

Genitive vs. *von*-PP arguments in German NPs: Grammatical and use conditions

Antonio Machicao y Priemer 

Humboldt-Universität zu Berlin

Pauline Reiß 

Friedrich-Schiller-Universität Jena

Giuseppe Varaschin 

Humboldt-Universität zu Berlin

Proceedings of the 32nd International Conference on
Head-Driven Phrase Structure Grammar

Centro de Linguística da Universidade de Lisboa, Lisbon, Portugal

Stefan Müller, Shûichi Yatabe (Editors)

2025

Berlin: Berlin Universities Publishing

pages 69–92



Keywords: NP, genitive, von PP, functional preposition, argument structure, alternation, parasitic head, social meaning, register, use-conditions

Machicao y Priemer, Antonio, Pauline Reiß & Giuseppe Varaschin. 2025. Genitive vs. *von*-PP arguments in German NPs: Grammatical and use conditions. In Stefan Müller & Shûichi Yatabe (eds.), *Proceedings of the 32nd International Conference on Head-Driven Phrase Structure Grammar*, Centro de Linguística da Universidade de Lisboa, Lisbon, Portugal, 69–92. Berlin: Berlin Universities Publishing. DOI: 10.83067/hpsg.2025.4.

 CC-BY 4.0

Abstract

In German, NP arguments of nominal heads can be marked either with genitive case or with the preposition *von*. Although these types of marking are truth-conditionally equivalent, they have distinct contextual distributions, with genitive-marked arguments being associated with speakers who present themselves socially as more educated than *von*-marked arguments. This paper has a threefold objective. First, we offer a lexicalist account of the *von* vs. genitive alternation for NP arguments in German, deriving the pool of possible variants. Second, we formulate an extension of the HPSG architecture that allows us to predict the differences in usage preferences in a way that captures a graded notion of register and contextual appropriateness – in the spirit of Varaschin, Machicao y Priemer & Lu (2024). Our extension crucially involves the statement of use-conditional constraints, which associate independently licensed linguistic objects to gradable social meanings. Finally, we propose a corpus methodology to empirically validate hypotheses about social meaning assignments via use-conditional constraints and the principle of social meaning composition, focusing on the interaction between the *von* vs. genitive alternation and *-ung* (‘-ment’) nominalisations.

1 Introduction

In German, NP arguments of nominal heads can be marked either with genitive case (GEN) or with a PP headed by the preposition *von* (*von*-PP), which assigns dative case to the embedded NP (Machicao y Priemer 2017, Machicao y Priemer & Müller 2021, Kopf & Bildhauer 2024), cf. (1).

- (1) die Behandlung { des Patienten / von dem Patienten }
the treatment the.GEN patient.GEN of the.DAT patient.DAT
‘the treatment of the patient’

While the GEN NP and the *von*-PP are truth-conditionally equivalent, they differ w.r.t. the contexts in which they usually appear. For instance, in more educated or more formal contexts, the GEN realisation is preferred over the realisation with a *von*-PP. In contrast, in neutral or colloquial contexts the *von*-PP is preferred. Similarly, the contexts in which *ung*-nominalisations¹ occur tend to be associated with a higher degree of education and formality (cf. nominal vs. verbal style). In this paper, we focus on verbs nominalised with the *-ung* suffix, as these are cases in which the argument structure of the verbal base is inherited by the derived nominal stem. Moreover, using *-ung* nouns as a proxy for education/formality enables us to examine which form (i.e. GEN vs. *von*) shows a stronger correlation with *-ung* in the data.

¹The nominalisation suffix *-ung* in German is comparable with the suffix *-ment* in English. For further details on this kind of nominalisation, cf. Ehrich & Rapp (2000), Dölling (2015), Machicao y Priemer (2017).

The goals of this paper can be summarised as follows: (i) We propose an account of the grammatical conditions for the *von* vs. GEN alternation, which also captures the argument-marking behaviour of the preposition, which we define as a *parasitic head* (Levine 2010). (ii) We show how the HPSG architecture can be used to model not only the grammatical properties of the construction, but its use conditions, thereby predicting the contexts in which the construction occurs. (iii) We outline a method for empirically validating the theoretical implications of our analysis by means of corpus data.

2 Describing the variation

Within the German verbal domain, NP arguments of verbal heads marked with structural case are realised with nominative (NOM) and accusative (ACC) case, cf. (2a). When the verbal base is nominalised (e.g. with the suffix *-ung*) its argument structure is inherited by the derived nominal (Bierwisch 1989, Grimshaw 1990, Machicao y Priemer 2017, Machicao y Priemer & Müller 2021), cf. (2a) vs. (2b). According to the Case Principle (Przepiórkowski 1999: 65), GEN functions as the realisation of structural case within the nominal domain.

- (2) a. *Tim* behandelt *den* *Patienten*.
Tim.NOM treats the.ACC patient.ACC
‘Tim treats the patient.’
b. *Tims* Behandlung *des* *Patienten*
Tim.GEN treatment the.GEN patient.GEN
‘Tim’s treatment of the patient’

When realised postnominally,² the inherited arguments with structural case (i.e. GEN) can also be realised with a PP headed by the preposition *von* ‘of’ (cf. Machicao y Priemer 2017, Kopf & Bildhauer 2024), cf. (3a) & (3b) vs. (2b). In this case, *von* selects an NP realised with dative (DAT).

- (3) a. die Behandlung *vo-m* *Patienten*
the treatment of-the.DAT patient.DAT
‘the treatment of the patient’
b. die Behandlung *von Tim*
the treatment of Tim.DAT
‘the treatment of Tim’

As a result of this alternation – and assuming that selection is strictly local (cf. selectional localism, Pollard & Sag 1987, 1994, Sag 2007, 2012)

²In this paper, we concentrate on the alternation of GEN vs. *von* PP in postnominal position. For an account of the prenominal vs. postnominal realisation of arguments, cf. Machicao y Priemer & Müller (2021).

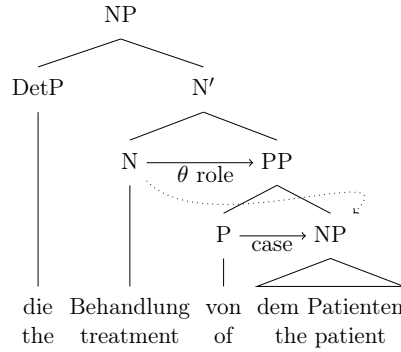


Figure 1: θ -role and case assignment

– the NP *dem Patienten* receives DAT from the preposition *von*. However, since the θ -role is determined and assigned by the head noun *Behandlung*, only the PP (and not the NP itself) can receive this θ -role, cf. Fig. 1. Note that the interpretation of the NP (or *von*-PP) arguments of the head noun parallels that of the corresponding arguments in the verbal domain, cf. (2a). Within the nominal domain an additional *possessor* interpretation is also available, which we likewise treat as an argument of the head noun (based on distributional evidence, cf. Machicao y Priemer & Müller 2021). Classifying these elements as arguments of the head noun – rather than as modifiers (contra Kopf & Bildhauer 2024) or arguments of an external head (contra Kratzer 1996) is crucial for an accurate account of the phenomenon. That is, one puzzle we encounter in dealing with this structure is how θ -role and case can be assigned by two different heads.

Furthermore, as pointed out in Kopf & Bildhauer (2024), not all *von*-PPs undergo the alternation. For instance, some nouns – e.g. *Abhängigkeit* ‘dependence’ – select a *von*-PP, cf. (4a). In this case, the *von*-PP is selected also by the adjectival base *abhängig* ‘dependent’, cf. (4b), and is inherited by the nominalised head.³

- (4) a. die *Abhängigkeit* { von seinen Eltern / *seiner Eltern }
the dependence on his parents his.GEN parents.GEN
INTENDED: ‘the dependence *on* his parents’
- b. Der Student ist [AP von seinen Eltern *abhängig*].
the student is on his parents dependent
‘The student depends on his parents.’

Another case in which the alternation does not take place is when the *von*-PP is a true modifier of the noun (e.g. temporal or local), cf. (5). In this

³This is similar to the selection of an *of*-PP by the verb *think*, which cannot alternate either, cf. *I think of Tim* vs. *Tim’s thinking*.

case, the preposition assigns case as well as a θ -role to the NP it combines with – in comparison to the *von* used in the alternation.

- (5) der Blick { vo-m Turm / *des Turms }
 the view from-the.DAT tower the.GEN tower.GEN
 INTENDED: ‘the view from the tower’

For an adequate description of the alternation, it is necessary to take into account that not all types of argument NPs can be realised with GEN case. The cases allowing both realisations (i.e. GEN as well as *von*-PP) include proper names, cf. (6a) & (7a),⁴ NPs with a definite determiner, cf. (6b) & (7b), and NPs with an indefinite determiner, cf. (6c) & (7c).

- | | |
|--|--|
| <p>(6) die Untersuchung ...
 the examination</p> <p>a. Tims
 Tim.GEN</p> <p>b. des Lehrers
 the.GEN teacher.GEN</p> <p>c. eines Lehrers
 a.GEN teacher.GEN</p> | <p>(7) die Untersuchung ...
 the examination</p> <p>a. von Tim
 of Tim.DAT</p> <p>b. vo-m Lehrer
 of-the.DAT teacher.DAT</p> <p>c. von einem Lehrer
 of a.DAT teacher.DAT</p> |
|--|--|

That is, a full NP can alternate, cf. (6) & (7), whereas an N' cannot serve as an argument either of the head noun (8a) or of the preposition (8b). This follows from an independent reason: German singular bare nouns are not fully saturated phrases, hence they cannot occur in argument positions.

- (8) a. *die Untersuchung *Lehrers*
 the examination teacher.SG.GEN
- b. *die Untersuchung *von Lehrer*
 the examination of teacher.SG.DAT

On the other hand, singular bare mass nouns (e.g. *Honig* ‘honey’) and plural count nouns (e.g. *Lehrer* ‘teachers’) are considered full NPs as well, as they can normally occupy argument positions without a determiner. However, they can appear only as arguments of the preposition *von*, cf. (10), and not as GEN marked arguments of the noun, cf. (9). Even more puzzling is the fact that singular bare mass nouns and plural count nouns without a determiner can appear as GEN marked arguments of the noun when they are modified, cf. (11a) & (11b) vs. (9).⁵ This is unexpected, since the modifier does not

⁴As a reviewer mentioned, it is not the case that proper names with GEN are allowed postnominally in all German dialects, i.e. (6a). See footnote 10 for further details.

⁵For the sake of completeness, it should be mentioned that *von* can combine with modified as well as with unmodified singular bare mass and plural count nouns.

turn a noun into a full NP. This can be seen with singular bare count nouns, which do not become fully saturated (i.e. NPs) by the addition of a modifier, cf. (8a) vs. (12)⁶ vs. (11b). Furthermore, given locality constraints, the selecting nominal head should not be able to determine whether its argument is modified or not.

- (9) die Untersuchung {*Honigs / *Lehrer}
the examination honey.SG.GEN teacher.PL.GEN
- (10) die Untersuchung {von Honig / von Lehrern}
the examination of honey.SG.DAT of teacher.PL.DAT
- (11) die Untersuchung ...
the examination
- a. *süßen* Honigs
sweet.SG.GEN honey.SG.GEN
- b. *korrupter* Lehrer
corrupt.PL.GEN teacher.PL.GEN
- (12) *die Untersuchung *korrupten* *Lehrers*
the examination corrupt.SG.GEN teacher.SG.GEN

That is, we have to account for the fact that the N-head and *von* combine with full NPs rather than with N', cf. (6)–(7) vs. (8). However, they differ w.r.t. their selectional restrictions: *von* can select “simple” or “complex” NPs, cf. (10), whereas the N-head selects only complex GEN NPs, cf. (11).

3 Grammatical constraints for alternation

Having identified the issues related to the alternation, we can now specify the grammatical constraints for our proposal. First, we require a lexical rule (LR) that derives *-ung* nouns and permits the inheritance of arguments from the verbal to the nominal stem. We adopt the LR from Machicao y Priemer & Müller (2021), cf. (13), which states that the arguments with structural case and the PP arguments of the verb are inherited by the derived *-ung* noun, cf. (2a) & (2b).⁷

⁶(12) is also not grammatical with the preposition *von*: **die Untersuchung von korruptem Lehrer* ‘the examination of corrupt teacher’.

⁷Following Przepiórkowski (1999), we divide the ARG-ST list into parametrised lists to inherit only the lists of arguments with structural case (i.e. GEN in the nominal domain) and of PP arguments, but not of arguments with lexical case.

(13) LR for *-ung*-nominalisation (cf. Machicao y Priemer & Müller 2021)

$$\left[\begin{array}{l} \text{stem} \\ \text{PHON } \boxed{1} \\ \text{SYNSEM|LOC|CAT } \left[\begin{array}{ll} \text{HEAD} & \text{verb} \\ \text{ARG-ST} & (\boxed{2} \text{list}_{str}) \oplus \text{list}_{lex} \oplus (\boxed{3} \text{list}_{pp}) \end{array} \right] \end{array} \right] \mapsto \left[\begin{array}{l} \text{ung-n-stem} \\ \text{PHON } \boxed{1} \oplus \langle \text{ung} \rangle \\ \text{SYNSEM|LOC|CAT } \left[\begin{array}{ll} \text{HEAD} & \text{noun} \\ \text{ARG-ST} & \boxed{2} \oplus \boxed{3} \end{array} \right] \end{array} \right]$$

For the mapping of elements from the ARG-ST list to the valence lists (e.g. COMPS), we assume argument structure mapping constraints (*as-mapping*) (cf. Manning & Sag 1998, Davis & Koenig 2000, Van Eynde 2015, Machicao y Priemer & Fritz-Huechante 2018, Machicao y Priemer & Müller 2021). To capture the fact that only “complex” NPs can be realised with GEN, (11) vs. (9), we make use of the LEX(ICALICALITY) feature (cf. Pollard & Sag 1987: 72–73).⁸ This feature distinguishes (roughly) between lexical and phrasal signs: LEX+ objects correspond to “words”, while LEX− objects result from the phrasal combination of two objects. Accordingly, “complex” NPs are characterised as output of a phrasal constraint, e.g. *head-adjunct-phrase* in (11). We implement this feature specification in the *as-mapping* constraint, such that only full NPs with the specification LEX− can be realised as complements of N.⁹ With this constraint in place – which is independently needed for all nouns – we can rule out N′ complements that are either LEX+ (8a) or LEX− (12), as well as full NPs that are LEX+, e.g. (9).¹⁰

(14) Constraint on ARG-ST to COMPS mapping

$$\left[\begin{array}{ll} \text{n-as-mapping} \\ \text{HEAD} & \text{noun} \\ \text{ARG-ST} & \langle \dots, \boxed{1} \text{ NP[LEX-]}, \dots \rangle \end{array} \right] \Rightarrow \left[\text{COMPS } \langle \dots, \boxed{1}, \dots \rangle \right]$$

As noted w.r.t. Fig. 1, the preposition *von* assigns DAT (more specifically *lexical dative*) to the NP, while it is the N-head that assigns the θ -role, resulting in a locality issue. Similar constructions are attested in other languages as well (for Dutch prepositions, see Van Eynde 2004; for *of* in English, see Van Eynde & Kim 2022; for *de* in Spanish, see Machicao y Priemer 2017; and for the *ass* camouflage construction, see Levine 2010). Following Levine

⁸See Arnold & Sadler (1992) and Deng et al. (2025) for further implementations of this feature in English and Mandarin Chinese.

⁹Recall that by virtue of the nominalisation LR in (13), only NPs with structural case are represented in the ARG-ST list of nouns.

¹⁰As mentioned above, the LEX+ vs. LEX− distinction corresponds “roughly” to the lexical–phrasal contrast. For instance, proper names such as *Tim* in (7a) can be classified as LEX−. This classification is also reflected in their behaviour w.r.t. modification and definiteness, and – depending on the analysis – they may even be derived via a phrasal constraint. Following our analysis, German dialects that do not allow proper names in GEN in postnominal position (cf. footnote 4), classify them as LEX+.

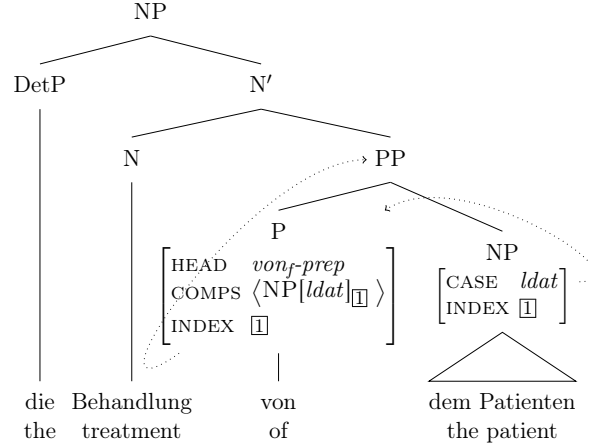


Figure 2: Parasitic head

(2010), we analyse this preposition as a *parasitic head*.¹¹ Specifically, we posit a distinct subtype of the PFORM value for the (functional) preposition *von* (PFORM *von_f*) – different from the one used in modification, cf. (5), which assigns a temporal or locative θ -role. This functional *von* (i.e. *von_f*) acts as a case assigner, selecting a NP with lexical dative (*dem Patienten* in Fig. 2) and taking the INDEX value of its complement (*dem Patienten*) as its own. Thus, the preposition can project its INDEX value to the maximal projection, which in turn receives a θ -role from the N-head (*Behandlung* in Fig. 2). Consequently, the complement PP of the N-head can bear either the agent or patient role, as illustrated in (3). The lexical entry for the preposition is provided in (15).¹²

(15) Lexical entry for functional *von*

PHON	$\langle von \rangle$
HEAD	$\left[\begin{array}{c} preposition \\ PFORM\ von_f \end{array} \right]$
COMPS	$\langle NP[ldat]_{\underline{1}} \rangle$
INDEX	$\underline{1}$

The final component of the grammatical analysis concerns the LR for the alternation. Although arguments realised with the preposition *von* are “more

¹¹A different approach worth considering is the functor analysis of prepositions (cf. Van Eynde 2004, Van Eynde & Kim 2022). However, adopting this approach would require more substantial changes to the overall system, since *von*-PPs would be analysed as (*von*-marked) NPs. As these issues are orthogonal to the concerns addressed here, we leave the exploration of this approach to future research.

¹²Consider that *von_f* is not only needed in this construction. It fulfils the requirements to reintroduce the external argument in passive constructions in German, where the preposition assigns case but no θ -role. For similar cases in Dutch, see Van Eynde (2004).

widespread” than those marked with GEN (cf. Sec. 4 & 5), the LR governing the alternation must take as its input an N-head that selects an NP with structural case. Recall that not all *von*-PPs participate in the alternation, cf. (5) & (4). Even in cases where *von_f* is used – this could be assumed for *Abhängigkeit* ‘dependence’ in (4) – the alternation is not possible. The crucial factor, then, is that a NP with structural case can be realised as a *von_f*-PP, whereas the reverse does not necessarily hold. Therefore, we propose the LR in (16), which states that a NP with structural case can be realised as a PP with the PFORM *von_f*.

(16) LR for GEN to *von_f*-PP alternation

$$\left[\begin{array}{c} n\text{-stem} \\ \text{CAT|ARG-ST} \end{array} \langle \dots, \text{NP}[\text{str}]_{\boxed{\text{I}}}, \dots \rangle \right] \mapsto \left[\text{CAT|ARG-ST} \langle \dots, \text{PP}[\text{PFORM } \textit{von}_f]_{\boxed{\text{I}}}, \dots \rangle \right]$$

With these constraints in place, we can correctly predict distribution of the alternation in the cases illustrated in Sec. 2. All constraints used are in any case needed for independent reasons, i.e. in other constructions. We just needed to strengthen the restrictions on these constraints, for instance, restricting the *as-mapping* constraint for nouns w.r.t. the LEX value. Furthermore, we provide a more adequate description of the functional counterpart of the preposition *von*, which is not only needed for this construction, but can be used for other phenomena as well, e.g. passive. In the next section, we show how the grammatical and use conditions interact to predict in which contexts the alternation occurs.

4 Grammar and use conditions

In the previous section, we proposed a set of constraints that capture the range of realisations of NP arguments in German. However, as the brief remarks about the situational distribution of each variants suggest, speakers know more than the purely structural licensing conditions for NP arguments: they also know when (and by whom) each type of argument is preferably used. The question we address in this section is the following: how do we relate these two types of knowledge – knowledge of grammatical constraints and knowledge of usage preferences – in a way that allows our theory not only to predict the pool of possible variants, but also to say something about their relative frequencies in different linguistic and situational contexts?

Building on prior work (Asadpour et al. 2022; Machicao y Priemer, Müller, Schäfer & Bildhauer 2022, Varaschin, Machicao y Priemer & Lu 2024), we answer this question by formulating a theory of register in HPSG. The cornerstones of this theory are what we call *use-conditional constraints* (UCCs), which take the general form in (17):

(17) *description of linguistic structure* $\mathcal{S} \Rightarrow$ *description of $\mathcal{S}$ ’s context*

Formally, UCCs are just ordinary implicational constraints. On the left hand side, a UCC identifies a range of independently licensed structures by means of a partial description, which can be a type or a more complex feature description. The right hand side is a partial specification of the features of the context that are required to make the structure acceptable. In the case of HPSG, the context is represented as a feature of signs themselves, under the attribute `CONTEXT` (Green 1994). So UCCs, like other constraints, simply state necessary conditions on the well-formedness of signs.

By using the HPSG formalism to state UCCs, we are hypothesizing that use-conditional knowledge is part of linguistic competence (Wilcock 1999, Paolillo 2000, Bender 2001, 2007, Asadpour et al. 2022; Varaschin, Machicao y Priemer & Lu 2024; i.a.). It is useful, however, to distinguish between those aspects of linguistic competence that pertain to the well-formedness of CAT, CONT and their correspondence (i.e. the domain of narrow syntax and semantics) from those that pertain to contextual appropriateness. There is suggestive evidence that the former are acquired sooner and processed more rapidly than use-conditional information (Münster & Knoeferle 2018, Soto & Almeida 2021, Plesca et al. 2024; i.a.). This is ensured by keeping UCCs separate from the core grammatical constraints.¹³ The set of UCCs is what Culy (1996) calls the “user’s manual” of a language.

In the case of register variants, UCCs constrain contexts via the assignment of *social meanings* to linguistic structures (Burnett 2019, Beltrama 2020, Salmon 2022). We use this term in a broad sense to refer to any non-truth-conditional semantic content that ascribes some socially-relevant property to a context coordinate, for instance, the speaker (s_c) or the addressee (a_c). We are assuming, therefore, a multidimensional semantics in the tradition of Kaplan (1999), Potts (2005), McCready (2010) and Gutzmann (2015), whereby a sentence like (18a) would have two layers of meaning: a *truth-conditional* (TC) and *social meaning* (SM), as in (18b–c).

- (18) a. Die Behandlung des Patienten war gut.
 the treatment the.GEN patient.GEN was good
 b. TC: $\{w : \text{the treatment of the patient was good in } w\}$
 c. SM: $\{c : s_c \text{ is presenting as educated in } c\}$

As (18) makes clear the TCs of an utterance u (TC_u) characterize a set

¹³For the sole purpose of defining the models of a grammar, it would often not matter whether we encode constraints on `CONTEXT` via dedicated UCCs or introduce them as parts of other (independently required) constraints. For instance, suppose we want to say a form derived via a lexical rule carries a use-condition. We could either state the constraint on the `CONTEXT` as part of the statement of the lexical rule or define a separate UCC identifying the outputs of the lexical rule in the antecedent. We opt for the latter strategy to capture the idea that UCCs and core grammatical constraints have a different cognitive status. Note, however, that some UCCs – namely, those that identify antecedents that are not mentioned in other parts of the grammar – are not reducible to other constraints.

of worlds. An utterance u is *true* in a specific world w iff $w \in TC_u$. In turn, the SMs of an utterance u (SM_u) characterize a set of contexts by ascribing a property to one of its participants. An utterance u is *usable* in a particular context c iff $c \in SM_u$. We refine this notion in gradient terms below.

SMs can be distinguished from most instances of TC meanings on the basis of the following properties (see Varaschin, Machicao y Priemer & Lu 2024 for more discussion):

- (19) a. **Independence:** SMs are immune to the effect of TC operators (e.g. negation, modals, conditionals, interrogatives).
- b. **Unplugability:** SMs project over presuppositional plugs like propositional attitude predicates, they cannot be evaluated w.r.t. to any parameter other than the local context.
- c. **Indexicality:** SMs predicate something of the present utterance situation or its participants (Potts 2007).
- d. **Gradability:** SMs hold of entities to different degrees.

The first two properties are the most important ones, as they reflect the way SMs interact with TC content. To illustrate (19a), note that (20) conveys the GEN-induced SM we stated in (18c), in spite of the fact that the very proposition to which GEN contributes its TC content is negated. (21) illustrates the related property in (19b): there is no reading in which the sentence conveys the SM that Lena (i.e. the attitude holder) is the one presenting as being educated. This SM projects to the speaker even under a reading where the existential presupposition associated with the definite *die Behandlung des Patienten* is not assumed to be satisfied in the world where the matrix clause is evaluated, but only in Lena’s belief worlds. That is, even if we assume that no patient was treated in the actual world, the SM that comes from GEN still projects. (22) illustrates the only context where SMs can be semantically embedded: direct quotation.

- (20) Die Behandlung des Patienten war *nicht* gut.
the treatment the.GEN patient.GEN was not good
- (21) Lena *glaubt*, dass die Behandlung des Patienten gut war.
Lena thinks that the treatment the.GEN patient.GEN good was
- (22) Lena sagte: “die Behandlung des Patienten war gut.”
Lena said the treatment the.GEN patient.GEN was good

We propose to model properties (19a)–(19b) by placing SMs as the values of the C(ONVENTIONAL)I(MPLICATION) attribute inside CONTEXT. In virtue of general linking constraints, TC operators (e.g. negation) and propositional attitudes only take their scopal arguments from elements under CONT. There is no TC operator that is a function over SM types. We model property (19c) by stating social-meaning relations in such a way that they always take a

C(ONTEXTUAL)-IND(ICE)S value as one of their arguments. The property in (19d) is modeled by requiring SM relations to take a DEG(REE) argument (i.e. an interval from 0 to 1), as in Potts & Kawahara (2004) and McCready (2019). As in standard degree semantics, each point d in this interval stands for an equivalence class (see, e.g., Kennedy 2001). In this case, ds are equivalence classes of contexts where the entity argument of the SM has the SM to a d -extent. This means that the wider the d range is the more contexts the SM is compatible with.

That said, we assume the following UCCs, cf. (23)–(25), for the phenomena at hand.¹⁴

(23) **UCC** for NPs with arguments with structural case (i.e. GEN):

$$\left[\begin{array}{c} \text{HEAD} \quad \textit{noun} \\ \text{ARG-ST} \quad \langle \dots, \text{NP}[\textit{str}], \dots \rangle \end{array} \right] \Rightarrow \left[\text{CTXT} \left[\begin{array}{c} \text{C-INDS|SPEAKER} \quad \boxed{1} \\ \text{CI} \left\langle \dots, \left[\begin{array}{c} \textit{educated} \\ \text{ARG1} \quad \boxed{1} \\ \text{DEG} \quad [.5, 1] \end{array} \right], \dots \right\rangle \end{array} \right] \right]$$

(24) **UCC** for NPs with *von*-PP arguments:

$$\left[\begin{array}{c} \text{HEAD} \quad \textit{noun} \\ \text{ARG-ST} \quad \langle \dots, \text{PP}[\text{PFORM} \textit{von}], \dots \rangle \end{array} \right] \Rightarrow \left[\text{CTXT} \left[\begin{array}{c} \text{C-INDS|SPEAKER} \quad \boxed{1} \\ \text{CI} \left\langle \dots, \left[\begin{array}{c} \textit{educated} \\ \text{ARG1} \quad \boxed{1} \\ \text{DEG} \quad (0, .7] \end{array} \right], \dots \right\rangle \end{array} \right] \right]$$

(25) **UCC** for *-ung* nouns:

$$\textit{ung-n-word} \Rightarrow \left[\text{CTXT} \left[\begin{array}{c} \text{C-INDS|SPEAKER} \quad \boxed{1} \\ \text{CI} \left\langle \dots, \left[\begin{array}{c} \textit{educated} \\ \text{ARG1} \quad \boxed{1} \\ \text{DEG} \quad [.6, 1] \end{array} \right], \dots \right\rangle \end{array} \right] \right]$$

Each of these SMs assigns to a class of structures a SM that picks out different sets of contexts where the speaker is presenting as being *educated* to a specific degree. (23), for instance, assigns to *nouns* that have an argument marked with structural case (i.e. GEN) a social meaning that represents a set of contexts where the speaker is presenting as being educated to mid-to-high extent. This captures the intuition that GEN sound more formal or educated in German. (25) assigns a similar SM (understood as a range of admissible contexts) to words that result from *ung*-derivations.

Of course this proposal is, as it stands, highly simplified. Forms subject to register variation are also typically not associated with a specific SM

¹⁴As in McCready (2019), the particular DEG intervals are estimated based on observations from the literature and on native speakers' intuitions about the situational distribution of the corresponding expressions. The kind of quantitative corpus-based methodology we discuss in Sec. 5 could be used to arrive at or motivate such intervals in a more data-driven fashion. But what is most important here is not the exact numerical values, but their relative differences and how they interact when different variants are used together.

(e.g. *educated*), but with an indexical field of inferentially related SMs (cf. Eckert 2008). This can be modelled by associating these structures with an underspecified type that includes *educated* as a subtype, so that SMs enforced by UCCs are only resolved to maximal types like *educated* in concrete communicative situations. For simplicity, we do not implement underspecification here, but a concrete worked-out proposal is developed in Varaschin, Machicao y Priemer & Lu (2024).

The UCCs account for the SMs that the grammar assigns to individual pieces of structures. To calculate the SM of a full utterance, we need a principle of composition. A statement of the principle proposed by Varaschin, Machicao y Priemer & Lu (2024) is shown in (26). We say that two predication P_1 and P_2 are repeated predication *iff* they are of the same type and have the same ARG values, but possibly different DEG(REE) values.

(26) **Local CI Composition Principle**

For each phrase,

- a. If the CI values of its daughters *do not have repeated predication*s, then the CI value of the phrase is the concatenation of the CI values of its daughters.
- b. If the CI values of its daughters *have repeated predication*s $SM_1, \dots SM_n$ then the CI value of the phrase is the concatenation of the CI values of its daughters: (i) minus $\langle SM_1 \rangle, \dots \langle SM_n \rangle$; (ii) plus a list of predication of the same type and with the same ARG values as $SM_1, \dots SM_n$, but with a DEG value consisting in the intersection between the DEG values of $SM_1, \dots SM_n$.

In cases where repeated SMs do not intersect, the mother phrase will have a SM with an empty DEGREE value, which represents that the set of contexts where it is admissible is the empty set. This is arguably what happens with honorific mismatches (Kaur & Yamada 2022, i.a.). Crucially, however, having an empty DEGREE value does not mean that the structure is ungrammatical (i.e. not a model of the grammar), only that it is unusable, given that no utterance context will be an element of the empty set.

If the DEGREE values of repeated predication do intersect, they get progressively narrowed as one moves upwards in the tree. The core claim is that SM composition within a clause is monotonic: constraints on the context accumulate (i.e. each SM-bearing daughter *adds* restrictions) either via concatenation or via intersection of repeated SM predication. As we said, narrower DEGREE intervals represent a narrower set of admissible contexts.

In the case of German NPs, the core empirical prediction this approach makes, when combine with the UCCs in (23)–(25) is the following: *-ung* nouns and *von_F*-PP argument combinations should be more contextually restricted than combinations between *-ung* and GEN. Figs. 3–4 illustrate the effects of SM composition in each of these cases. What is interesting here is that, due

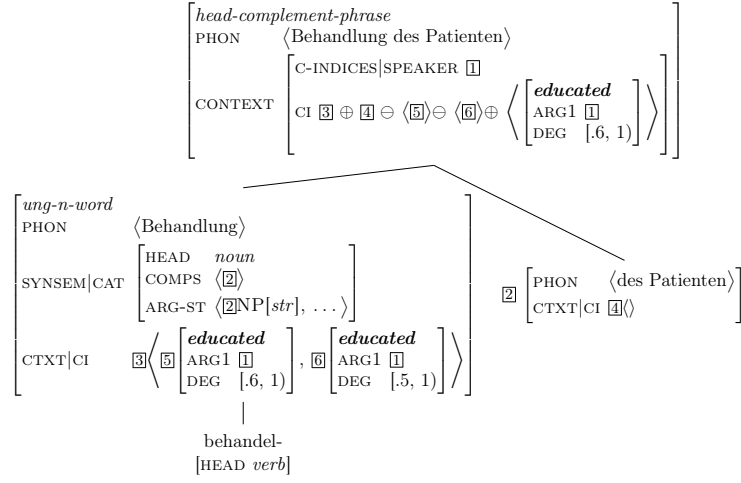


Figure 3: SM composition with *-ung* and GEN arguments

to the structures targeted by the UCCs, repeated SM predications have to be present at the lexical level. The composition principle in (26) applies when such words undergo phrasal combination.

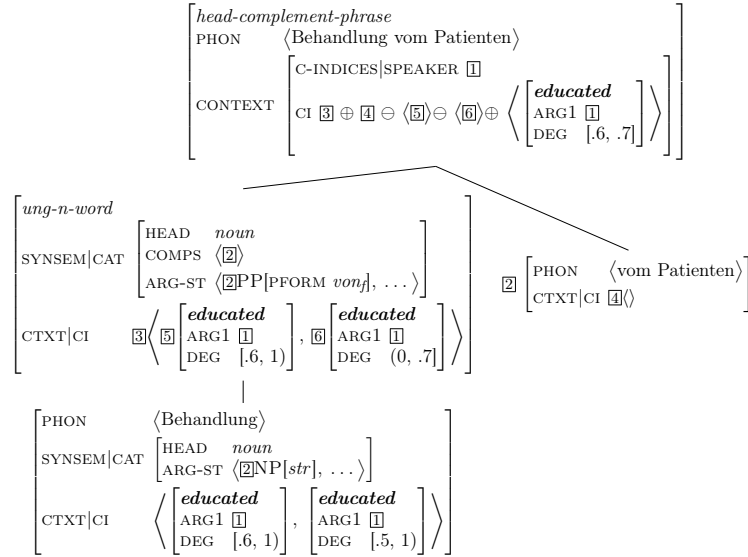


Figure 4: SM composition with *-ung* and *von_f*-PP arguments

If felicitous, the output of Local CI Composition (26) updates the global context – i.e. the context that pertains to the broader discourse beyond the sign (Paolillo 2000). Contextual update is subject to the following constraint:

(27) **Felicity constraint** (cf. McCready 2019: 31)

For every utterance u expressing a SM α , if the **prior global context** of u is specified as having a SM α' , where α and α' are repeated predication, then the DEGREE values of α and α' have to intersect.

The bigger the overlap between the SMs in the prior global context and those in u 's CONTEXT|CI value, the more appropriate u is with respect to the context. On this view, we can think of a register as a cluster of linguistic constraints whose associated models are required (by virtue of UCCs) to carry SMs that are appropriate in the same global contexts. Therefore, whether or not a form *belongs* to a register R is a matter of degree, i.e. it depends on how much its SMs overlap with the contextual parameters associated with R .

5 Empirical Validation

In this section, we argue that the proposed set of grammatical constraints and the proposed UCCs with respective DEG-values need to be empirically validated. Furthermore, we argue that all claims about register-sensitive grammatical structures need to be supported by empirical data. In order to investigate the register sensitivity of GEN and *von_f*-PPs, we analyse corpus data from the PreCOXX25-LDA corpus (Schäfer, Bildhauer, Reiß, Pankratz & Müller 2025). Our goal is to test whether our theoretical predictions are empirically supported and, if necessary, make adjustments to our theoretical modelling. In this section, we will outline our methodological approach, which should be seen as a proposal on how to motivate numerical values in UCCs and SMs.¹⁵

The PreCOXX25-LDA corpus is a large-scale corpus containing 21,775,285 tokens across 2,475 documents. All documents show a wide variety of language representation, ranging from internet forums to sports reports and legal texts. We chose the corpus for its large number of tokens and linguistic heterogeneity. The corpus was built by Schäfer, Bildhauer, Reiß, Pankratz & Müller (2025) with a probabilistic approach based on Latent Dirichlet Allocation (LDA) (Blei 2012). LDA is a statistic model originally used for topic modelling, which posits that documents are mixtures of topics and topics are mixtures of words. As a result, documents are being assigned weights representing the probability that each document is associated with a specific topic. In the process of constructing the corpus, it was used to discover latent register dimensions to describe the probabilities of documents to belong to different potential registers (or *pregisters*), allowing the documents to be described as a *mixture* of potential registers. In the experiment, 25 so-called preregisters

¹⁵A complete and commented version of the code is available via Github https://github.com/paureiss/GEN_vs_VON_HPSG_2025.git.

were extracted. Each of these can be interpreted as a cluster of documents sharing similar linguistic features.

Since the analysis of register in language requires the analysis of situational and functional parameters, the clustering of texts into registers was completed by manual annotation. For the analysis, a subset of documents was extracted from the corpus and manually annotated by four human experts. The documents were classified according to situational and functional parameters: *Education*, *Interactivity*, *Proximity* and *Narrativity*, achieving a substantial to almost perfect inter-rater agreement across categories. Due to the combination of LDA-based analysis and human annotation, the corpus provides a robust empirical basis for analysing register-sensitive grammatical structures and enables the nuanced analysis of register mixtures in texts.

In our analysis, we focus on the *Education* parameter due to its established association with nominal style and genitive constructions (cf. Biber 1988). In the manual annotation, *Education* was annotated using binary labels, following “whether a document requires an elevated educational background (EEB) on the side of the reader” (Schäfer, Bildhauer, Reiß, Pankratz & Müller 2025: 16) and achieved an overall substantial inter-rater agreement (Fleiss’s $\kappa = .67$).

For our analysis, we begin by ranking the registers by their *Education* annotation scores (drawn from the manual annotation), normalised on a continuum using *min-max* scaling. Registers with a larger proportion of documents annotated as *high education* receive a higher *Education* score whereas registers with a larger proportion of documents annotated as *low education* receive a lower score. These serve as proxies for the education parameter. We then compute whether the number of extracted GENS and *von_f*-PPs correlate with the assigned *Education*-scores. This methodology enables the comparison with the hypothesized DEG-values in the corresponding constructions.

For the analysis, we conducted a pre-study in order to pre-validate the hypothesis that genitives and *von_f*-PPs correlate in terms of their register usage. The pre-study was conducted without further annotation in order to ascertain the hypothesis’s plausibility and to obtain a preliminary insight.

To visualize the relationships between *gens*, *von_f*-PPs and the *Education*-scores, we created a heat-map (cf. Fig. 5) to identify patterns, correlations, and anomalies within the dataset. The colour gradient ranging from blue (-1.0) to red ($+1.0$) represents the strength of correlation, blue indicating a negative and red indicating a positive correlation. The heat-map reveals indications of weak to moderate correlations in the expected direction. GENS and *von_f*-PPs show a perfect negative correlation of -1 as the existence of one excludes the other. *ung*-nouns generally show a moderate positive correlation with the *Education* parameter (0.49). While genitives show a weak positive correlation (0.28) with *Education*, *von_f*-PPs show a weak negative correlation (-0.28) with it – as expected. This results enables us to infer that genitives

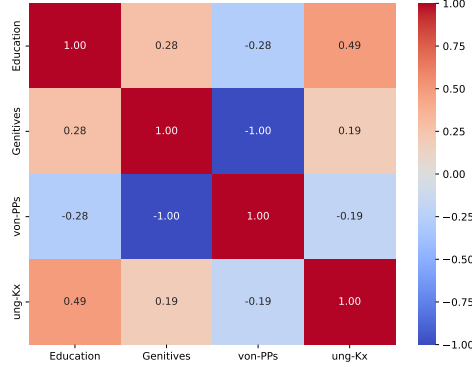


Figure 5: Correlation heatmap of all target variables, based on Spearman’s rank correlation coefficient (Kx = Construction)

and *-ung*-nouns trend positively with higher educational registers.

While the heat-map provides valuable insights, it is only a simplification of the studied structures. Still, these findings support the broader hypothesis of register-sensitive distribution patterns, as expected by our theoretical modelling. Therefore, we can deduce that the combinations of *von*-phrases with *-ung*-nouns should have small DEG values for their educated SM, rendering them more contextually restricted.

Nonetheless, validation via manual annotation is essential. As stated in Sec. 2, the extent to which a speaker has the freedom to choose between the two constructions depends on the context. As a second step, we therefore conducted a manual annotation experiment to empirically validate the degree of optionality between GENs and *von*-PPs in our dataset. The study aimed to determine whether speakers truly have the freedom to choose between the GEN and *von*-PP in different contexts or if certain contexts favour one construction over the other. In order to maintain and verify the quality of annotations, the annotation was performed by two expert annotators. The quality of the annotation is calculated via inter-rater agreement (IRA) using Fleiss’s κ .

In total, we conducted two annotation rounds, with the endeavour to fine-tune the annotation guidelines and to ensure sufficient inter-annotator agreement. The first round consisted of $n = 250$ items of two exemplary registers (register 10 and register 24), resulting in $n = 500$ in total. The registers were chosen based on their general high number of *ung*-derived nouns and their contrasting *Education*-scores. Subsequent to the revision of discrepancies in the annotation guidelines, a second round was conducted to check the operationalisation of the annotation and comprehensibility and clarity of the annotation guidelines. The second round consisted of $n = 250$

items from the same registers as well as an additional pre-register (10, 14 and 24), resulting in $n = 750$ in total. The final IRA of Fleiss’s $\kappa = .701$ indicates substantial agreement between the annotators, suggesting that the annotation guidelines were clear and consistently applied.

The annotation guidelines are based on the discussed literature in Sec. 2. First, we excluded all elements comprising proper names as attributes as they differ with regard to their syntactic properties and are, therefore, not suitable for comparison. In a second step, we filtered the data based on optionality by excluding all items where speakers do not have choice over the construction. Further, we excluded all items that are fixed constructions or idioms as in (28a). Additionally, all *von*-PPs were excluded if *von* had a temporal or local interpretation as in (28b) instead of being interpreted as an argument of the head noun. If the *von*-PP did not contain a determiner as in (28c) they were excluded as well, as GEN nouns are obligatorily marked with a determiner. This implies that no equivalent construction exists with GEN. Lastly, all items were excluded if *von* was part of the argument structure of the verbal base as in (28d).

- (28) a. Tag der Arbeit
day the.GEN work.GEN
‘Labour Day’
b. der Blick vo-m Turm
the view from-the.DAT tower.DAT
‘the view from the tower’
c. Beförderung von Hunden
transport of dogs.PL.DAT
‘transport of dogs’
d. Ausschluss vo-m Wahlrecht
exclusion from-the.DAT voting.right.SG.DAT
‘exclusion from voting rights’

The stepwise testing allowed us to establish the annotation guidelines and confirm their consistency, thus providing a solid foundation for a methodologically robust analysis.

Considering the two previously outlined annotation rounds it is evident, that establishing clear and consistent guidelines is essential prior to conducting a larger-scale empirical study. Despite the fact that the dataset is not sufficiently large and does not allow for a sophisticated statistical analysis, both annotation rounds still serve as a pre-study to pre-validate our hypothesis and to fine-tune the guidelines for the annotation. Only three registers were annotated to gain preliminary insights into the data. Preliminary findings suggest that GEN and *-ung* nouns trend positively with higher education registers and that *von*-PPs show a slight negative correlation. While a larger

statistical study is still pending, our methodology sets a strong foundation for validating the theoretical predictions regarding the encoding of register-sensitive grammatical structures, particularly in relation to education-based register variations. Once a larger study has been conducted, the data can be used to model the respective DEGREE values in the corresponding UCCs.

The presented methodological approach represents one of several possible methods for empirical validation. It has potential applications for testing formal hypotheses about register and social meaning. Still, we wish to stress that numerical values in UCCs and SMs should ultimately not be arbitrary but empirically verified. This can be done on the basis of a quantitative corpus methodology, or average ratings derived from matched-guise experiments (cf. e.g. Lüdeling et al. 2024; Varaschin, Machicao y Priemer & Lu 2024).

In this section, we presented a method combining probabilistic modelling (LDA) and human annotation for validating register-sensitive grammatical encoding. Future empirical work should expand the annotation to establish robust empirical intervals. While the analysis provides evidence supporting our hypothesis, further research is needed to refine the DEG values for our theoretical model.

6 Concluding remarks

In this paper, we proposed (i) structural constraints that predict when GEN and *von*-PP arguments of NPs in German *can* be used: (ii) use-conditional constraints (UCCs) that model when the two forms *would* be used, and (iii) a new method to empirically validate the predictions made by the model that results from (i)–(ii). The latter relied on the automatic classification of texts into potential registers (i.e. *pregisters*) by means of the Latent Dirichlet Allocation (LDA) model and on their subsequent manual annotation for situational parameters (Schäfer, Bildhauer, Reiß, Pankratz & Müller 2025).

A crucial step in our methodological proposal was to take the LDA-induced preregisters ordered by their annotation-derived *education* scores as proxies for the *education* SM parameter itself. This assumes that how much a structure appears on a preregister assigned to a particular degree of education reflects how appropriate it is in a context where the speaker presents as having that particular degree of education.¹⁶ This appropriateness can then be compared to the hypothesized DEGREE values we assign to the corresponding *educated* SM in UCCs like (23)–(25). If the empirically induced *education* intervals don’t match the hypothesized values, the latter can be revised accordingly.

¹⁶This is a simplification (since many SM parameters are active in any situation), but it serves as a heuristically useful approximation for empirically validating hypothesized SMs. In this respect, this technique is similar to the use of written vs. spoken corpora as proxies for *formality*, as proposed in Sauerland (2022). The difference is that our association between corpora and situational parameters is independently established by annotators.

A further step is to develop some technique for automatically extracting SM DEGREE values from the distribution of different variants across different registers annotated for that SM. There are potentially different ways this could be done. One could, for instance, derive such values from the relative proportion of different variants across registers ranked by their scores for a particular SM (e.g. *education*), normalized to [0,1]. The degree interval could be defined as the range where the variant appears with frequency above some threshold (e.g. above the median frequency, or above a percentage of its maximum frequency). A simpler alternative, following the logic of Sauerland (2022), is to calculate the frequency of a variant in the register ranked highest with respect to a particular SM annotation, divide it by the frequency of the same variant in the register ranked lowest with respect to the same SM annotation and normalise the resulting number to correspond to a value between [0,1]. This provides a single point estimate reflecting a variant’s degree of association with contexts associated with a SM. To convert this to an interval, one could establish bounds using confidence intervals derived from the empirical variance, or by setting the point estimate as the interval’s midpoint with width determined by the variant’s dispersion across registers.

On the theoretical side, our proposal says that genitive arguments are more frequent in ‘more educated’ contexts because they are grammatically required (by UCCs) to have SMs that largely overlap with these contexts. The higher prevalence of genitives in registers with high education scores was confirmed by our preliminary data analysis. According to the SM composition principle in (26), *-ung* derived Ns and *von*-PPs are predicted to have a restricted distribution because their hypothesized SMs define sets of admissible contexts with a narrow intersection. The preliminary data analysis we carried out favors this intersective approach to local SM composition. This goes against proposals like McCready’s (2019), who argues that DEGREE values of SMs of the same type are composed by calculating a numerical average.

7 Acknowledgements

We thank the audience at the 32nd HPSG Conference (Centro de Linguística da Universidade de Lisboa), three anonymous reviewers, and Sigi Bahr, Felix Bildhauer, Elin McCready, Andrés Saab, Manfred Sailer, and Frank Van Eynde for invaluable comments and conversations. All remaining errors are ours. This research was funded by the Deutsche Forschungsgemeinschaft (DFG) – SFB 1412, 416591334, Project A04.

References

- Arnold, Doug J. & Louisa Sadler. 1992. Noun-modifying adjectives in HPSG. *Working Papers in Language Processing* (35).

- Asadpour, Hiwa, Shene Hassan & Manfred Sailer. 2022. Non-*wh* relatives in English and Kurdish: Constraints on grammar and use. In Stefan Müller & Elodie Winckel (eds.), *Proceedings of the 29th International Conference on Head-Driven Phrase Structure Grammar, Nagoya University & Institute for Japanese Language and Linguistics*, 6–26. Frankfurt am Main: University Library. DOI: 10.21248/hpsg.2022.1.
- Beltrama, Andrea. 2020. Social meaning in semantics and pragmatics. *Language and Linguistics Compass* 14(9). 1–20. DOI: 10.1111/lnc3.12398.
- Bender, Emily M. 2001. *Syntactic variation and linguistic competence: The case of AAVE copula absence*. Stanford, CA: Stanford University. (Doctoral dissertation).
- Bender, Emily M. 2007. Socially meaningful syntactic variation in sign-based grammar. *English Language & Linguistics* 11(2). 347–381. DOI: 10.1017/S1360674307002286.
- Biber, Douglas. 1988. *Variation across Speech and Writing*. 1st edn. Cambridge University Press. DOI: 10.1017/CB09780511621024.
- Bierwisch, Manfred. 1989. Event nominalization: Proposals and problems. In Wolfgang Motsch (ed.), *Wortstruktur und Satzstruktur* (Linguistische Studien: Reihe A. Arbeitsberichte 194), 1–73. Berlin: Akademie der Wissenschaften der DDR, Zentralinstitut für Sprachwissenschaft.
- Blei, David M. 2012. Probabilistic topic models. *Communications of the ACM* 55(4). 77–84. DOI: 10.1145/2133806.2133826.
- Burnett, Heather. 2019. Signalling games, sociolinguistic variation and the construction of style. *Linguistics and Philosophy* 42(5). 419–450.
- Culy, Christopher. 1996. Null objects in English recipes. *Language Variation and Change* 8(1). 91–124.
- Davis, Anthony & Jean-Pierre Koenig. 2000. Linking as constraints on word classes in a hierarchical lexicon. *Language* 76(1). 56–91.
- Deng, Chenyuan, Antonio Machicao y Priemer & Giuseppe Varaschin. 2025. Unifying modifiers, classifiers and demonstratives. In Gabriela Bîlbîie & Gerhard Schaden (eds.), *Empirical issues in syntax and semantics: selected papers from CSSP 2023*, 57–90. Berlin: Language Science Press. DOI: 10.5281/zenodo.15450432.
- Dölling, Johannes. 2015. Sortale Variation der Bedeutung bei *ung*-Nominalisierungen. In Christian Fortmann, Anja Lübke & Irene Rapp (eds.), *Situationsargumente im Nominalbereich* (Linguistische Arbeiten 562), 49–92. Berlin: de Gruyter Mouton.
- Eckert, Penelope. 2008. Variation and the indexical field. *Journal of sociolinguistics* 12(4). 453–476. DOI: 10.1111/j.1467-9841.2008.00374.x.
- Ehrich, Veronika & Irene Rapp. 2000. Sortale Bedeutung und Argumentstruktur: *ung*-Nominalisierungen im Deutschen. *Zeitschrift für Sprachwissenschaft* 19(2). 245–303.
- Green, Georgia M. 1994. The structure of context: The representation of pragmatic restrictions in HPSG. In James H. Yoon (ed.), *Proceedings*

- of the 5th Annual Conference of the Formal Linguistics Society of Mid-America, vol. 1/2 (Studies in the Linguistic Sciences), 215–232. Urbana, IL: Department of Linguistics, University of Illinois.
- Grimshaw, Jane. 1990. *Argument structure*. Cambridge: MIT Press.
- Gutzmann, Daniel. 2015. *Use-conditional meaning: Studies in multidimensional semantics*. Oxford: Oxford University Press.
- Kaplan, David. 1999. The meaning of *ouch* and *oops*.
- Kaur, Gurmeet & Akitaka Yamada. 2022. Honorific (mis)matches in allocutive languages with a focus on Japanese. *Glossa: A Journal of General Linguistics* 44(1). DOI: 10.16995/glossa.7669.
- Kennedy, Christopher. 2001. Polar opposition and the ontology of ‘degrees’. *Linguistics and Philosophy* 24. 33–70. DOI: 10.1023/A:1005668525906.
- Kopf, Kristin & Felix Bildhauer. 2024. The genitive alternation in German. *Corpus Linguistics and Linguistic Theory*. 1–35. DOI: 10.1515/cllt-2024-0017.
- Kratzer, Angelika. 1996. Severing the external argument from its verb. In Johan Rooryck & Laurie Zaring (eds.), *Phrase structure and the lexicon* (Studies in Natural Language and Linguistic Theory 33), 109–137. Dordrecht: Kluwer Academic Publishers.
- Levine, Robert D. 2010. The ass camouflage construction: Masks as parasitic heads. *Language* 86(2). 265–301.
- Lüdeling, Anke, Luka Szucsich, Lars Erik Zeige, Aria Adli, Artemis Alexiadou, Malte Belz, Miriam Bouzouita, Oliver Bunk, Malte Dreyer, Markus Egg, Anna Helene Feulner, Jürg Fleischer, Natalia Gagarina, Aron Hirsch, Stefanie Jannedy, Pia Knoeferle, Thomas Krause, Silvia Kutscher, Liu Mingya, Beat Lütke, Antonio Machicao y Priemer, Katja Maquate, Laura Merino Hernández, Roland Meyer, Christine Mooshammer, Stefan Müller, Uli Sauerland, Antje Sauermann, Viola Schmitt, Nicole Schumacher, Dina Serova, Stephanie Solt, Jozina Vander Klok, Elisabeth Verhoeven, Richard Waltereit, Melanie Weirich & Heike Wiese. 2024. Register: Language users’ knowledge of situational-functional variation – Frame text of the second phase proposal for the CRC 1412. *Register Aspects of Language in Situation (REALIS)* 3(1). 1–53. DOI: 10.18452/29131.
- Machicao y Priemer, Antonio. 2017. *NP-arguments in NPs: An analysis of German and Spanish noun phrases in Head-Driven Phrase Structure Grammar*. Berlin: Humboldt-Universität zu Berlin. (Doctoral dissertation). DOI: 10.18452/20109.
- Machicao y Priemer, Antonio & Paola Fritz-Huechante. 2018. Korean and Spanish psych-verbs: Interaction of case, theta-roles, linearization, and event structure in HPSG. In Stefan Müller & Frank Richter (eds.), *Proceedings of the 25th International Conference on Head-Driven Phrase Structure Grammar*, 153–173. Stanford: CSLI Publications. DOI: 10.21248/hpsg.2018.10.

- Machicao y Priemer, Antonio & Stefan Müller. 2021. NPs in German: Locality, theta roles, possessives, and genitive arguments. *Glossa: A Journal of General Linguistics* 6(1). 1–38. DOI: 10.5334/gjgl.1128.
- Machicao y Priemer, Antonio, Stefan Müller, Roland Schäfer & Felix Bildhauer. 2022. Towards a treatment of register phenomena in HPSG. In Stefan Müller & Elodie Winckel (eds.), *Proceedings of the 29th International Conference on Head-Driven Phrase Structure Grammar, Nagoya University & Institute for Japanese Language and Linguistics*, 86–101. Frankfurt am Main: University Library. DOI: 10.21248/hpsg.2022.5.
- Manning, Christopher D. & Ivan A. Sag. 1998. Argument structure, valence, and binding. *Nordic Journal of Linguistics* 21(2). 107–144. DOI: 10.1017/S0332586500004236.
- McCready, Elin. 2010. Varieties of conventional implicature. *Semantics and Pragmatics* 3. 8–1. DOI: 10.3765/sp.3.8.
- McCready, Elin. 2019. *The semantics and pragmatics of honorification: Register and social meaning* (Oxford Studies in Semantics and Pragmatics 11). Oxford: Oxford University Press.
- Münster, Katja & Pia Knoeferle. 2018. Extending situated language comprehension (accounts) with speaker and comprehender characteristics: Toward socially situated interpretation. *Frontiers in psychology* 8. 1–12.
- Paolillo, John C. 2000. Formalizing formality: An analysis of register variation in Sinhala. *Journal of Linguistics* 36(2). 215–259. DOI: 10.1017/S0022226700008148.
- Plesca, Ana-Maria, Katja Maquate & Pia Knoeferle. 2024. Register and standard language knowledge in interaction. *Register Aspects of Language in Situation* 3(2). 1–35. DOI: 10.18452/28042.
- Pollard, Carl & Ivan A. Sag. 1987. *Information-based syntax and semantics. volume 1: Fundamentals*. Stanford: CSLI Publications.
- Pollard, Carl & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago, IL: University of Chicago Press.
- Potts, Christopher. 2005. *The logic of conventional implicatures*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199273829.001.0001.
- Potts, Christopher. 2007. The expressive dimension. *Theoretical Linguistics* 33(2). 165–198. DOI: <https://doi.org/10.1515/TL.2007.011>.
- Potts, Christopher & Shigeto Kawahara. 2004. Japanese honorifics as emotive definite descriptions. In Robert B. Young (ed.), *Proceedings of the 14th Semantics and Linguistic Theory Conference (SALT)*, vol. 14, 253–270. Northwestern University: The Linguistic Society of America. DOI: 10.3765/salt.v14i0.2917.
- Przepiórkowski, Adam. 1999. *Case assignment and the complement/adjunct dichotomy: A non-configurational constraint-based approach*. Eberhard-Karls-Universität Tübingen. (PhD thesis). <https://publikationen.uni-tuebingen.de/xmlui/handle/10900/46147> (19 July, 2015).

- Sag, Ivan A. 2007. Remarks on locality. In Stefan Müller (ed.), *Proceedings of the 14th international conference on Head-Driven Phrase Structure Grammar*, 394–414. Stanford, CA: CSLI Publications.
- Sag, Ivan A. 2012. Sign-based construction grammar: A synopsis. In Hans C. Boas & Ivan A. Sag (eds.), *Sign-based construction grammar*, 61–197. Stanford, CA: CSLI Publications.
- Salmon, William. 2022. Social markers and dimensions of meaning. *Journal of Pragmatics* 192. 98–115. DOI: <https://doi.org/10.1016/j.pragma.2022.02.014>.
- Sauerland, Uli. 2022. Quantifying the register of German quantificational expressions: A corpus-based study. In Nicole Gotzner & Uli Sauerland (eds.), *Measurements, numerals and scales: essays in honour of Stephanie Solt*, 261–273. Cham: Springer.
- Schäfer, Roland, Felix Bildhauer, Pauline Reiß, Elizabeth Pankratz & Stefan Müller. 2025. Probabilistic register modelling. <https://rolandschaefer.net/wp-content/uploads/2024/07/modellingregisters-1.pdf> (25 September, 2025).
- Soto, Marije & Wellington Couto Almeida. 2021. Entre a agramaticalidade e a variação: concordância verbal, sociolinguística e neurociência da linguagem. *ReVEL* 19(36). 1–29.
- Van Eynde, Frank. 2004. Minor adpositions in Dutch. *Journal of Comparative Germanic Linguistics* 7. 1–58.
- Van Eynde, Frank. 2015. *Predicative constructions: From the Fregean to a Montagovian treatment*. Stanford: CSLI Publications.
- Van Eynde, Frank & Jong-Bok Kim. 2022. Pseudo-partitives in English: An HPSG analysis. *English Language and Linguistics* 27(2). 271–302.
- Varaschin, Giuseppe, Antonio Machicao y Priemer & Yanru Lu. 2024. Topic drop in German: Grammar and usage. In Stefan Müller (ed.), *Proceedings of the 31st international conference on Head-Driven Phrase Structure Grammar*, 193–217. Frankfurt am Main: University Library. DOI: 10.21248/hpsg.2024.12.
- Wilcock, Graham. 1999. Lexicalization of context. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and constructional aspects of linguistic explanation*, 373–387. Stanford, CA: CSLI Publications.