analysis in which the obligatory PP complement originates as a syntactic argument of the determinerless head noun and is raised to become a syntactic argument of the preposition. Most importantly, raising postnominal extensions would not shed any light on the syntactic distribution of BPPs, as postnominal extensions are optional in BPPs.

While Stvan’s 2009 analysis address a subset of PNCs in English that clearly differ from the ones under investigation in the present paper, her analysis of nouns is relevant. Stvan 2009 assumes that at least some PNCs are determined by a specific class of nouns. Among other characteristics, these nouns violate P II, and thus appear determinerless outside of PPs. For these nouns, modification is generally prohibited (Stvan 2009: 329–331). As will become clear in the present analysis, certain nouns have a strong facilitating effect on determiner omission in German as well, but these nouns only show the effect inside BPPs. This aspect will be further discussed in Section 4.3.

While not a formal aspect proper, we also note that the choice of prepositions under investigation is often unsystematic in previous proposals. Baldwin et al. 2006 chose the prepositions as, at, by, in, and on without providing a rationale. In Himmelmann’s seminal paper, the prepositions under investigation are the ones that allow determiner omission in the first place. This is a working proposition if the class of prepositions, or the class of prepositions taking part in PNCs, is very small. This assumption does not hold for German or English. As we already mentioned, German shows more than 20 prepositions allowing BPPs. Interestingly, Himmelmann (1998: 333) remarks: ‘In the Germanic languages, no generalisations [concerning the presence or absence of a determiner in PPs] are possible with respect to individual prepositions or subclasses of primary prepositions.’

2.2 The semantics of the preposition

The present proposal follows a research strand that was initiated by Himmelmann 1998, Baldwin et al. 2006, Kiss 2007, and Le Bruyn et al. (2012: 192–194) in assuming that the interpretation of the preposition plays an essential role in analysing determiner omission.

The relevance of the preposition senses has implications for the general approach to the analysis, and also to the analyses in itself. If the senses of the preposition are relevant for determiner omission, then it becomes necessary to identify all senses so as to characterise which sense supports or inhibits determiner omission. Without a prior definition of the senses, this cannot be achieved. Working with the senses of preposition again justifies it to start with individual prepositions, because even without a full inventory of prepositions, it should be clear that not all prepositions share all senses.
Interestingly the majority of analyses covering PNCs starts by using individual senses without a background of a sense inventory and arbitrarily divided interpretations of prepositions into (primary) senses and uses. As a consequence, the preposition form is often identified with a prominent sense of the preposition, particularly so if the preposition shows a spatial sense. The following statement from Baldwin et al. (2006: 171) is indicative: ‘a significant number of spatial prepositions (...) occur in [P+N combinations] in both temporal and stative uses’. As Baldwin et al. 2006 do not provide a distinction between senses and uses, the implications of the aforementioned statement remain dubious. The least thing that would be required for an analysis of PNCs based on a distinction of senses would be a distinction of the senses mentioned. In ignoring the polysemy of the preposition, the research question is actually evaded. Following a similar line of reasoning, Trawiński et al. (2006: 188) restrict their considerations on two senses of the German preposition in: a spatial one, and a ‘its metaphorical non-spatial meaning’. Kiss et al. (2016: 99–110) distinguish nine different super-senses of in.

Baldwin et al. (2006: 171) argue that preposition senses can be divided antonymically into positive (inclusive) and negative (exclusive) ones. They reach the conclusion that ‘positive prepositions’ occur more often in PNCs than ‘negative’ ones, referring to contrasts like in vs. out and on vs. off. This is an interesting assumption but naturally requires an extension to all prepositions showing such antonymous senses (including the prior definition of the senses). As will become clear in Section 4, the present analysis of the prepositions mit and ohne provide evidence against such a conclusion. The two prepositions share several sub-senses, and can be characterized as being antonymic with respect to these senses. If mit assumes a MERELOGICAL sense, indicating that something is part of something else, ohne assumes the same sense indicating that something is NOT part of something else. As we will see, ohne is the only preposition in German that occurs more often in BPPs than in PPs, an observation that holds for other languages as well. The realisation of BPPs with these two prepositions is in fact tied to senses of the prepositions, but antonymy plays a role as shared senses show the same tendencies.

Unfortunately, neither Baldwin et al. 2006 nor Himmelmann 1998 try to provide criteria for the proper segregation of senses. Baldwin et al. (2006: 172), e.g., assume a distinction between lexicalized and grammaticalized prepositions, where the latter, more abstract meanings are assumed to be less prone to determiner omission, without actually defining the distinction. How do we gauge senses like SPATIAL, MERELOGICAL, and MODAL-INSTRUMENTAL in terms of their abstractness/concreteness? In any case, we can show that their behaviour in BPPs sharply differs, as will be seen below. With regard to spatial interpretations, Himmelmann (1998: 321) draws a distinction between generalized locative meanings and concrete local relations. He assumes that prepositions bearing the former
meanings are susceptible to determiner omission, while prepositions bearing the latter meanings do not. Stative locative and directional interpretations are subsumed under generalized locative meanings. Proximal locatives and superessives – which is a sense expressing vertical relations as can be found with English above – are taken to be concrete local relations. Himmelmann does not provide criteria to identify the individual senses, and also does not account for his decision to assign proximals and superessives to the concrete local relations, and other stative (and directional) senses to the generalized locative meanings. While the identification of the senses could possibly be achieved in relevant contexts, the assignment of the individual senses to the super-senses remains dubious. In particular, it should be noted that proximal and superessive interpretations form a subset of the stative locatives.14

The annotations on which the analysis in Section 4 is based employ the sense inventory developed in Kiss et al. (2016). For reasons of space, the inventory itself cannot be presented here.

3. METHODOLOGY

The analysis presented in Section 4 is based on the concept of ANNOTATION MINING. Annotation mining is a two-step process. In the first step, linguistic data available in a corpus are annotated at all possible linguistic levels. This means that the annotations may apply to different entities within the corpus (the lemma, the phrase, the sentence), and that different linguistic levels are annotated (such as morphology, syntax, semantics). The annotated data are then subjected to further analysis by making use of Generalized Linear Mixed Modelling (Zuur et al. 2009).

A dependent feature is predicted by summing up coefficients for a number of independent features in a Generalized Linear Mixed Model (GLMM). Here, the dependent feature is the presence or absence of a determiner, and the model provides a probability that a determiner will be realised, given the coefficients of the features present in the sentence. Since the majority of the features in the present model are categorial (e.g. indicating whether a noun is the result of a nominalisation, or which of the senses of a preposition is given), the calculation is reduced to the coefficients15, which are summed up and fed into a link function (the inverse logit in the present case) to provide the probability of determiner realisation.

The model also uses RANDOM FEATURES, which differ from ‘ordinary’ – so-called FIXED – features, in various respects; while fixed features are finite by definition, random features may be sampled from a possibly infinite population, and hence are infinite themselves. The preposition’s senses, the syntactic complements and modifiers of the noun are typical fixed features, bearing only a finite number of values. Random features, as we employ them, are the nouns that head the nominal
complements of the preposition in a PP. There is a possibly infinite number of different nouns, and different samples may contain different subsets of this infinite set.

3.1 Annotated features

The annotated features that form the basis for the development of the GLMMs can be divided into three groups:

The first group comprises global features such as the sentence type (verb-initial clause, verb-second clause, verb-final clause) and information about the contexts in which the clause as a whole appeared, including information about PPs occurring in headlines, in quotations (particularly of poems), and in (media) titles. BPPs occurring in such contexts have generally been excluded from the analysis. Determiner omission appears to be a fairly general, yet stylistic operation in these environments.

The examples were part-of-speech-tagged with STTS (Stuttgart-Tübingen Tagset; Schiller et al. 1999) and syntactically parsed with the MALT-Parser (Nivre et al. 2007), using a Dependency Grammar (Osborne 2015) and the TiGER annotation for syntactic relations (Brants et al. 2004). With regard to the PP as a whole, the grammar provides information about its governor and its syntactic position, i.e. whether the PP occurs in the so-called German Mittelfeld between the position of the finite verb in main clauses and the position of the finite verb in subordinate clauses, or whether it occurs in sentence-initial (topicalisation) position.

The second group comprises features related to the preposition, and the PP as a whole. The senses of the preposition are annotated at the level of a super-sense (if a super-sense is relevant), and at the level of the most specific sense, according to the guidelines developed in Kiss et al. (2016).

The third group comprises features related to the noun and the nominal projection.

To satisfy P I (Countability), the candidates for (B)PPs were filtered by a classifier that determined whether the nominal head of the complement can be analysed as a count noun. This prior classification step resulted in a set of 4,413 count noun lemmata. Only PPs containing such a lemma as head of the nominal complement were considered for the analysis.

Morphological information is provided for inflection, nominalisation, further derivations, possible derivational suffixes, as well as compounding.

At the level of lexical semantics, we use GermaNet (Hamp & Feldweg 1997) to provide a rough semantic description. The nouns are mapped to their ‘unique beginners’, which form the top part of GermaNet’s taxonomy, so that each noun is described by the set of unique beginners it belongs to.

The analysis provides information about the pre- and postnominal extensions of the noun, so that we are able to capture information pertaining to P IV (Phrasality). We have aggregated prenominal
modification under the term ADJECTIVAL MODIFICATION. It reflects that adjectival modification is the primary source for prenominal extensions in German nominal projections. The relevant postnominal extensions are genitive NPs, appositions (including titles of media), prepositional modifiers and prepositional complements, relative clauses, complement clauses, and of course the case of no extension at all. The parser produces more relations than the aforementioned, which are aggregated under the feature OTHER EXTENSIONS. This aggregation reflects that the majority of these features were erroneously assigned, and also that they appear rather rarely.

The annotations provide individual analyses of instances of the data, employing established facts on their syntax and semantics (among other linguistic levels). Even before the annotated data are fed into GLMMs, their aggregation may already provide insights that may not be visible at the individual instances; in particular, interesting gaps of otherwise possible combinations may be observed (cf. Section 3.3.)

### 3.2 Generalized linear mixed modelling

A full technical introduction of Generalized Linear Mixed Models (GLMMs) lies beyond the scope of the present paper, so we will introduce the most important properties of the model only informally. Linear models usually provide a numeric prediction for a given value on the basis of the summation of numerical feature values and their coefficients. A formula for such a model is provided in (12).

\[ Y \sim \beta_0 + (\beta_1 X_1) + ... + (\beta_n X_n) + \varepsilon \]

Here, \( Y \) is the term to be predicted, \( \beta_0 \) is the intercept term, and \( \varepsilon \) is an error term, which reflects that no model can perfectly predict the data (the error term will not be considered henceforth, but it should be clear that no statistical model reaches a perfect prediction).

The INTERCEPT TERM can be conceptualized as providing basic information in the absence of the other predictors. In GLMMs, the intercept is defined to provide REFERENCE information, which we will discuss below. \( X_1 \) to \( X_n \) are the features, and \( \beta_1 \) to \( \beta_n \) are the corresponding coefficients. If a coefficient is 0, then the whole feature becomes irrelevant, and smaller coefficients express smaller influences than larger coefficients. At first sight, the application of such a model to linguistic questions seems improbable since linguistic features are usually categorial, and multiplying a number with a categorial feature does not make sense. In addition, whatever could be summarised from such multiplications would at best be a number in the interval between \( \pm \infty \), the interpretation of which remained mysterious. Fortunately, a linear model as the one in (12) can be transformed into a meaningful model in a few steps.
The first step consists in re-coding categorial variables through contrast coding (Chambers & Hastie 1992), so that the presence or absence of a feature is coded by 1s and 0s. Features receive the value 1 if present, and 0 if absent. As a consequence, the relevant predictors in the model are restricted to the coefficients itself (since $\beta \times 1 = \beta$, and $\beta \times 0 = 0$). All coefficients can either be positive or negative, indicating which of the outcomes they favour. For the purposes of the present analysis, a negative coefficient indicates that the feature value is in favour of determiner omission, while a positive coefficient indicates that it is in favour of determiner realisation.

The second step (which is actually interdependent with the first step) is to integrate reference values into the intercept. We know that we have more than one sense per preposition, and we also know that we can choose from different values for the postnominal extension of the noun in a (B)PP, so one sense, and one postnominal extension will be taken to be the reference, and the coefficients for the other features will be determined in relation to this reference. Let us briefly illustrate this with the preposition *mit*. We assume for PP headed by *mit* (in fact for all PPs included in the present analysis) that NO EXTENSION will be the reference value for postnominal extensions. With regard to the senses of *mit*, we assume that the MERELOGICAL sense – presented in (2a) and (7b) – will be the reference (this sense is genuine to the prepositions *mit* and ohne). This means that examples containing the preposition *mit* will provide an intercept, which contains this information, and further coefficients must be added if the sense of the preposition differs from the MERELOGICAL sense, or if there is a postnominal extension present.

The third step provides the actual transformation from a numeric predictor to a probability of determiner realisation or omission. To achieve this goal, the outcome of the prediction is fed into a link function: the INVERSE LOGIT maps predictors from $-\infty$ to $+\infty$ to the interval between 0 and 1, and thus provides the desired probability.¹⁹

Let us consider two schematic examples to illustrate this. In the first example, the preposition assumes the MERELOGICAL sense, and no postnominal extension is present. In the second example, we assume that the meaning of the preposition is CAUSAL, and that a postnominal genitive NP is present. Both examples are schematically depicted in (13).

(13) (a) [mit\textsubscript{mero}\ldots N]  
(b) [mit\textsubscript{caus} \ldots [NNP\text{gen}]]

We assume a model (actually, we are taking the values from the model discussed in Section 4) in which the intercept receives a value of −0.84 (that is our $\beta_0$). In line with what we said before, the negative value indicates that the model suggests determiner omission with a certain likelihood on the basis of the intercept. For (13a), with the MERELOGICAL sense, and no postnominal extension, we can finish the calculations here, and determine the probability of a determiner realisation, by feeding
the intercept term – which is the only term – into the inverse logit. This will yield a probability of 30.2 % for determiner realisation. In (13b), we have the intercept term, and in addition, two further features that have to be added to it to account for the differences between the reference values and the actual values of the features. The coefficient for a postnominal genitive NP is 2.30, the coefficient for the sense CAUSAL is 3.15. Both coefficients are positive, hence suggest determiner realisation. They have to be added to the intercept together with the value of the contrast coding, resulting in the formula in (14).

(14) \[ Y \sim -0.84 + 2.30 \times \text{contrast(POSTNOMINAL GENITIVE NP)} + 3.15 \times \text{contrast(CAUSAL)} \]

The contrast coding yields a 1 for both features in the case of (13b). The value for \( Y \) will be hence be 4.61 (= –0.84+2.3+3.15), which – when fed into the inverse logit, yields a probability of 99.02 % in favour of determiner realisation.

The models presented in Section 4 make use of the same set of fixed features, so that fixed features showing relevance for the analysis across models suggest themselves as general properties of BPPs, while the presence or absence of features in individual models indicates a preposition-specific condition.

A model with contrast coding for categorial values, reference values contained in the intercept, and a probability for a possible outcome out of two is called a Generalized Linear Model. It becomes a Generalized Linear Mixed Model by adding a random component. In the present model, we want to gauge the influence of the individual (4,413) noun types on determiner realisation, so the noun lemmata are taken as random features in the model. They differ from the features introduced so far in that they are not drawn from a finite vocabulary. As we will gauge their influence on the intercept only, the model is called a RANDOM INTERCEPT MODEL, and the formula in (12) is extended by adding this influence (\( \theta_0 \)) to the intercept term:

(15) \[ Y \sim (\beta_0 + \theta_0) + (\beta_1 X_1) + ... + (\beta_n X_n) + \varepsilon \]

The formula in (15) provides an average value for \( \theta_0 \), which is interesting insofar, as the fixed features \( X_1 \) to \( X_n \) can be interpreted as the rule-based component of this model, while \( \theta_0 \) is a reflection of idiosyncrasy. If the average value for \( \theta_0 \) would be large, then we would have to assume that idiosyncratic values overwhelm the rule-based component of the model – which would lead to the conclusion that the whole process modelled is more idiosyncratic than rule-based. In the analysis in Section 4, it will turn out that for three of the four prepositions, \( \theta_0 \) is rather small. The exception is the model for \( \text{über} \), which suggests that BPPs with \( \text{über} \) result from the lexical influence of individual nouns. On the flipside, we will see that the influence of the random component can be neglected for the preposition \( \text{ohne} \), the behaviour of which appears to be almost completely rule-based. For the two
prepositions *mit* and *unter*, there are individual nouns that may exert an influence over the rule-based component. These nouns can be identified by a high individual $\theta_0$, so that the model allows a mixed description, taking rule-based and individual factors into account.

As will become clear in the analyses, not all of the features are actually employed. This is so since the models do not only determine a coefficient for the features, but also determine whether the coefficient’s value is actually the result of a mere accident. If this were the case, we could not be certain whether the actual value of the coefficient would be 0, so the model provides a likelihood whether the coefficient could be 0. If this likelihood is too large (usually, one relies on a boundary of 5 %), the feature showing the coefficient will be excluded from the analysis for the simple reason that even if it were present, it would be multiplied with its coefficient – the value of which would be 0 – and hence would not play a role. There are some interesting exclusions on this basis, which will be discussed below.

3.3 On introspection and (un-)acceptability

The employment of Annotation Mining and Generalized Linear Mixed Modelling should not be mistaken as a principled statement against introspective judgments on acceptability. The methodology used here should be seen as an addition, and not as a replacement. In fact, introspective judgments play two important roles in the present analysis.

First, annotation mining is based on annotations at all kinds of linguistic levels, and for many of these levels, the annotations have been developed on the basis of introspective judgments. As with many other annotations syntactic annotations derive at least partially from reference corpus data, and hence on the introspective judgments of the annotators. Although these judgments may have been filtered by the sheer quantity of data in the corpus, they form a basis nevertheless. So, the present analysis relies at least partially on now covert introspective judgments. But the present analysis also relies on overt introspective judgments. Some of these judgments are completely uncontroversial, as e.g. the optionality of the determiner in (2). Of course, the optionality can be backed up by the corpus data itself, since the pertinent nouns appear with and without a determiner if embedded under *mit*. The situation becomes controversial when we turn to data that are claimed to be unacceptable.

It is common sense in linguistics that unacceptable data cannot be found in a corpus. Yet, some qualifications are necessary here. Linguistic phenomena follow a so-called Large Number of Rare Events (LNRE) distribution (Baayen 2009); even if the linguistic ‘events’ were finite in nature, this means that the vast majority of them only occurs rarely. Of course, we assume infinity for linguistic events, but we are also aware that infinity should be established among the linguistic tokens, and not among their descriptions, in terms of grammars or other devices. Compositionality in particular
implies that the number of linguistic rules be finite, and it would be counter-intuitive at least to assume that linguistic rules follow an LNRE distribution, too. What does this mean for the question at hand? Our claim is that patterns of unacceptability can actually be derived from a corpus, given that the corpus consists of annotated data (cf. Section 3.1).

As an illustration of an interesting gap, consider the distribution of complement clauses and senses for the prepositions mit and unter with and without a determiner.20 If a determiner is present, complement clauses may show up as postnominal extensions with various senses of the two prepositions. But in the absence of a determiner, we do not find a single example showing a complement clause.21 A cross-classification of senses and postnominal extensions provides a schematic syntactic and semantic analysis for PPs – an analysis which gives rise to the suspicion that determiners are obligatory, if postnominal complement clauses are present. Given the size of the samples, it becomes conspicuous that postnominal complement clauses do not occur if a determiner is missing. This observation gives rise to an elicitation of introspective judgments, which reveals, in accord with the above observations, that speakers strongly disfavour the absence of a determiner in such patterns. We thus feel justified to assume that the lack of a determiner actually leads to unacceptability in such cases. Similar considerations apply to other instances of unacceptability reported in the present paper.

3.4 Summary statistics

The analysis is based on data sets for the prepositions mit, ohne, über, and unter as summarised in Table 1.

<table>
<thead>
<tr>
<th>Prep.</th>
<th>Number of observations</th>
<th>Frequency PP</th>
<th>Frequency BPP</th>
<th>Number of noun types</th>
<th>Meanings of P</th>
</tr>
</thead>
<tbody>
<tr>
<td>mit</td>
<td>15,281</td>
<td>11,360</td>
<td>3,921</td>
<td>1,683</td>
<td>14</td>
</tr>
<tr>
<td>ohne</td>
<td>2,965</td>
<td>425</td>
<td>2,540</td>
<td>696</td>
<td>8</td>
</tr>
<tr>
<td>über</td>
<td>2,123</td>
<td>1,905</td>
<td>218</td>
<td>690</td>
<td>9</td>
</tr>
<tr>
<td>unter</td>
<td>6,067</td>
<td>5,140</td>
<td>927</td>
<td>670</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 1: Summary statistics of the data sets

A few comments are required here. First note that the proportion of BPPs for ohne is much higher than the respective proportion of PPs. This preposition differs from all other prepositions in this respect. Apparently, determiner omission has become regular with this preposition – a conclusion that
does not only hold for *ohne* in German, but also for *zonder* in Dutch (Le Bruyn et al. 2012), and for *without* in English. This distribution also has consequences for the GLMM for *ohne*, where we are looking for factors that inhibit determiner omission.

Secondly, the data set for *mit* appears much larger than the other data sets. The data sets for *mit*, *ohne*, and *unter* are the complete data sets found in the sample from the *Neue Zürcher Zeitung* corpus based on the 4,413 count noun lemmata, after the filtering procedures (elimination of headlines, titles etc.) were applied. The larger total number of observations for *mit* thus simply reflects that relevant examples with *mit* occur more often and with a larger subset of the 4,413 noun lemmata than relevant examples for *ohne* and *unter*. In the case of *über*, the data set only shows the comparatively small number of 218 BPPs listed in Table 1. We have thus decided to sample PPs from the much larger data set for *über*, and not to look at the complete data set. The rationale behind this decision is that we are looking for features facilitating determiner omission. Looking at the complete data set would thus not have added further data showing determiner omission, but would have reduced the proportion of the BPPs.

4. FACTORS OF DETERMINER OMISSION

The analysis of determiner realisation in PPs headed by *mit*, *ohne*, *unter*, and *über* is the subject of this section. In the first sub-section, we will discuss what we call the rule-based component of the analysis, which is based on the linguistic features included in the analysis as fixed factors. Subgroups of the features re-occur with the different prepositions, and common properties can be attributed across the prepositions, even if an analysis cannot be based on one and the same feature for the prepositions. As the four prepositions do not share senses, we do not expect that one and the same sense can be made responsible for determiner omission across the prepositions. What we can show, however, is, that there are specific senses that either inhibit or facilitate determiner omission.

The second section will deal with the random features, which introduce (currently) irreducible lexical idiosyncrasy into the analysis: the lexical heads of the nominal complements of the prepositions. Baldwin et al. (2006: 166, 169f.) and Le Bruyn et al. (2012: 188) have emphasised that the behaviour of a subset of PNCs is largely determined by the noun contained in the construction, so-called ‘N-based’ PNCs. A discussion about the particular influence of individual nouns on determiner realisation in German BPPs thus will have to take into account whether the findings to be reported here allow a mere transfer of the earlier assumptions on N-based PNCs to BPPs, or whether they suggest that a different treatment is required. This issue will be addressed in Section 4.3.

The analysis has been carried out in R with the libraries LME4 (Bates et al. 2015), BLME (Chung et al. 2013), and LSMEANS (Lenth 2016). All graphics have been produced with the library GGPLOT2.
(Wickham 2009). The models are Bayesian Generalized Linear Mixed Models (implemented using BLME), which are more robust than Generalized Linear Mixed Models when QUASI SEPARATION becomes an issue. Quasi separation occurs if a predictor (almost) perfectly splits the data. In the models reported here, the conspicuous gap of complement clauses not occurring in BPPs constitutes a case of quasi separation (cf. Section 3.3).

In order to keep the presentation of the models in Section 4.1 transparent, we will present the models without the random components in this section.

4.1 The categorial component of the models

4.1.1 The model for mit

The model for mit given in (16) contains the largest set of features, which can be subdivided into POSTNOMINAL EXTENSIONS, the SENSES OF THE PREPOSITION, and three individual features.

The first group provides information about the influence of syntactic extensions of the head noun on determiner omission. Recall from Section 3.2 that a positive coefficient indicates that the feature facilitates determiner realisation, while a negative value indicates that the feature facilitates determiner omission. All postnominal extensions bear a positive coefficient and hence suggest that postnominal extensions inhibit determiner omission.

The senses of the preposition provide information about the role of the different senses of the preposition for determiner omission. The annotated senses follow the analysis in Kiss et al. (2016). All senses listed for mit show positive coefficients, suggesting that the senses listed inhibit determiner omission (but see immediately below).

A closer scrutiny of the aforementioned groups reveals that features are missing in both groups: there is no coefficient for the feature NO POSTNOMINAL EXTENSION, and there is no coefficient for the MERELOGICAL sense of mit, which has already been introduced in (2) and (7), nor for two other senses of the preposition, viz. DEPENDENCY and ASSOCIATION. As for the latter two, the model will only contain SIGNIFICANT coefficients. The role and representation of significant features will be discussed below.

The features NO POSTNOMINAL EXTENSION and MERELOGICAL sense (as well as others, to be discussed below), however, are not listed because they are contained in the INTERCEPT TERM. The intercept term provides information about determiner realisation and omission, when the so-called reference values of the features can be applied. If a feature is binary, the reference value will be the absence of the feature (which will pertain to the two other features to be discussed: ADJECTIVAL MODIFICATION, and NOMINALISATION), if a feature contains more than two values – as with
postnominal extensions and senses of the preposition, one feature is chosen as a reference value. In the models for *mit* and *ohne*, the reference value for the preposition’s sense is the **MEREOLOGICAL** sense (as is the **SPATIAL** super-sense in the models for *unter* and *über*), and in all models, the reference value for postnominal extension is that there is no such extension. We note that the intercept term is negative, thus indicating that determiner omission is likely in the presence of the reference values.

The feature **ADJECTIVAL MODIFICATION** represents the presence or absence of a pre-nominal modifier. Its reference value is that no such modifier is present. The feature **NOMINALISATION** describes whether the noun contained in the (B)PP is the result of a derivation from a verbal base. Again, the reference value assumes that the noun has not been derived. In sum, the model predicts that determiner omission is more likely than determiner realisation if the sense of the preposition is **MEREOLOGICAL**, if neither a prenominal nor a postnominal extension is present, and if the noun is not the result of a derivation.

(16) Fixed Component of Generalized Linear Mixed Model for *mit*:

**Number of observations:** 15,281; **determiner:** yes 11,360, no 3,921

| Feature                  | Coefficient | Pr(>|z|)     |
|--------------------------|-------------|-------------|
| (Intercept)              | -0.83544    | 3.79e-11 ***|
| Adjectival Modification  | -0.98361    | < 2e-16 ***  |
| **Postnominal extensions** |             |             |
| Postnominal Genitive     | 2.31048     | < 2e-16 ***  |
| Apposition               | 3.05075     | 5.42e-13 *** |
| PP complement            | 2.51778     | < 2e-16 ***  |
| PP modifier              | 2.43593     | < 2e-16 ***  |
| Clausal complement       | 6.54965     | 1.26e-07 *** |
| Relative clause          | 2.67701     | 1.51e-10 *** |
| Other extensions         | 1.30259     | < 2e-16 ***  |
| **Senses of the preposition** |       |             |
| causal                   | 3.20760     | < 2e-16 ***  |
| conditional              | 2.20951     | 1.95e-11 *** |
| event                    | 5.30834     | < 2e-16 ***  |
| indicator                | 4.11596     | 5.15e-09 *** |
| modal                    | 2.60083     | < 2e-16 ***  |
| participation            | 4.53443     | < 2e-16 ***  |
| point-of-reference       | 4.51736     | < 2e-16 ***  |
| realisation              | 5.70331     | < 2e-16 ***  |
| restrictive              | 4.24290     | 1.72e-08 *** |
| temporal                 | 3.15569     | 3.55e-08 *** |
| governed                 | 4.89785     | < 2e-16 ***  |
| **Derivational morphology of the noun** |       |             |
| Nominalisation           | -0.55040    | 0.00596 ***  |
Fixed features are selected only if they contribute to the analysis. The model has in fact access to further morphological information, such as compounding, but such features do not play a role for the analysis, because they are not significant.

Whether or not a feature plays a role is determined by calculating how sure we can be that the coefficient of a given feature value is not actually 0, i.e. that the value given in the model is not only different from 0 by chance. We accept coefficients only that show a probability of being 0 of less than 5 %, which is determined through the Wald test statistic (abbreviated by $z$). In the following, we have omitted the actual Wald test statistic $z$ and only represent the relevant likelihood of the coefficient being 0 in the column headed by $Pr(>|z|)$ – as in the model in (16). The stars provide indication for the strength of the likelihood: one star indicates a likelihood between 5 % and 1 %; two stars indicate a value between 1 % and 0.1 %; three stars indicate a value of less than 0.1 %, possibly bordering to 0. The value $<2e$-16 indicates that the actual value is smaller than the smallest value which can be calculated by the system used. As we can see, all remaining coefficients receive three stars.

With regard to the senses of mit, we note that 12 of the 14 senses are actually significant, while two senses are not significantly different from the reference value, so that changing from the MEROLOGICAL sense to the senses DEPENDENCY or ASSOCIATION (cf. Kiss et al. 2016: 111) will not change the prediction. The fact that ADJECTIVAL MODIFICATION supports determiner omission (given its negative coefficient) is surprising in light of the claim found in Baldwin et al. (2006: 167) that ‘modification is seldom unrestricted [in P+N combinations]’.

The value of the intercept is approximately -0.84. If the intercept term would be the only term indicating whether or not a determiner should be realized in the case of mit, it would suggest determiner realisation with a likelihood of only 30 %, which is the probability provided by the inverse logit of -0.84. We can thus conclude that the MEROLOGICAL sense and the absence of pre- and post-nominal extensions support determiner omission. Let us now take a look how the model deals with the examples, which predictions are made, and how they come about, starting with examples (2a, b), repeated here under (17).23

(17) (a) Von einem Peso mit (einem) Che-Guevara-Bildnis über authentische Zigarren bis zu Medikamenten wird vieles feilgehalten.

A variety of things are kept for sale, beginning with a peso with an effigy of Che Guevara, including authentic cigars, up to medical drugs.

(b) Mit {(einer besonderen), besonderer} Genehmigung können sie bei der Messe...
with special permission can they at the mass verwenden.
fermented grape juice use.

‘They can use fermented grape juice during the mass by special permission.’

Example (17a) is predicted to allow determiner omission by the model as it is completely covered by the intercept: the interpretation of the preposition is MERELOGICAL, there is no pre- or postnominal extension in the nominal projection, and the noun Bildnis (‘effigy’) is not a derived nominal. We can thus calculate the likelihood of determiner realisation by feeding the intercept term (-0.84) into the inverse logit, yielding a likelihood of 30.24% for determiner realisation.

Next consider (17b). Here, the sense of the preposition is CONDITIONAL, which inhibits determiner omission with a positive coefficient of 2.21. But (17b) does not only differ from (17a) with regard to the sense of the preposition: in addition, the noun is modified by an AP, and it can be classified as a derived nominal, two features which support determiner omission due to their negative coefficients. Taking all factors together, the likelihood for determiner realisation reaches a value of 46.01%. (We will return to (17b) in Section 4.2.)

Let us compare these results with the results for the examples provided in (7), repeated here under (18).

(18) (a) Abuhalima hat am 6. März mit *(einem) Flugzeug der
Abuhalima has on.the 6. March with an airplane of.the
South African Airways die USA verlassen.
South African Airways the USA left
‘Abuhalima left the USA on 03/06 with an airplane of the South African Airways.’

(b) Das Ministerium erhielt die Kopie eines Telegramms des damaligen
the ministry received the copy of.a telegram of.the former
sowjetischen stellvertretenden Verteidigungsministers mit *(dem) Befehl,
Soviet deputy defence secretory with the order,
Wallenberg sei festzunehmen.
Wallenberg is.SBJV to.be.apprehended
‘The ministry received a copy of a telegram of the former soviet deputy
defence secretary, containing the order that Wallenberg should be
apprehended.’
Die Räuber konnten zusammen mit *(einem) dritten Komplizen fliehen

The robbers were able together with a third accomplice flee.

‘The robbers were able to flee together with a third accomplice.’

The preposition shows the sense MODAL (INSTRUMENTAL) in (18a), and a postnominal genitive NP. Hence, we find two features with positive coefficients that accordingly inhibit determiner omission. The GLMM for mit predicts a likelihood for determiner realisation of 98.33 %, matching the ungrammaticality of determiner omission in the example.

While the sense MODAL facilitates determiner realisation, it does so to a much lesser degree than the sense PARTICIPATION (as will become clear shortly, the former makes determiner realisation about ten times more likely, but the latter about a hundred times). Consequently, the model predicts that the determiner has to be realized in (18c) with a likelihood of 95.58 %. In (18b), we find the MERELOGICAL sense again, but the postnominal extension is a complement clause. The presence of a complement clause is an even stronger indicator for determiner realisation, and the example is predicted to contain a determiner with a likelihood of 99.67 %.

The influence of the individual fixed features on the outcome can be made more transparent in a plot of the odds ratio (Agresti 2007: 28–33). The odds ratio indicates how many times more likely or less likely the positive outcome is, given the presence of the feature. In the present case, a positive outcome is the realisation, and a negative one the omission of the determiner.

The odds ratios for the significant coefficients of the fixed factors in the model for mit are provided together with their 95 % confidence intervals in Figure 1. The dotted line is provided for the value 1. This value would indicate that the feature makes the outcome neither less nor more likely, i.e. does not have an influence at all. If a feature shows a value below the dotted line, it makes determiner realisation less likely (it shows a negative coefficient), if it occurs above the dotted line, it makes it more likely (it shows a positive coefficient). The odds ratios are given on a logarithmic scale, since the dotted line separates the effects symmetrically, leading to equidistance for the values below and above 1. Hence a feature value with an odds ratio of 0.1 is exactly as far away from 1 as is a feature value with an odds ratio of 10; thus allowing the former to outweigh the latter.
The majority of the postnominal extensions makes determiner realisation more than ten times more likely. If no postnominal extension is present, which is indicated in the \textit{INTERCEPT}, then determiner realisation becomes less likely, as we already have illustrated. In the following, we will only present the odds ratios, and discuss the implications on their basis.

One could wonder whether the odds ratio in Figure 1 does reveal even more structure regarding the individual features. To this end, we have applied Tukey’s Honest Significance Distance Test (Tukey’s HSD, Baayen 2009: 114–116) to determine whether the features differ from each other.

With regard to the postnominal extensions, it turns out that the features \textsc{Postnominal Genitive}, \textsc{PP Complement}, and \textsc{PP Modifier} differ significantly from \textsc{Complement Clause}, but not from each other. These differences can be interpreted so that the former already make determiner realisation more likely, but the presence of the latter makes determiner omission impossible. And this is what we have seen in (18b) – cf. also the discussion in Section 3.3.
With regard to the senses of the preposition, we can identify two groups of features: the senses CAUSAL, CONDITIONAL, and MODAL make determiner omission more likely, but they sharply differ from the senses PARTICIPATION, POINT OF REFERENCE, REALISATION, EVENT, and GOVERNED, the positive odds ratio of which is much larger, resulting in obligatory determiner realisation even if there are no postnominal extensions, as in (18c).

With the exception of (2c), which will be dealt with in Section 4.2, the apparently erratic distribution of determiner realisation and omission exemplified in (2) and (7) could be accounted for by identifying the features provided in (16) and Figure 1: One sense of the preposition, the MEREOLOGICAL sense, facilitates determiner omission, while the other senses inhibit it. POSTNOMINAL EXTENSIONS also inhibit determiner omission, while ADJECTIVAL MODIFICATION and NOMINALISATION support it. Since these features can be present at the same time, it might be possible that a postnominal extension is present, but a determiner can still be omitted, particularly so, if the noun is modified by an AP or is the result of nominalisation. This is the situation found in (8), where we find a postnominal genitive NP suggesting determiner realisation together with three properties that suggest determiner omission: the MEREOLOGICAL sense, the presence of an AP, the derived nature of the head noun.
4.1.2 The model for ohne

The odds ratios for the model of *ohne* are presented in Figure 2.

![Figure 2: Odds ratios and confidence intervals for significant coefficients in the GLMM for *ohne*](image)

The model for *ohne* contains two features that have not reached significance in the model for *mit*. They relate to the position of the PP in the German clause (cf. Müller 2015), and to the element with which the PP is combined. Basically, the model says that PPs that modify a verb (*GOVERNOR: V*), and PPs occurring in sentence initial position in German root clauses (*PHRASE: V2*) are (a little) less likely to omit a determiner – both values are positive. The two features *GOVERNOR: V* and *PHRASE: V2* are specific for *ohne* – they are the only features that do not show up in the other models, and will be discussed in Section 5 in this respect.

Let us look at the predictions of the model with respect to the data presented in (9), repeated here under (19).
(19) (a) Die Sozialdemokraten auf der anderen Seite bekundeten Mühe mit ihrer Rolle als Regierungspartei ohne (eine) Parlamentsmehrheit. ‘The Social Democrats struggled with their role as being a governing party without having a majority in parliament.’

(b) Ohne (den) Eingriff in die Entwicklung müsste 1998 mit einem Gesamtaufwand von 1362 Millionen Franken für den Betrieb der Akutspitäler gerechnet werden. ‘Without an intervention into the development, the total costs for running the emergency hospitals will amount to 1,362 million SFR.’

(c) Man kann ein so kleines Kind nicht schlagen, ohne *(den)Vorsatz, es entsprechend zu treffen. ‘One cannot beat such a small child without the intention to actually hit it.’

The model for **ohne** introduces a high negative intercept (-3.74858), which – when fed into the inverse logit – yields a probability of determiner realisation of only 2.3 %, which is what is predicted for (19a). In addition to assuming that no postnominal extension is present, that the sense of the preposition is MERELOGICAL, that there is no ADJECTIVAL MODIFICATION, and that the noun is not the result of the derived nominal, the INTERCEPT also assumes that the PP is NOT governed by a verb, and that it does NOT occur in the verb-initial position in a German main clause – all conditions are met in (19a). Given the high negative intercept, the model predicts thus the aforementioned low likelihood of determiner realisation, which is in line with the observation that determiner realisation is almost completely optional with **ohne**. The odds ratio shows that all other significant features except for NOMINALISATION increase the likelihood of determiner realisation. As with **mit**, the postnominal extension COMPLEMENT CLAUSE stands out as a strong inhibitor of determiner omission. We note a much lower number of significant senses, which do not establish a hierarchy (at least not in the sense of Tukey’s HSD test).

Next, consider (19b). This example shows the CONDITIONAL sense of the preposition, as well as a prepositional complement of the noun. Also, the PP appears in verb-initial position, and is
dependent on a verb. Taken together, the likelihood of determiner realisation is still only 41 %, and again, the determiner can easily be omitted. This is different in (19c), where the following features come together: first, the PP modifies a verb, which is subsumed under the feature GOVERNOR: V. This feature only has a weak influence (cf. Figure 2), but the sense CONDITIONAL, and the postnominal COMPLEMENT CLAUSE strongly facilitate determiner realisation. The features thus lift the likelihood of determiner realisation to 95.8 %.

The prepositions mit and ohne share five senses, for which an antonymous relation can be established: MERELOGY, CONDITIONAL, MODAL, RESTRICTIVE, and PARTICIPATION. For each sense, mit provides an inclusive interpretation (as e.g. X is part of Y, X is carried out by use of Y, etc.), while ohne provides the exclusive or privative interpretation (e.g. X is NOT part of Y, X is NOT carried out by use of Y etc.). The results of the analysis contradict Baldwin et al.’s 2006 suggestion that inclusive senses allow determiner omission more easily than exclusive senses. Determiner omission occurs more often – and according to the strong negative intercept: more easily – with ohne than with mit, despite the exclusive senses of the preposition.

4.1.3 The model for unter

The odds ratios of the model for unter are provided in Figure 3. There is a high positive odds ratio for the INTERCEPT, indicating that determiner realisation is much more likely than determiner omission. As in the models for mit and ohne, the significant postnominal extensions (PP COMPLEMENT and MODIFIER, POSTNOMINAL GENITIVE NP, RELATIVE CLAUSE, APPOSITION and COMPLEMENT CLAUSE) further inhibit determiner omission.

The intercept contains the SPATIAL sense of the preposition, hence spatial interpretations of the preposition facilitate determiner realisation. The three senses STATE, MODAL, and CONDITIONAL have an inhibiting effect on determiner realisation. Interestingly, the largest inhibiting effect is provided by the feature Nominalisation. Its negative coefficient is almost identical to the intercept (INTERCEPT: 4.8599; NOMINALISATION: -4.5757). We will discuss the effect of nominalisation in more detail below.

In addition to the SPATIAL sense of the preposition, the intercept contains reference values that are already familiar from the former models: it assumes that there is no POSTNOMINAL EXTENSION, and that the noun is not the result of NOMINALISATION.
With regard to the features postnominal extension and senses of the preposition the Tukey HSD test does not reveal anything that is not already apparent from Figure 3: All other postnominal extensions differ significantly in their influence on determiner realisation from complement clauses, which make determiner omission more or less impossible. This is quite similar to the situation observed with *mit*.

The consequences of the model can be illustrated with the examples in (10), repeated here under (20).

(20) (a) Als Folge dieser Ostdrift wurde die westeuropäische Polarluft selbst unter *(einer gleichzeitigen), gleichzeitiger* Teilung ostwärts verschoben.

as consequence this.GEN eastern.drift PASS.AUX the western.European polar.air even under a concurrent concurrent division eastward shifted
‘The Western European polar air moved eastwards, and was simultaneously divided as a consequence of this eastern wind current.’

(b) Es ist denkbar, dass man solche Veränderungen unter *(einem)* Mikroskop wahrnehmen könnte.

‘It is conceivable that one such transformations under a microscope could be observed under a microscope.’

(c) Israelische Sicherheitsleute vermuten, dass sich ein Selbstmordattentäter unbemerkt unter *(eine)* Ansammlung religiöser Juden mischen würde.

‘Israeli security officers assume that a suicide attacker tried to blend in with a crowd of religious Jews.’

The examples (20a, b) exhibit MODAL senses: CONCOMITANT CIRCUMSTANCE in (20a) and INSTRUMENTAL in (20b). Yet, determiner omission is possible in the former, but not in the latter.

Example (20a) contains an additional inhibiting feature: the head noun is a derived nominal. The cumulative effect of the two features with negative coefficients leads to a negative prediction of approximately -3.79, the inverse logit of which yields a probability of 2.21% for determiner realisation. The inhibiting features are missing in (20b), which accordingly shows a probability of 68.66% for determiner realisation. (We will return to example (20b) in Section 4.2. In addition to the fixed component, the random component also speaks in favour of determiner realisation in this case.) Example (20c) shows a non-modified derived nominal as the head noun, but the sense of the preposition is SPATIAL, and there is a POSTNOMINAL GENITIVE NP present. The probability for determiner realisation of 94.23% is not surprising.

4.1.4 The model for über

The odds ratio for über, presented in Figure 4, differs strongly from the other three odds ratios: there are only significant fixed factors with positive coefficients in the model.
Figure 4: Odds ratios and confidence intervals for significant coefficients in the GLMM for über

The fixed component predicts determiner omission to be impossible with über, an assumption that has been contradicted by the examples provided in (11). Contrary to the observations, all examples in (11) should require the obligatory realisation of the determiner. The examples, however, can be accounted for, if we consider the random component of the model.

4.2 The random component of the models

Before we will discuss how the interaction between the random and the fixed component of the models is achieved formally, we will elucidate the linguistic interpretation of the fixed and random components. As an illustration, consider the model for mit. The features given in Figure 1 can be interpreted as establishing syntactic and semantic patterns that either allow or prohibit determiner omission. An illustration for the first type is given in (21a), an illustration for the second one in (21b).
The position of a possible determiner is indicated through DET in (21). Since determiner omission is optional, a determiner could always be inserted in this position. But (21a) is a pattern in which no determiner has to be inserted, while (21b) requires its presence. The effect of the random component can be characterized as follows: in a pattern like (21b), where determiner omission should not happen, determiner omission still seems to happen with a high likelihood, depending on how the N⁰ position is filled. The random component does not provide a reason for this change, but properties that have already been captured in the fixed component cannot be made responsible for it. Consequently, the random component either captures latent properties of certain nouns that have not been established yet, or it models true cases of idiosyncrasy. In the current analysis, we cannot go deeper than this regarding the properties of the nouns that make up the random component. Interestingly, we find that determiner realisation can also become more likely in patterns like (21a), a point to which we will return in Section 4.3.

Let us now turn to the formal combination of the random and the fixed component. In a random intercept model, the effect of the random component is provided as a standard deviation of the intercept value. A standard deviation is on the same scale as the term of which it is a standard deviation. As an illustration, consider Figure 5, which shows the distribution of the logistic function.

Figure 5: Illustration of influence of the random component on the linear predictor
Let us assume that the fixed component has provided a linear prediction of 1 on the x-axis. This value is mapped to a probability of determiner realisation – $P(\text{Determiner})$ – of 73% by the logistic function, as indicated by the data point (1, 0.73) in Figure 5. If only the fixed factors were considered, the model would thus predict determiner realisation with high probability. A random intercept model gauges the influence of the random features – the nouns in our case – on the intercept. Let us assume a particular noun that bears an individual value of -3. This value must be added to – in this case: subtracted from – the intercept term, according to the formula (15), provided in Section 3, and repeated here under (22).

$$Y \sim (\beta_0 + \theta_0) + (\beta_1 X_1) + \ldots + (\beta_n X_n) + \varepsilon$$

In (22), $\beta_0$ is the intercept term, and $\theta_0$ is the added random component. Consequently, $\theta_0$ – bearing a negative value – can be subtracted from the linear predictor, yielding -2 instead of 1. Of course, the change in the linear predictor (on the x-axis in Figure 5) is reflected in a change in the mapping of the logistic function, from a 73% likelihood without the random component to a mere 12% likelihood after the random component, i.e. the individual influence of the noun present in the PP has been added, as indicated in the data point (-2, 0.12). As a result, the fixed component predicts a determiner realisation, while the combination of fixed and random components makes determiner realisation rather unlikely.

In the summary of the models, the random component is given as the standard deviation of the intercept. This means that we get a value that is averaged over all possible instances of the random component. We can, however, extract the individual random components for the nouns that occur with the respective prepositions in the corpus. In the following analysis of the random components we will provide individual values for nouns, which show a stable influence in one direction. This means that the noun either has a positive or negative influence on determiner realisation. The individual variance introduced by the nouns (provided as a standard deviation) will be listed together with their 95% confidence intervals, to make clear that the influence is actually one-sided. We will start our discussion with the random component for the prepositions mit and unter.

4.2.1 The random components in the models for mit and unter

The random components in the models for mit and unter are given in (23).

<table>
<thead>
<tr>
<th>Preposition</th>
<th>Variance</th>
<th>Standard Deviation</th>
<th>Intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td>mit</td>
<td>4.988</td>
<td>2.233</td>
<td>-0.8350</td>
</tr>
<tr>
<td>unter</td>
<td>9.687</td>
<td>3.112</td>
<td>4.8599</td>
</tr>
</tbody>
</table>
The summary of the two random components indicates that in both models, the random component has an influence on determiner omission. Consider the preposition *unter*. If we had to deal with an example that would show all features referenced in the intercept, and no further, then we would yield a likelihood of determiner realisation of over 99 %. If we further assumed that the variance introduced by the random component could be deducted from the linear prediction of the fixed component (which requires transforming it into the standard deviation to get a number on the same scale), then the likelihood would drop to 85 %. Of course, this does not look like a large deduction, but we should keep in mind that we have used averaged values here. If we look into the individual values, a different picture emerges, as can be illustrated with the example in (24).

(24) Das Wertpapier wird von der Bankengruppe *unter Lizenz von* the security PASS.AUX by the bank.consortium under licence of

**FTSE International** angeboten.

FTSE International offered

‘The security if offered by the bank consortium under licence from FTSE International.’

The preposition in (24) shows an extended sub-sense of the spatial sense, **RELATIONSHIP OF POWER**. Also, the noun *Lizenz* (*license*) shows a postnominal extension (a PP COMPLEMENT). Taken together, these features predict a likelihood of determiner realisation of 99.89 %. The determiner, however, is conspicuously missing in (24). This can be accounted for if the standard deviation introduced by the noun *Lizenz* (*license*) is added to the linear predictor. Since this value is negative (-7.26), it is actually subtracted, and the prediction for determiner realisation is lowered to 40.81 %.

A summary of the nouns that exert a random influence on the linear predictor for *unter* is provided in Figure 6.
In Figure 6, we find a large variety of nouns with an individual negative influence between -1.54 and -8.28, but we also find nouns with a positive random influence, with values between 0.83 and 4.83. Such nouns are responsible for determiner realisation in patterns that would otherwise suggest determiner omission, so that the general optionality of determiner realisation in PPs is barred if such a noun occurs as the nominal head in a BPP. One such noun is actually *Mikroskop* (‘microscope’) that we already have seen in example (10b) and repeated as (20b).

(20) (b) Es ist denkbar, dass man solche Veränderungen unter *(einem) Mikroskop* wahrnehmen könnte.

It is conceivable that one such transformations under a microscope could be observed under a microscope.
The preposition shows a MODAL (INSTRUMENTAL) sense in (20b), which facilitates determiner omission. But since no other features support determiner omission in the example, the prediction is still in favour of determiner realisation (68.66 %). And the determiner cannot be omitted in this example. If the positive random influence of the noun is taken into account, the likelihood for determiner realisation is raised to 99.00 %. We will return to the nouns that exert positive random influence in Section 4.3.

A summary of the nouns that exert a random influence on the linear predictor for mit is provided in Figure 7.

![Figure 7: Individual nouns with their influence on the linear predictor of the fixed component for mit](image)

We have not discussed (2c) in Section 4.1, which is repeated below. The sense of the preposition is MODAL (INSTRUMENTAL), and hence the model predicts a likelihood of determiner realisation of
85.38 %. The noun *Kreditkarte* (‘credit card’), however, exerts a strong negative influence on determiner realisation and reduces the likelihood of determiner realisation to 24.09 %.

(2) (c) Wer mit *(einer) Kreditkarte* zahlte, sollte sein Konto im Auge behalten.

‘Those who pay with a credit card should keep an eye on their bank account.’

In addition to (2c), let us have an additional look at (2b), repeated as (17b).

(17) (b) *Mit *(einer besonderen, besonderer) Genehmigung* können sie bei der Messe

‘They can use fermented grape juice during the mass by special permission.’

Here, the fixed component yields a likelihood of determiner realisation of 46.01 %, which is very close to 50 %, i.e. no decision at all. But the random component applies to this example as well. The presence of the noun *Genehmigung* (‘permission’) lowers the likelihood of determiner realisation to a mere 1.1 % (cf. Figure 7). Similar considerations apply to example (8), which is repeated here under (25).

(25) *Es muss energisch der Meinung widersprochen werden,*

‘It is necessary to strongly object against the opinion, according to which companies with a statutory anchoring of an opting out will accrue a disadvantage from it.’

The fixed component yielded a prediction of 48.52 %. The noun *Verankerung* (‘anchoring’) shows an additional negative influence, which turns the prediction of determiner realisation in this case to 15.84 %.

As was already the case with the random component for *unter*, not only do we find nouns that exert a negative influence, but also nouns that exert a positive one, a point to which we will return in Section 4.3.
4.2.2 The random components in the models for ohne and über

The random components for the prepositions *ohne* and *über* form two extremes of random features. The random component for *ohne* is irrelevant for the prediction. But the fixed component for *über* prohibits determiner omission altogether. As a consequence, determiner omission happening with *über* can only be accounted for by the random component of the model. A summary of the two random components is provided in (26).

<table>
<thead>
<tr>
<th>Preposition</th>
<th>Variance</th>
<th>Standard Deviation</th>
<th>Intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td>ohne</td>
<td>1.101</td>
<td>1.049</td>
<td>-3.749</td>
</tr>
<tr>
<td>über</td>
<td>20.650</td>
<td>4.545</td>
<td>4.545</td>
</tr>
</tbody>
</table>

The random component for *über* in (26) tells us that considerable variance in the model is accounted for by the nouns that are realized within the PPs headed by *über*, which in effect means that it is largely dependent on the nouns whether or not a determiner is realized. We have extracted the nouns that show a stable one-sided influence, and listed them with their individual values and their 95 % confidence intervals in Figure 8.
Figure 8: Individual random effects (i.e. nouns) in the model for über

In contrast to the random components discussed so far, the random component for über does not show nouns with positive influence. The effect of the random component can be illustrated with the examples in (11), repeated here under (27).

(27) (a) In den vergangenen Wochen haben fast alle Münchner Gegner über(eine) ungerechte Behandlung geklagt.

In the past weeks have nearly all Munich opponents over an unfair treatment complained

‘Almost all opponents from Munich have complained about an unfair treatment during the past weeks.’

(b) Das Programm soll über(das) Kabelnetz in der Deutschschweiz verbreitet werden.

The channel shall over the cable.net in the
German-speaking.Switzerland distributed PASS.AUX
‘The channel is to be distributed through the cable net in German-Speaking Switzerland.’

c) Die Behandlung des Bundesrats hatte die Diskussion über
the treatment of the federal councillor had the discussion over
(eine) Rationierung in der Medizin entfacht.
a rationing in the medicine sparked
‘The treatment of the federal councillor had sparked the discussion about scaling down in medicine.’

d) Die Aufregung über *(die) Tunnelfinanzierung haben den aussenpolitischen
the commotion over the financing of the tunnel have the foreign political
Erfolg der Ministerin völlig überschattet.
success the minister completely overshadowed
‘The minister’s success in foreign policy has been completely overshadowed by the commotion caused by the financing of the tunnel.’

The BPP shows a CAUSAL sense in (27a). The likelihood for determiner realisation is given as 98.94 % by the fixed component (it is irrelevant in this case that the noun is a derived nominal since this feature does not play a role in the model for über). The noun Behandlung, however, provides a strong negative value to be added to the intercept (-5.86). Thus, the likelihood of determiner realisation is reduced to 9.76 % by adding the random component, i.e. the influence of Behandlung on determiner realisation in PPs headed by über. Similar considerations apply to (27b). In (27b), the noun Kabelnetz has the effect of reducing the likelihood from 99.40 % to 31.70 %. In (27c), the noun Rationierung reduces the likelihood of determiner realisation from over 99 % to a mere 18.73 %. Finally, the pertinent noun in (27d) – Tunnelfinanzierung – does not belong to the random component of über, and hence, the omission of the determiner is not licensed.

In contrast to the random component for the other prepositions, the random component for ohne can be neglected. First, it contains only nine nouns with a one-sided influence, seven of which facilitate determiner realisation. Secondly, the individual influence of the nouns ranges from 1.79 to -1.42, which must be set in relation to the high negative intercept of the fixed component (-3.749). If we take the intercept as a reference point, the presence of one of the positively influential nouns would only yield a likelihood of 11 % for determiner realisation.
4.3 Does the random component provide evidence for N-based PNCs?

The inclusion of a random component in the models for BPPs has emphasized the role of the noun within a BPP. This, however, does not seem to be a new finding. It is widely acknowledged that the behaviour of a subset of PNCs is determined by the noun contained in the construction. This goes so far that Le Bruyn et al. (2012: 188) call a subclass of PNCs ‘N-based’. A similar class is singled out in Baldwin et al. (2006: 166, 169f.) and Stvan (2009: 329–331). A discussion of the particular influence of individual nouns on determiner realisation in German BPPs thus has to make clear whether this is just a corroboration of the existence of ‘N-based’ BPPs, or whether the elements of the random component require a different treatment.

Le Bruyn et al. (2012: 188) point out the following properties of N-based Dutch PNCs: First, the nouns belong to a semantic class of COMMON LOCATIVE NOUNS. By the same line of reasoning, Baldwin et al. (2006: 169–170) introduce a class of INSTITUTIONALIZED NOUNS. While a similar class of nouns can be identified in German, they show a syntactic behaviour differing from PNCs. These nouns occur in constructions where the preposition and the determiner seem to be amalgamated (so-called ‘Verschmelzungsformen’), as illustrated by (28a).

(28) (a) Er ist im Gefängnis.
   he is in the jail
   ‘He is serving a sentence.’

(b) Er ist in einem Gefängnis.
   he is in a jail
   ‘He is located in a prison.’

(c) *Er ist in Gefängnis.
   he is in jail

In (28a), we find many of the criteria developed by Stvan (2009), in particular, we find a pragmatic enrichment, which is not present in a construction, where the preposition and the determiner occur independently. While (28a) can only mean that ‘he is serving a sentence’, (28b) shows an ordinary spatial interpretation. Crucially, BPPs are ungrammatical in such contexts, as is illustrated in (28c). German also defies the claim found in Baldwin et al. (2006: 164) that ‘articles are regularly omitted in expressions of similar semantic type across languages’. They use institutions, seasons, and metaphors as examples (it should be noted here that institutions etc. refer to the meaning of the embedded nouns, not to the meaning of the phrase as a whole). It does not hold for similar constructions in German that the article can be omitted (English examples taken from Baldwin et al. 2006): at school vs. in *(der) Schule, in winter vs. im Winter vs. in *(einem/dem) Winter, at large vs. auf *(der) Flucht.
Secondly, the nouns occur across prepositions. Le Bruyn et al. 2012 illustrate this with *uit bed* (‘out of bed’), *naar bed* (‘to bed’), and *in bed* (‘in bed’). Thirdly, the nouns cannot be exchanged with near-synonymous nouns or nouns belonging to the same lexical field in the construction, as is illustrated with the contrasts between *zolder* (‘attic’) and *kelder* (‘basement’), and *school* (‘school’) and *universiteit* (‘university’) in the following examples.

(29) (a) *naar* zolder vs. *naar* *(de)* kelder
    to attic to the cellar

(b) *op* school vs. *op* *(de)* universiteit
    at school at the university

In addition, Le Bruyn et al. 2012 as well as Baldwin et al. (2006: 169f.) assume that N-based combinations require a stereotypical pragmatic enrichment, or a semantics similar to a weak definite (Aguilar-Guevara & Zwarts 2011). Furthermore, Baldwin et al. 2006 and Stvan (2009: 317–318) show that nouns in English N-based PNCs may occur without a determiner outside of PNCs, so that they systematically violate P II (Restriction to P).

The properties listed above do not apply to the nouns identified in the random components of the models for the German prepositions under discussion.

First, we note that the nouns that have been listed in Figures 6, 7, and 8 hardly belong to a common semantic class. Secondly, we do not find pragmatic enrichments or a necessary interpretation as a weak definite if the nouns occur within a BPP.

Thirdly, the behaviour across prepositions is much more complex than the pattern provided by Le Bruyn et al. 2012. This is particularly evident with one factor that has not been considered at all in the discussion by Baldwin et al. 2006, Le Bruyn et al. 2012 and Stvan 2009: the models for *unter* and *mit* (as well as the one for *ohne*) do not only contain nouns that facilitate determiner omission, but also nouns that inhibit determiner omission. The least we would expect would be that nouns occurring across prepositions show the same effects (i.e. either facilitate or inhibit determiner omission). But the very same noun may support determiner omission with one preposition, and inhibit it with another, as can be witnessed with *Vorbehalt* (‘prerequisite’) when combined with *mit* and *unter*: With the first preposition, the noun has a strong negative influence (−3.636), and hence facilitates determiner omission, with the second preposition, it shows a weak positive influence (0.836), and hence facilitates determiner realisation.

Fourthly, we note that the contrast described by Le Bruyn et al. 2012, and illustrated in (29), does not pertain to similar cases in German. Consider the sub-sense **BOUNDARY** of the spatial senses of the prepositions *über* and *unter*. This sense describes that something has transgressed an upper or lower bound, and is found in phrases like *über/unter Plan* (‘above/below what was planned’), and *über/unter Budget* (‘above/below the scheduled budget’).
The nouns belong to the same lexico-semantic field and both nouns occur with \( \text{über} \) and \( \text{unter} \), exerting negative influence on determiner realisation. Further semantically similar nouns – such as \textit{Tarif} (‘pay scale’) – could be substituted, counter the idea that the noun has to remain constant. Also, there is no stereotypical enrichment: examples containing the phrases receive a completely compositional interpretation, indicating that the external argument of the PP has either transgressed the upper or lower bound of what is the internal argument of the preposition.

Finally, we should repeat that the nouns found in the random components do not occur without a determiner outside of PPs.

In sum, the nouns contained in the random components of the models cannot be equated with the nouns identified in the proposals of Baldwin et al. 2006, Stvan 2009, and Le Bruyn et al. 2012, since they do not share their characteristics.

5. **Conclusions and Prospects**

We have set out observing that the behaviour of BPPs in German is apparently erratic, and that grammaticality judgments are not easily elicited. Using Annotation Mining and Generalized Linear Mixed Modelling, we have been able to discern three different traits of the phenomenon: there are general grammatical rules at work, but they interact with lexeme-specific rules, as well as with idiosyncratic behaviour governed by individual nouns. Given this state of affairs, the apparently erratic behaviour of BPPs in German does not come as a surprise any longer.

Across all models, we have identified \textsc{postnominal extensions} as a factor that inhibits determiner omission. This conclusion is even corroborated by the model for the preposition \textit{über}, although the fixed component of this preposition does not contain any feature supporting determiner omission.

For two prepositions, we were able to establish a hierarchy between complement clauses on the one hand, and all other extensions on the other. For these prepositions, the presence of a complement clause (including non-finite complements) makes determiner omission impossible, while the other extensions may still coincide with determiner omission, if other factors facilitating determiner omission are present as well.

\textsc{Nominalisation} appears to be a second general factor supporting determiner omission across prepositions. The model for \textit{über} did not contain this factor, as it did not reach significance, but even in this model, the calculated value for \textsc{nominalisation} was negative.

At the lexeme-specific level, we have been able to identify senses of individual prepositions that support determiner omission. This holds in particular for the \textsc{mereological} sense of the prepositions \textit{mit} and \textit{ohne}. As an interesting observation, we can note that for \textit{mit}, the \textsc{mereological} sense is actually the only sense of the preposition that supports determiner omission (for \textit{ohne}, we also find
support with the MODAL sense, the coefficient of which is not sufficiently different from the reference value, and hence the sense is not contained in the model).

For the preposition *mit*, which shows the largest inventory of senses, we have also been able to establish a hierarchy of the senses. While all other senses apart from the MERELOGICAL sense support determiner realisation, their individual strength differs. It is thus much more likely to omit a determiner with the senses CONDITIONAL, MODAL, and CAUSAL, provided that other features supporting determiner omission are present, while the senses POINT-OF-REFERENCE, PARTICIPATION, GOVERNED, EVENT, and REALISATION much more strongly inhibit determiner omission, up to impossibility.

In addition to POSTNOMINAL EXTENSIONS, NOMINALIZATION, and the senses of the preposition, we have identified ADJECTIVAL MODIFICATION as relevant for determiner omission (in the case of *mit*). Presently, it must remain an open question why ADJECTIVAL MODIFICATION differs in its effect in the case of *ohne*, where is actually supports determiner realisation. Across the prepositions, we have only identified two features that are confined to a single preposition (*ohne*): GOVERNOR: V, and PHRASE: V2. These features might be an artefact of an interaction of the senses of the preposition with possible positions of the PP, which is not directly reflected in the model: If the PP is governed by a verb or realized in clause initial position of main clauses, the preposition may not show a sense that requires modification of a noun. This is obvious in the first case, i.e. if the PP modifies a verb; as for the second feature, this seems to be reflection of the fact that extraction from NP is much more restricted than extraction from inside VP. Consequently, PPs in clause-initial position are taken to modify verbs, not nouns. But the sense that mainly facilitates determiner omission with *ohne* is MERELOGICAL, which may only emerge if the PP modifies a noun.

The interaction of the general rules (realise a determiner if the nominal complement shows postnominal extensions, omit a determiner if the noun is a derived nominal) with the individual senses, as well as with the other lexeme-specific factors leads to the ostensibly erratic pattern observed in examples like (2), (7), (8), (9), (10), and (11).

This situation is aggravated by the idiosyncratic influence exerted by individual nouns. The idiosyncrasy introduced by an individual noun can conspire with a more general pattern, as was observed in (8), where the postnominal extension suggests determiner realisation, while the MERELOGICAL sense as well as the adjectival modification suggest determiner omission. Together with the noun *Verankerung* (‘anchoring’), the dilemma is resolved, and the determiner omitted. A similar influence is found in the model for *unter*, as was further illustrated with (24). A particularly interesting case is example (10b), because the nominal head inside the PP actually exerts a positive influence, i.e. makes determiner omission even more unlikely.
The analysis of *ohne* can be fully based on the fixed component, i.e. on the interaction of post-nominal extensions, syntactic dependencies, senses of the preposition, and adjectival modification. The possibility of determiner omission with *über* is fully based on the random component, and hence it is not the structure nor the interpretation of the PP that allows determiner omission in the examples in (11), but the sheer presence of the respective nouns.

As this point, we have to add a few remarks to features that are NOT relevant in the analysis. Two such features are of particular importance: compounding and the semantics of nouns.

Baldwin et al. (2006: 168) make the interesting observation that some PNCs in English require MORPHOLOGICAL modification, i.e. they are only possible if the noun contained in the P+N combination is a compound, as illustrated by *at *(eye) level or *at *(company) expense. The examples raise a problematic aspect, which carries over to the present analysis, and will be discussed briefly below: the role of LOCALITY in a formal analysis of BPPs. What we note now, however, is that compounding has been annotated in the present analysis. Depending on the data sets, between 10 and 23 % of the nouns were compounds. But the category COMPOUND does not play a role in the analysis of BPPs – the feature never reached significance. From the perspective of language theory assuming lexical integrity, this is what one would expect if morphology and syntax are taken to be two different realms. Of course, the data presented by Baldwin et al. 2006 point into a different direction. But further investigations would require a broader research on the construction.

A second feature that does not play a role in the present analysis is the semantics of the noun, which means that the respective features never reached significance, and hence their inclusion did not improve the models. In the annotation stage, we have used GermaNet (Hamp & Feldweg 1997) to provide a rough annotation of noun senses. Now it may very well be that the annotations provided in this way were to sketchy to actually derive semantic classes for the noun. What is conspicuous in this respect, however, is that semantic classes also could not be detected from the random component. The random components would at least have hinted at semantic classes, i.e. we would have found nouns that fall into the same semantic field. Such a conclusion could not be derived from the random component, so that we have to assume that the sense of the preposition is a major factor for determiner omission, but the interpretation of the noun is not.

Let us finally return to a possible implementation of the results reached in the present paper in a theoretical framework, and its implications. As we have shown, there are at least syntactic and semantic factors to be considered in the analysis, so that a theoretical framework which covers both, such as HPSG (Pollard & Sag 1994) appears to be suitable. In such an analysis, we would code the senses of the preposition disjunctively and would require different subcategorization frames depending on the senses of the preposition. But it would also be necessary to deal with the apparent non-
locality of the postnominal extensions. In a framework like HPSG, where local schemata are employed for syntactic combination, information about the realisation of postnominal extensions are not usually projected up in a structure. While a technical implementation is feasible, it would obscure the theoretical point why such non-local information sharing is necessary in the analysis. Given the observations by Baldwin et al. 2006 on the boundary between morphology and syntax, this seems to be an issue, which has to be resolved in future research.
REFERENCES


*Author’s address:* Sprachwissenschaftliches Institut, Ruhr-Universität Bochum, 44780 Bochum, Germany
tibor@linguistics.rub.de
The research reported herein would not have been possible without the members and affiliates of the ‘PNC Project’, to which I am grateful: Daniel Abbassi, Katharina Börner, Monika Duzy, Ron Hoffmann, Halima Husic, Katja Keßelmeier, Antje Müller, Johanna Poppek, Claudia Roch, Nino Simunic, Tobias Stadtfeld, Jan Strunk, and Vanessa Weidmann. Parts of this paper have been presented at Brandeis University, Simon Fraser University, University of Alberta (Edmonton), Friedrich-Alexander-Universität Erlangen-Nürnberg, Universität Leipzig, Stanford University, and the Norwegian University of Science and Technology Trondheim. I would like to thank the audiences for their comments. In addition, I would like to thank three anonymous reviewers for their comments and suggestions. I am indebted to Katharina Börner and Anneli von Könemann for their assistance with the manuscript. Finally, I would like to thank the Deutsche Forschungsgemeinschaft (DFG) for their support under grant KI-759/5.

I would like to thank an anonymous reviewer to raise the issue that P II has to consider omission salva veritate in addition to omission salva congruitate. One could argue that P II shares some similarity with stricter versions of the CANDIDATE SET, as defined in Optimality Theory (Legrendre et al. 1998: 258).

In the following examples, indefinite and definite determiners are chosen so that the examples including a determiner appear as neutral as possible. In addition to the definite determiner’s contingency on aspects such as uniqueness and familiarity, definiteness is not necessarily marked by the determiner alone, but may emerge from e.g. embedding the NP under a definite NP, from uniqueness presuppositions, and from embedded structures, among others. In such cases, the insertion of an indefinite determiner would be infelicitous. Consequently, I have included an indefinite determiner if no further marking indicates definiteness (examples (2), (7a, c), (9a), (10), (11a, c)), and a definite determiner if the NP as a whole showed definiteness markings or uniqueness presuppositions (examples (7b), (9b, c), (11b, d)).

We would like to thank an anonymous reviewer for pointing out the significance of the data in (5).

The conditional interpretation of ohne negates the precondition.

As in the case of the conditional interpretation, the preposition ohne in its instrumental sense provides information about instruments NOT used.

It is an interesting question (raised by Martin Haspelmath and an anonymous reviewer) how the interpretation of a determiner can be recovered if the determiner is omitted. It should be noted here that the interpretation of BPPs is not necessarily indefinite, thus indicating that the absence
of a determiner allows both indefinite and definite interpretations. This conclusion can be reached by looking at (near) minimal pairs containing a definite determiner and no determiner, as e.g. carried out in Kiss (2007). An illustration is provided in (i) and (ii), where the NP in (ii) receives a definite interpretation.

(i) Er bedrohte sein Opfer mit der Pistole.
   ‘He threatened his victim with a gun.’

(ii) Er bedrohte den 44-jährigen Angestellten mit gezogener Pistole.
   ‘He threatened the 44-year-old employee with the gun he had pulled.’

The omission of the determiner would require a change in the declension class of the adjective, which then would be realized as drittem. The ungrammaticality after omission of the determiner is retained.

8 We will discuss the identification of ungrammatical examples in Section 3.3.

9 The Duden tries to accommodate for a subclass of BPPs and assumes – in rule 395 – that determiner omission becomes possible in PPs in the genre of law and naval language. The examples given above, however, do not fall under either. A similar treatment for English P+N combinations is found in Huddleston & Pullum (2002: 409–410). They extend the treatment from FIXED EXPRESSIONS to FIXED FRAMES.

10 Bresnan et al. 2007 employed logistic regression to determine the distribution of NP vs. PP arguments of ditransitive verbs.

11 This tradition is implicitly addressed in Baldwin et al. 2006, where PNCs are characterized as ‘multi-word expressions’.

12 An anonymous reviewer has pointed out that the factors discussed below could equally well license a determiner that is not pronounced. Involving a determiner that is not pronounced instead of assuming that the determiner is simply not present, is, at the present state of the art, mainly a question of scientific economy. If such a determiner without pronunciation is required to analyze different areas of grammar, one could possibly use it here as well. The crucial point is that the conditions presented below will remain the same.

13 Kiss et al. (2016: 224–227) discuss stative proximal locative interpretations of three German prepositions. They induce evidence that determiner omission is disfavoured with these senses.

14 This is so because the coefficients are multiplied with the value of the feature. If the feature is categorial, then its value is 1 if it is present, and 0 if it is absent. Consequently, we are left with the value of the coefficient \( n \) in case the feature is present \( (1 \times n = n) \), and with \( 0 \) \( (0 \times n = 0) \), if the feature is absent.
Advertisements are another area in which truncations are used, but the corpus did not contain advertisements. We would like to thank an anonymous reviewer for pointing out the role of advertisements.

For a survey of German clause structure, cf. Müller 2015.

The classifier distinguishes between count nouns and non-count nouns and thus leaves it open whether one identifies classes that are count to a certain degree (as suggested in Allan 1980). The nouns that have been identified can be assumed to satisfy the criteria imposed on strict count nouns in Allan 1980.

The inverse logit is defined as $e^Y/(1+e^Y)$, where $e$ is Euler’s number and $Y$ is the result of the linear predictor.

It should be noted here that ohne is the only preposition that actually shows complement clauses in BPPs. This observation is in line with the analysis of Section 4, which assumes that determiner omission has become almost perfectly regular with ohne. The preposition über also shows the distributional gap reported here. As the discussion in Sections 4.1 and 4.2 will reveal, however, determiner omission is dependent on individual lexical items for this preposition.

We have applied a Cramér-von-Mises test, which shows that the cross classification of senses and postnominal extensions differs in its distribution, depending on whether a determiner is present or not.

We have also developed a GLMM for mit containing a sample of the original data with 6,000 instances. The model based on the smaller sample provides results that are very similar in quantity (relative influence) and quality (for or against determiner omission) for all fixed features, and also for the random features that are shared between the two models. We would like to thank an anonymous reviewer for raising this issue.

Example (2c) is falsely predicted to prohibit determiner omission (due to the modal (instrumental) sense of the preposition), while the determiner can in fact be dropped in this example. We will return to this case when we discuss the influence of the random features in Section 4.2.

If this were not the case, we would expect that significant fixed features become insignificant after inclusion of the random component, cf. Baayen (2009: 305).