Interpretation and Structure: Conditions for Determiner Omission in Nominal Complements of Prepositions

Prof. Dr. Tibor Kiss
Universität Stuttgart, 24.06.2010

Sprachwissenschaftliches Institut
Some basic facts on PNCs

- Preposition Noun Combinations (PNCs) are phrases containing at least a preposition and a singular count noun, but no determiner.

  (1) auf Anfrage (after being asked), ohne Gewinnchance (without a chance to win), unter Androhung (under threat), mit Vorbehalt (with reservation), ...

- They can be extended prenominally and postnominally (and can also appear in governed positions).

  (2) auf parlamentarische Anfrage (after being asked in parliament), bei absolut klarer Zielsetzung (given a clearly present aim), mit schwer beladenem Rucksack (with heavily loaded backpack), unter sanfter Androhung (under gentle threat), ...

  (3) Er wehrt sich gegen die Forderung nach Stilllegung einer Verbrennungsanlage. he defends REFL against the demand for closedown an incineration plant ‘He defies the demand for closing an incineration plant.’

- English speaking linguists use the terms „determinerless PP“ and „bare PP“, respectively. These are obvious misnomers.
A definition

- **Preposition Noun Combination (PNC):**
  - A PNC is a phrase consisting of a preposition and a determinerless nominal projection (governed by the preposition), the syntactic head of which is realized in the singular and can be classified as a count noun.
  - We assume that for PNCs, it should always be possible to add a determiner to the nominal projection, and hence turn the PNC into an ordinary PP.
Why PNCs are special

- **German:**
  - *Duden rule 442* requires that determiners are obligatory for singular count nouns – otherwise they are assumed to be ungrammatical
  - There are some exceptions captured in *Duden rule 395*
    - The rule supposes that exceptions can be listed
    - Exceptions are postulated to be from specific genres and registers

- **Crosslinguistic:**
  - “However, ‘irregularities’ ... recur in language after language, provided they have article-like elements, and can therefore **claim to be regular from a crosslinguistic point of view.**”
    (Himmelmann 1998:316)

- **PNCs are** **productive** (Kiss 2007; Dömges et al. 2007)
Ways out: Countability

- **Countability** is a property of nominal projections, which is determined by contextually assigned features such as plural morphology, or realisation of a determiner (cf. Allan (1980), Borer (2005), Bale and Barner (2009)).

- According to this assumption, **nouns are not countable per se**, but will receive an appropriate interpretation if realized in a proper context.

  4. Fahrzeug (unspecified), Fahrzeug+PL = Fahrzeuge (+count), ein Fahrzeug (+count)

  5. Hier bekommen Sie mehr Fahrzeug für ihr Geld. ‘In this place, you get more car for your money.’

  6. Hier bekommen Sie mehr Fahrzeuge für ihr Geld. ‘In this place, you get more cars for your money.’

- But there is no such context in a PNC.
  - Consequently, there is no rule violation, and each and every PNC ought to be classified as grammatical.
Countability and Grammaticality

(7) *Er wehrt sich gegen ø Forderung nach Stilllegung einer Verbrennungsanlage.
he defends REFL against demand for closedown an incineration plant
‘He defies the demand for closing an incineration plant.’

(8) Auch Philippe Egli besteht auf einer eigenen Handschrift –
Also Philippe Egli insists on a own signature –

a) [PP unter der Voraussetzung/der Prämisse des Einverständnisses des Ensembles].
under the prerequisite/ the premise of the acceptance of the ensemble.
b) [PNC unter Voraussetzung des Einverständnisses des Ensembles].
under prerequisite of the acceptance of the ensemble.
c) *[PNC unter Prämisse des Einverständnisses des Ensembles].
under premise of the acceptance of the ensemble.

‘Philippe Egli insists on his own style as well, provided that the ensemble accepts.’
Other ways out: Properties of P and N in English

- Stvan (1998): N-type selection: PNCs are licensed through nouns of a given semantic type.
  - Such nouns may be defective in that they appear without a determiner in other contexts (10) (Baldwin et al. 2006).
  - The nouns induce pragmatic extensions (11).

  (9) on campus, from school, at school, in jail, from jail, ...
  (10) School is over.
  (11) Mary’s husband is in prison.

- Baldwin et al. (2006): P-type selection: PNCs are licensed through selection restrictions imposed on the noun.
  - by+MEANS, on+MEDIUM (we have to assume that a selection restriction leads to an inherent disambiguation of the preposition)

  (12) by train, by plane, by bus, by pogo stick, by hydro-foil, ...
  (13) on disc, on CD, on DVD, on tape, on stick, on memory card, ...
But in German ...

- The properties of N type PNCs described in Stvan (1998) are typical for German Verschmelzungsformen (preposition-determiner-amalgamations).

  (14) Er ist im Gefängnis./Er ist in einem Gefängnis.

- (Why not even consider that the cases in English are covert Verschmelzungsformen!)

- The contrast in (14) cannot typically observed with PNC/PP pairs (already observed in Grimm 1987)

  (15) Milosevic unterschrieb auch unter ø/einer/der Androhung von Nato-Bombardementen nicht. Milosevic signed even under threat/the threat/a threat of bomb attack-by-NATO not 'Milosevic did not even sign under pain of NATO bomb attacks.'

  (16) Ursprünglich war der Artikel als Verbot der Beleidigung jeder Religion eingeführt originally was the paragraph as prohibition of-the libel of-any religion introduced worden, der Diktator engte ihn jedoch nach 1980 auf die Beleidigung des Islam ein, PASS, the dictator narrowed it however after 1980 to the libel of-the Islam PART, und er verschärfte ihn durch ø/eine/die Androhung der Todesstrafe. and he tightened it through ø/a/the threat of-the death penalty.

  ‘The paragraph had been initially introduced as a prohibition against insulting religions in general; however, after 1980 the dictator narrowed it to libelling the Islam and tightened it by putting it under pain of death penalty.'
But in German ...

- The productive combination referred to in Baldwin et al. (2006) for P type selection in PNCs is realized through $[_{PP} P [_{NP} \text{Art}_{def...} +N ...]]$ in German.

(17) Fahren wir mit (*dem) Bus/Auto/Skateboard/Lenkdrachen?
Drive we with the bus/car/skateboard/paraglider
‘Do we take the bus/car/skateboard/paraglider?’

(18) Fahren wir per *(dem) Bus/per *(der) Bahn.
Drive we by the bus by the train
In between summary

- The construction is regular.
- Speakers, however, are reluctant to coin new combinations or to judge combinations in isolation.
- We need a formal model identifying the licensing conditions for PNCs.
- We cannot rely on introspective judgments.
Annotation Mining

- Annotation Mining (Chiarcos et al. 2008) suggests itself as an alternative approach.
  - Ideally, large datasets are annotated automatically (categorial, morphological, syntactic, semantic, contextually)
  - Features are extracted from the data and used for abstraction.
  - The analysis is based on the level of extracted and abstracted features.
  - Focus on techniques from inferential statistics, and machine learning.
  - The results will be re-interpreted as categorial properties of the data.


- Implemented as XML standoff format (strict separation of data and annotation, using MMAX2 (Müller and Strube 2006) for inspection and manual annotation.)
**Example**

<table>
<thead>
<tr>
<th>lexical</th>
<th>Er</th>
<th>wehrt</th>
<th>sich</th>
<th>gegen</th>
</tr>
</thead>
<tbody>
<tr>
<td>POS</td>
<td>PPER</td>
<td>VVFIN</td>
<td>PRF</td>
<td>APPR</td>
</tr>
<tr>
<td>morphology</td>
<td>subst.3.nom.sg.masc</td>
<td>full.3.sg.pres.ind</td>
<td>refl.subst.3.acc.sg.*6</td>
<td>acc</td>
</tr>
<tr>
<td>countability</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>prep. meaning</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>noun meaning</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>syntactic</td>
<td>chunks</td>
<td>NC</td>
<td>VC</td>
<td>PC</td>
</tr>
<tr>
<td>relational</td>
<td>ext-depend.</td>
<td>int-depend.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dependency</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>global</td>
<td>hit</td>
<td>title: no</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Regression Forest Tagger (Schmid and Laws 2008)
SMOR (Schmid 2004)
Tree Tagger (Schmid 1995)
Count (StadtBeld 2010)
Prep.Meanings (Müller et al. 2010)

- Nanogrammar: Each hit is currently modelled by 92 features.

GermaNet (Lemnitzer and Kunze 2007)
HaGenLex (Hartrumpf et al. 2003)
Preposition Sense Annotation: which P?

- As a basis for formal annotation, we had to develop an inventory (in fact: a tagset) for preposition senses, as described in Müller et al. 2010.

- Currently, we consider the following prepositions

  \[(19) \text{ an, auf, bei, binnen, dank, durch, für, gegen, gemäß, hinter, in, mit, mittels, nach, neben, ohne, seit, über, um, unter, vor, während, wegen} \]

  - no postpositions, no secondary prepositions
  - no prepositions without regular NP complement
    - \textit{bis} is excluded, as it does not govern NPs
    - \textit{zwischen} is excluded, as it either governs two NPs or an NP obligatorily marked as plural.
  - no prepositions without case selection
The feature space for preposition senses

- Senses with subtypes (5)
  
  Spatial (= Local and Directional), Temporal, Causal, Modal, Presence

- Senses without subtypes (22)
  
  State, Comitative, Reduction/Extension, Participation, Subordination, Affiliation, Correlation/Interaction, Transgression, Order, Centre of Reference, Theme, Agent, Recipient, Substitute, Exchange, Comparison, Restrictive, Copulative, Adversative, Distributive, Realisation, Statement/Opinion

(cf. Müller et al. 2010)
Prepositions, polysemy, frequency

ohne is the only preposition that occurs more often with count nouns in PNCs than in PPs
Classification by logistic regression

- Analyzing the prepositions *ohne* and *unter*

<table>
<thead>
<tr>
<th>Preposition</th>
<th>Interpretations</th>
<th>PNCs</th>
<th>PPs</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>ohne</td>
<td>6</td>
<td>3159</td>
<td>591</td>
<td>3750</td>
</tr>
<tr>
<td>unter</td>
<td>10</td>
<td>857</td>
<td>4334</td>
<td>5181</td>
</tr>
</tbody>
</table>

- The problem of determiner omission in PNCs is rephrased as a problem of logistic regression: given the classes *with determiner* and *without determiner*, and the annotation classes as described, which annotation shifts odds in favor of determiner omission/determiner realization?

- Factors that *favour determiner omission* will receive *negative coefficients*.
- Factors that *favour determiner realization* will receive *positive coefficients*.

- The analyses have been carried out using R, and Harrell's (2001) DESIGN library in particular.
Logistic Regression Model for *ohne*

- `lrm()` from Harrell's (2001) **DESIGN** library is the function of choice.

Logistic Regression Model

\[
lrm(formula = \text{Det} \sim \text{UNG} + \text{ADJA} + \text{CAUSAL} + \text{COMITATIVE} + \text{PARTICIPATION} + \text{PRESENCE} + \text{DEP}_S + \text{DEP}_NP + \text{DEP}_PP + \text{GN}\_\text{RELATION} + \text{GN}\_\text{ATTRIBUTE} + \text{GN}\_\text{EVENT} + \text{GN}\_\text{ARTEFACT}, \\
data = \text{ohne.ung.data}, \ x = \text{T}, \ y = \text{T}, \ \text{penalty} = 0.8)
\]

Frequencies of Responses

<table>
<thead>
<tr>
<th></th>
<th>no</th>
<th>yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>3159</td>
<td>591</td>
<td></td>
</tr>
</tbody>
</table>
Logistic Regression Model for *ohne*

<table>
<thead>
<tr>
<th>Obs</th>
<th>Max Deriv</th>
<th>Model L.R.</th>
<th>d.f.</th>
<th>P</th>
<th>C</th>
<th>Dxy</th>
<th>R2</th>
<th>Brier</th>
</tr>
</thead>
<tbody>
<tr>
<td>3750</td>
<td>7e-06</td>
<td>1063.48</td>
<td>12.97</td>
<td>0</td>
<td>0.876</td>
<td>0.753</td>
<td>0.424</td>
<td>0.089</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coef</th>
<th>S.E.</th>
<th>Wald</th>
<th>Z</th>
<th>P</th>
<th>Penalty</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-2.4024</td>
<td>0.1109</td>
<td>-21.66</td>
<td>0.0000</td>
<td>0.00000</td>
<td>0.00000</td>
</tr>
<tr>
<td>UNG</td>
<td>-1.3579</td>
<td>0.1870</td>
<td>-7.26</td>
<td>0.0000</td>
<td>0.31599</td>
<td>0.31599</td>
</tr>
<tr>
<td>ADJA</td>
<td>1.1360</td>
<td>0.1188</td>
<td>9.57</td>
<td>0.0000</td>
<td>0.38505</td>
<td>0.38505</td>
</tr>
<tr>
<td>CAUSAL</td>
<td>1.2063</td>
<td>0.1302</td>
<td>9.26</td>
<td>0.0000</td>
<td>0.32459</td>
<td>0.32459</td>
</tr>
<tr>
<td>COMITATIVE</td>
<td>2.2821</td>
<td>0.5201</td>
<td>4.39</td>
<td>0.0000</td>
<td>0.06351</td>
<td>0.06351</td>
</tr>
<tr>
<td>PARTICIPATION</td>
<td>3.4027</td>
<td>0.4895</td>
<td>6.95</td>
<td>0.0000</td>
<td>0.07423</td>
<td>0.07423</td>
</tr>
<tr>
<td>PRESENCE</td>
<td>-0.7780</td>
<td>0.1463</td>
<td>-5.32</td>
<td>0.0000</td>
<td>0.43281</td>
<td>0.43281</td>
</tr>
<tr>
<td>DEP_S</td>
<td>5.0797</td>
<td>1.0542</td>
<td>4.82</td>
<td>0.0000</td>
<td>0.05258</td>
<td>0.05258</td>
</tr>
<tr>
<td>DEP_NP</td>
<td>2.9752</td>
<td>0.1718</td>
<td>17.32</td>
<td>0.0000</td>
<td>0.24268</td>
<td>0.24268</td>
</tr>
<tr>
<td>DEP_PP</td>
<td>2.1978</td>
<td>0.1487</td>
<td>14.78</td>
<td>0.0000</td>
<td>0.25686</td>
<td>0.25686</td>
</tr>
<tr>
<td>GN_RELATION</td>
<td>-1.0292</td>
<td>0.4072</td>
<td>-2.53</td>
<td>0.0115</td>
<td>0.13542</td>
<td>0.13542</td>
</tr>
<tr>
<td>GN_ATTRIBUTE</td>
<td>-1.3528</td>
<td>0.3038</td>
<td>-4.45</td>
<td>0.0000</td>
<td>0.22683</td>
<td>0.22683</td>
</tr>
<tr>
<td>GN_EVENT</td>
<td>-0.8431</td>
<td>0.1431</td>
<td>-5.89</td>
<td>0.0000</td>
<td>0.39470</td>
<td>0.39470</td>
</tr>
<tr>
<td>GN_ARTEFACT</td>
<td>-0.4117</td>
<td>0.1564</td>
<td>-2.63</td>
<td>0.0085</td>
<td>0.33654</td>
<td>0.33654</td>
</tr>
</tbody>
</table>
Logistic regression model for *ohne*

- **Causal**: 1,21  **Presence**: -0,78
  
  **Comitative**: 2,28  
  **Participation**: 3,40

  - While the first three senses favour the realization of a determiner, the last sense favours the omission of a determiner.
  
  - **Comitative**, **Participation**, and **Presence** are core senses of *ohne*, they are only shared with *mit (with)*.

- **Dep-S**: 5,08  
  **Dep-NP**: 2,97  
  **Dep-PP**: 2,20

  - The syntactic realization of a complement, a sentential complement in particular, shifts odds in favour of determiner realization.
Interpretations

(20) Sämtliche kurdischen Politiker sind davon überzeugt, dass ohne einen Machtwechsel in Bagdad die Kurdenfrage des Iraks nicht zu lösen sei. (causal-conditional, NZZ, 28.01.1993)


(22) Die abschließende Verhandlung der Nevada State Athletic Commission fand ohne den Angeklagten statt. (participation, NZZ, 10.07.1997)

Logistic Regression Model for *ohne*

- **GN-RELATION:** -1,03  
  **GN-ATTRIBUTE:** -1,35  
  **GN-EVENT:** -0,84  
  **GN-ARTIFACT:** -0,41

- Nouns appearing in the respective GermaNet Tops shift odds in **favour of**  
  **determiner omission.**
  
  - *relation*: an abstraction belonging to or characteristic of two entities or parts together (*Kameradschaft, Heiratsantrag*)
  - *attribute*: an abstraction belonging to or characteristic of an entity (*Tonhöhe, Gefühllosigkeit*)
  - *event*: something that happens at a given place and time (*Diffusion, Konjunkturprogramm*)
  - *artefact*: a man-made object taken as a whole (*Klosteranlage, Schutzvorrichtung*)

- If a P-type selection restriction analysis were be maintained, it should at least be reformulated in terms of a multifactorial condition (yet, disjunctions are ugly).
Logistic regression model for *unter*

- \texttt{lrm(formula = Det \sim UNG + ADJA + COMPOUND + GOVERNED +}
  
  \hspace{1cm} SPATIAL + CAUSAL + SUBORDINATION + ORDER +
  
  \hspace{1cm} TRANSGRESSION + DEP_S + DEP_NP + DEP_PP + GN_GROUP +
  
  \hspace{1cm} GN_COMMUNICATION + GN_LOCATION + GN_RELATION + GN_POSSESS +
  
  \hspace{1cm} GN_ATTRIBUTE + GN_ARTEFACT + GN_HUMAN + HL_AD + HL_AS,}

  \texttt{data = unter.ung.data, x = T, y = T, penalty = 0.35)}

Frequencies of Responses

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>no</td>
<td>847</td>
</tr>
<tr>
<td>yes</td>
<td>4334</td>
</tr>
</tbody>
</table>
Logistic regression model for *unter*

<table>
<thead>
<tr>
<th>Obs</th>
<th>Max Deriv</th>
<th>Model</th>
<th>L.R.</th>
<th>d.f.</th>
<th>P</th>
<th>C</th>
<th>Dxy</th>
<th>R²</th>
<th>Brier</th>
</tr>
</thead>
<tbody>
<tr>
<td>5181</td>
<td>7e-05</td>
<td>2245.64</td>
<td>21.85</td>
<td>0</td>
<td>0.937</td>
<td>0.874</td>
<td>0.596</td>
<td>0.068</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coef</th>
<th>S.E.</th>
<th>Wald</th>
<th>Z</th>
<th>P</th>
<th>Penalty</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.4379</td>
<td>0.1657</td>
<td>-2.64</td>
<td>0.0082</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>UNG</td>
<td>-0.8346</td>
<td>0.2259</td>
<td>-3.70</td>
<td>0.0002</td>
<td>0.12186</td>
<td>0.23725</td>
</tr>
<tr>
<td>ADJA</td>
<td>-1.0177</td>
<td>0.1432</td>
<td>-7.11</td>
<td>0.0000</td>
<td>0.23725</td>
<td>0.23725</td>
</tr>
<tr>
<td>COMPOUND</td>
<td>2.1719</td>
<td>0.2538</td>
<td>8.56</td>
<td>0.0000</td>
<td>0.23725</td>
<td>0.23725</td>
</tr>
<tr>
<td>GOVERNED</td>
<td>1.9894</td>
<td>0.3017</td>
<td>6.59</td>
<td>0.0000</td>
<td>0.23725</td>
<td>0.23725</td>
</tr>
<tr>
<td>SPATIAL</td>
<td>2.3237</td>
<td>0.2044</td>
<td>11.37</td>
<td>0.0000</td>
<td>0.21627</td>
<td>0.21627</td>
</tr>
<tr>
<td>CAUSAL</td>
<td>1.3047</td>
<td>0.2272</td>
<td>5.74</td>
<td>0.0000</td>
<td>0.12237</td>
<td>0.12237</td>
</tr>
<tr>
<td>SUBORDINATION</td>
<td>3.0529</td>
<td>0.2559</td>
<td>11.93</td>
<td>0.0000</td>
<td>0.23253</td>
<td>0.23253</td>
</tr>
<tr>
<td>ORDER</td>
<td>3.4228</td>
<td>0.1861</td>
<td>18.40</td>
<td>0.0000</td>
<td>0.28756</td>
<td>0.28756</td>
</tr>
<tr>
<td>TRANSGRESSION</td>
<td>4.4186</td>
<td>0.3677</td>
<td>12.02</td>
<td>0.0000</td>
<td>0.20323</td>
<td>0.20323</td>
</tr>
<tr>
<td>DEP_S</td>
<td>8.4717</td>
<td>4.0734</td>
<td>2.08</td>
<td>0.0375</td>
<td>0.07979</td>
<td>0.07979</td>
</tr>
<tr>
<td>DEP_NP</td>
<td>0.8551</td>
<td>0.1436</td>
<td>5.95</td>
<td>0.0000</td>
<td>0.26362</td>
<td>0.26362</td>
</tr>
<tr>
<td>DEP_PP</td>
<td>0.3043</td>
<td>0.2170</td>
<td>1.40</td>
<td>0.1609</td>
<td>0.20944</td>
<td>0.20944</td>
</tr>
</tbody>
</table>
## Logistic Regression Model for *unter*

<table>
<thead>
<tr>
<th></th>
<th>Obs</th>
<th>Max Deriv</th>
<th>Model L.R.</th>
<th>d.f.</th>
<th>P</th>
<th>C</th>
<th>Dxy</th>
<th>R2</th>
<th>Brier</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5181</td>
<td>7e-05</td>
<td>2245.64</td>
<td>21.85</td>
<td>0</td>
<td>0.937</td>
<td>0.874</td>
<td>0.596</td>
<td>0.068</td>
</tr>
</tbody>
</table>

- **GN_GROUP**
  - Coefficient: 0.5241
  - Standard Error: 0.2563
  - t-value: 2.04
  - P-value: 0.0409
  - C: 0.16679

- **GN_COMMUNICATION**
  - Coefficient: -0.9149
  - Standard Error: 0.1443
  - t-value: -6.34
  - P-value: 0.0000
  - C: 0.29059

- **GN_LOCATION**
  - Coefficient: 2.2704
  - Standard Error: 0.6208
  - t-value: 3.66
  - P-value: 0.0003
  - C: 0.07979

- **GN_RELATION**
  - Coefficient: -2.1161
  - Standard Error: 0.6022
  - t-value: -3.51
  - P-value: 0.0004
  - C: 0.04181

- **GN_POSSESS**
  - Coefficient: -0.8482
  - Standard Error: 0.3665
  - t-value: -2.31
  - P-value: 0.0206
  - C: 0.11316

- **GN_ATTRIBUTE**
  - Coefficient: -2.2847
  - Standard Error: 0.2741
  - t-value: -8.33
  - P-value: 0.0000
  - C: 0.28393

- **GN_ARTEFACT**
  - Coefficient: 0.4169
  - Standard Error: 0.1601
  - t-value: 2.60
  - P-value: 0.0092
  - C: 0.28393

- **GN_HUMAN**
  - Coefficient: 1.8870
  - Standard Error: 0.4999
  - t-value: 3.77
  - P-value: 0.0002
  - C: 0.17855

- **HL_AD**
  - Coefficient: -1.0253
  - Standard Error: 0.1888
  - t-value: -5.43
  - P-value: 0.0000
  - C: 0.16184

- **HL_AS**
  - Coefficient: -1.4214
  - Standard Error: 0.3804
  - t-value: -3.74
  - P-value: 0.0002
  - C: 0.10979

### HL-AD: Lauf, Diebstahl, Veränderung

### HL-AS: Umstand, Notlage, Auslastung
Comparison of the models

- Both models show high positive coefficients for the feature DEP-S.
- Both models show positive coefficients for features representing syntactic arguments.
- The coefficients for prenominal modification (ADJA) differ in both models.
  - Presence of adjectival modification shifts odds in favour of determiner omission with unter.
  - Presence of adjectival modification shifts odds in favour of determiner realisation with ohne.
- Both models show negative coefficients for the feature UNG.
  - However, the model for ohne performs better in bootstrap validation, if this factor is not considered.
## Brief assessment of the models: bootstrap validation

<table>
<thead>
<tr>
<th>ohne</th>
<th>index</th>
<th>training</th>
<th>test</th>
<th>optimism</th>
<th>corr. index</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_{xy}$</td>
<td>0.7526</td>
<td>0.7570</td>
<td>0.7500</td>
<td>0.0070</td>
<td>0.7456</td>
</tr>
<tr>
<td>$E_{max}$</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0096</td>
<td>0.0096</td>
<td>0.0096</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>unter</th>
<th>index</th>
<th>training</th>
<th>test</th>
<th>optimism</th>
<th>corr. index</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_{xy}$</td>
<td>0.8736</td>
<td>0.8744</td>
<td>0.8692</td>
<td>0.0052</td>
<td>0.8684</td>
</tr>
<tr>
<td>$E_{max}$</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0055</td>
<td>0.0055</td>
<td>0.0055</td>
</tr>
</tbody>
</table>
Beyond interpretation?

- At first glance, article omission with *ohne* is constrained by interpretation and the syntactic realization of noun complements. However ...

- The interpretations *comitative*, *participation*, and *presence* belong to the core meaning of *ohne* (and *mit*), in the sense that no other preposition shares this sense.
  - Observation: *ohne* will only assume the meaning *presence* if it modifies nouns.
  - Observation: other interpretations of *ohne*, including *comitative* and *participation* occur in verbal modification only (strictly speaking: in modification contexts where the modified element denotes an event).
Beyond interpretation?

- Verbal modification

(24) *Die Hauptverhandlung wird am 8. Februar notfalls auch ohne den Angeklagten fortgesetzt.*

The main hearing will at 08.02. in-case-of-need also without the defendant continued.

(NZZ, 28.01.1993)

(25) *Der Viertelmeilen-Sprint, einer der anfänglich vermuteten Höhepunkte, entwickelte sich ohne den großen Abwesenden von Zürich, Michael Johnson, zum erwarteten US-Duell zwischen Antonio Pettigrew und Jerome Young.*

The quarter-mile sprint, one of the incipiently assumed highlights developed REFL without the big absentee of Zürich, Michael Johnson to-the expected US duel between Antonio Pettigrew and Jerome Young.

(NZZ, 12.08.1999)
Beyond interpretation

- Nominal event modification

(26) Nur akademische Miles-Davis-Forscher dürfen von diesem only academic Miles-Davis-researchers will of this detaillierten Einblick in die Werkstatt profitieren und detailed insight in the workshop profit and stundenlang abgebrochene Versuche oder gar Durchläufe ohne hour-after-hour discontinued attempts or even passages without den Solisten mitverfolgen. the soloist trace (NZZ, 02.10.1996)

(27) Denn eine Siegerehrung ohne die Amerikanerin ist letztlich doch wohl because an award ceremony with the American is ultimately arguably ebenso die Ausnahme von der Regel wie die Durchführung einer Abfahrt as well the exception from the rule just-as the execution of-the downhill in zwei Läufen. in two runs (NZZ, 16.12.1995)
Beyond interpretation

- If we are on the right track, the rule should not be
  - “if interpretation $s_i$ then PNC and if interpretation $s_j$ then PP”,
- but instead
  - if modification of V (e) then PP, and if modification of N (not e) then PNC.
- It should be noted that a similar distribution of core meanings cannot be observed for unter.
  - Perhaps, PNCs require a differentiation of core meanings to become not only productive, but truly regular.
Conclusions

- We have presented logistic regression models for determiner omission with *ohne* and *unter*.

- The presence of syntactic complements of the noun seems to be a common structural denominator.
  - Sentential complements are particularly inclined not to be realized inside PNCs.
  - The reason might be that sentential complements require a referential anchor that can only be provided by the determiner.
  - If this is basically correct, we expect relative clauses not to appear in PNCs (which seems correct).

- While interpretations could be considered as factors for determiner omission, their respective presence or absence in PNCs might reflect deeper regularities as well (modification of V, modification of N).

- A true regularity of PNCs, however, seems to require a proper distribution of meanings of the preposition, be it dependent on structural or interpretational factors.