

# *Syntax natürlicher Sprachen*

Tutorium 08:

Grammatische Merkmale

Shuyan Liu

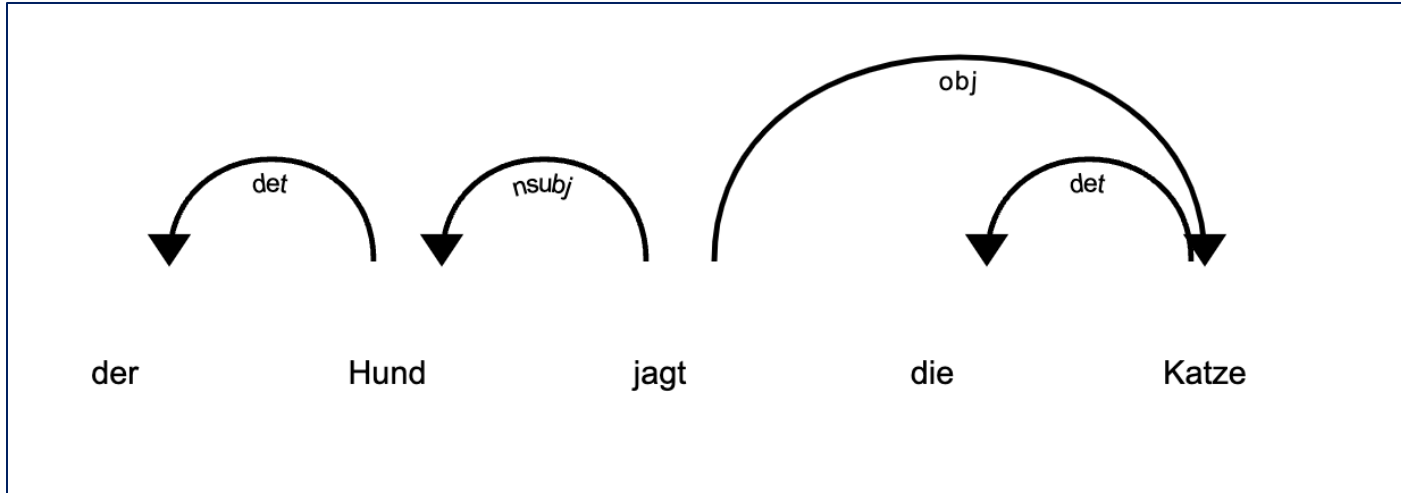
13.12.2024

Einige Beispiele kommt aus der Vorlesungsfolien, Aufgaben sowie Übungen.

Die Hauptteile der Slides dieser Woche stammen von Sarah Anna Uffelmann aus Wintersemester 2023/24 und wurden bearbeitet. Verwendung mit Dank.

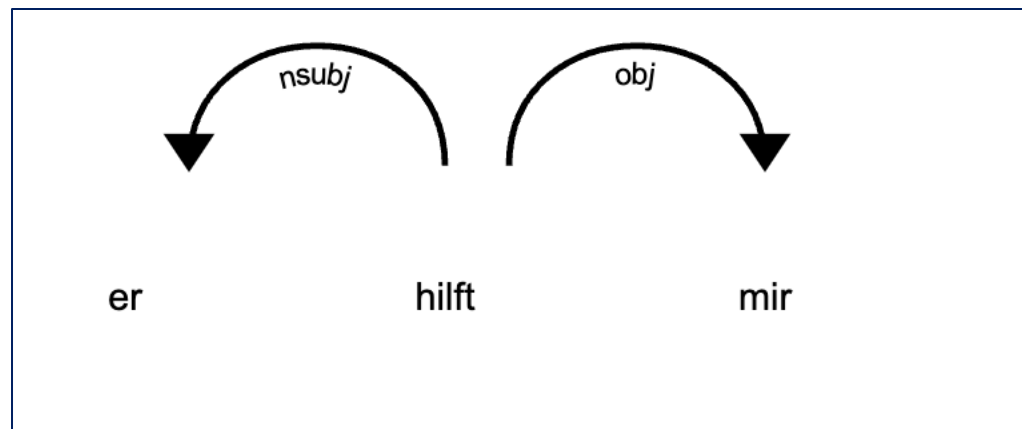
**Nachtrag zu Dependenz:  
iobj, obj, cop, aux**

# obj



## Akkusativobjekt

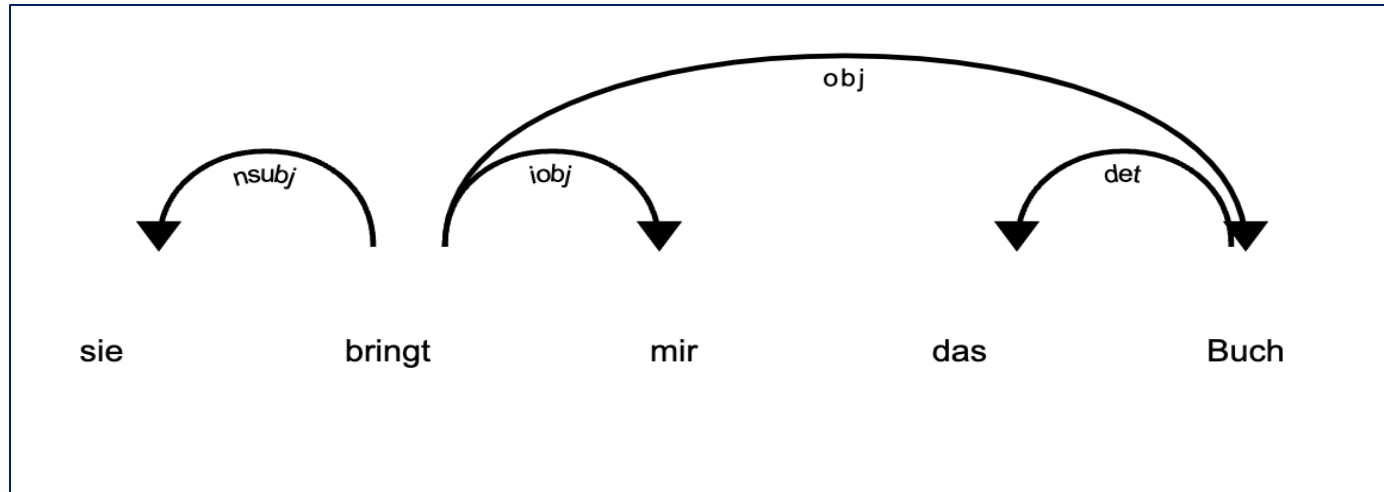
```
sent_nr = ""  
1 der 2 det  
2 Hund 3 nsubj  
3 jagt 0 ROOT  
4 die 5 det  
5 Katze 3 obj  
""
```



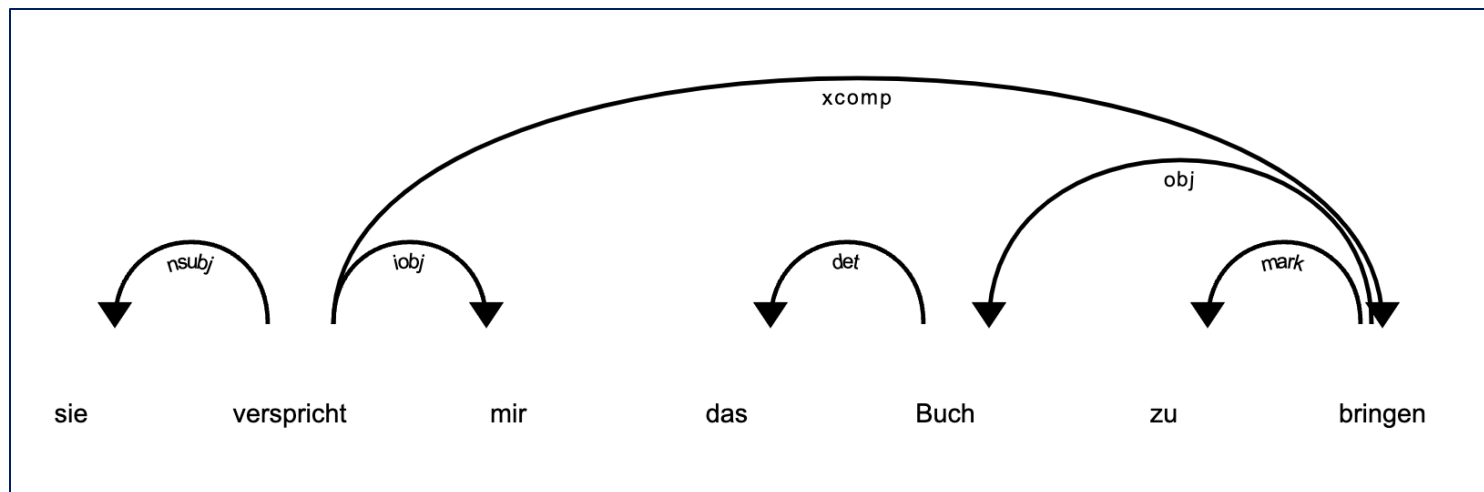
## Dativobjekt

```
sent_nr = ""  
1 er 2 nsubj  
2 hilft 0 ROOT  
3 mir 2 obj  
""
```

# obj und iobj



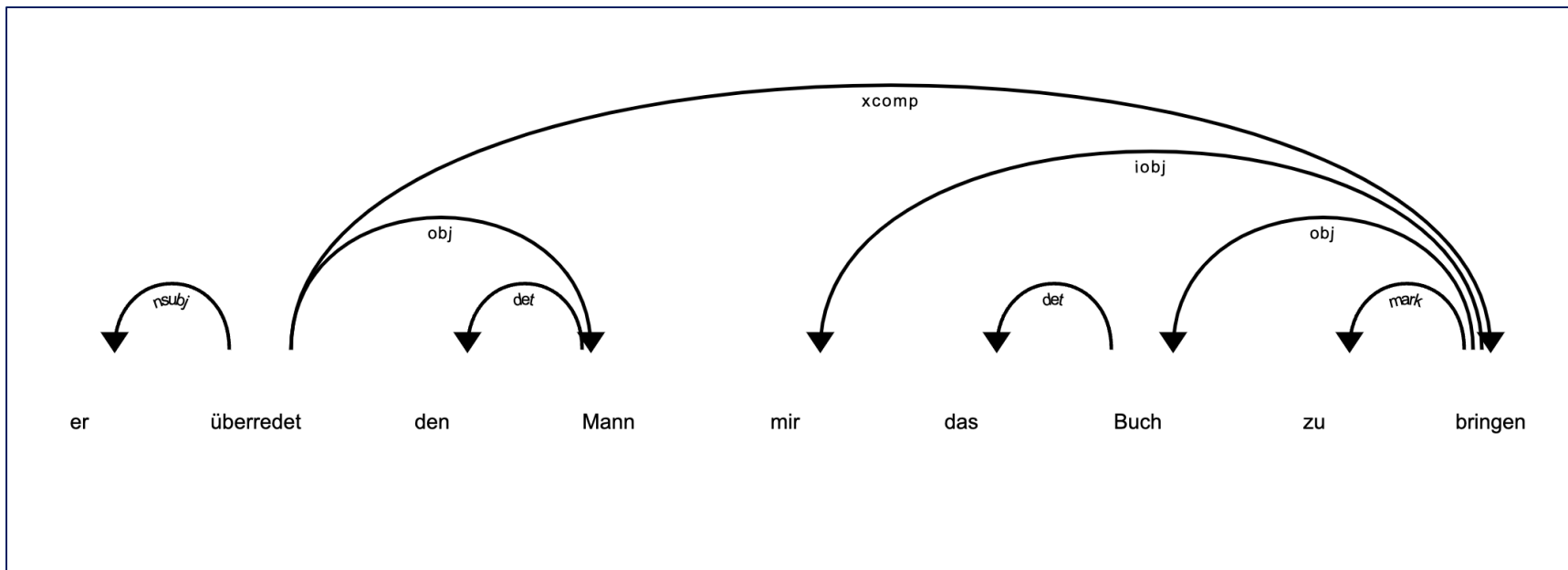
```
sent_nr = ""  
1 sie 2 nsubj  
2 bringt 0 ROOT  
3 mir 2 iobj  
4 das 5 det  
5 Buch 2 obj  
""
```



```
sent_nr = ""  
1 sie 2 nsubj  
2 verspricht 0 ROOT  
3 mir 2 iobj  
4 das 5 det  
5 Buch 7 obj  
6 zu 7 mark  
7 bringen 2 xcomp  
""
```

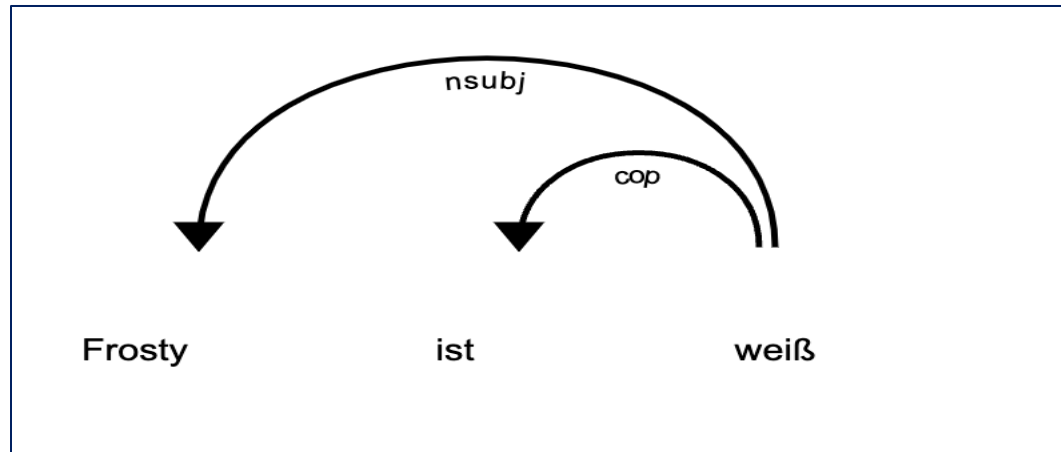
# obj und iobj

Es kann aber auch obj und xcomp (bzw. ccomp) in einem Satz geben!  
(Hier zusätzlich mit obj und iobj auch im eingebetteten Satz.)



```
sent_nr = ""  
1 er 2 nsubj  
2 überredet 0 ROOT  
3 den 4 det  
4 Mann 2 obj  
5 mir 9 iobj  
6 das 7 det  
7 Buch 9 obj  
8 zu 9 mark  
9 bringen 2 xcomp  
""
```

# cop

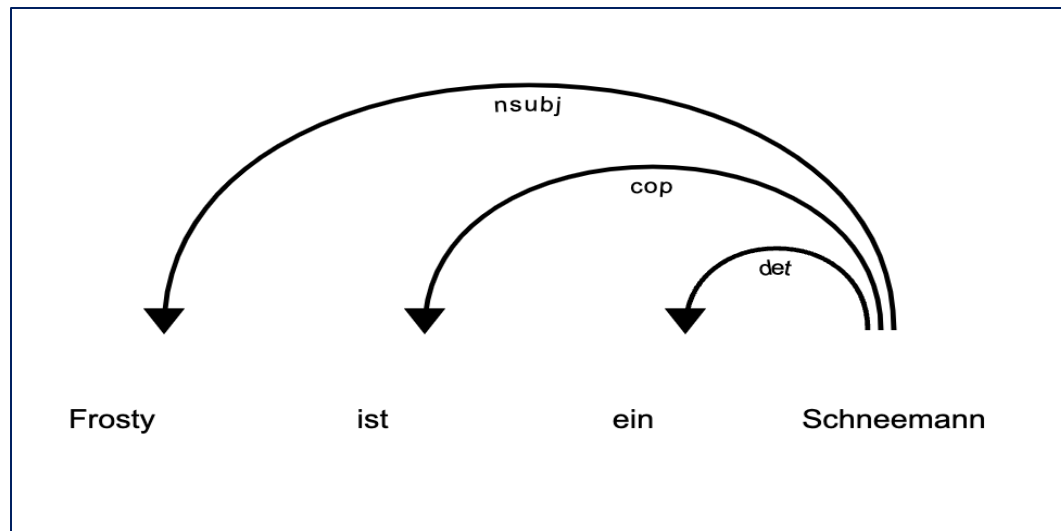


```
sent_nr = ""  
1 Frosty 3 nsubj  
2 ist 3 cop  
3 weiß 0 ROOT
```

""

**Prädikative  
Zuschreibung:**

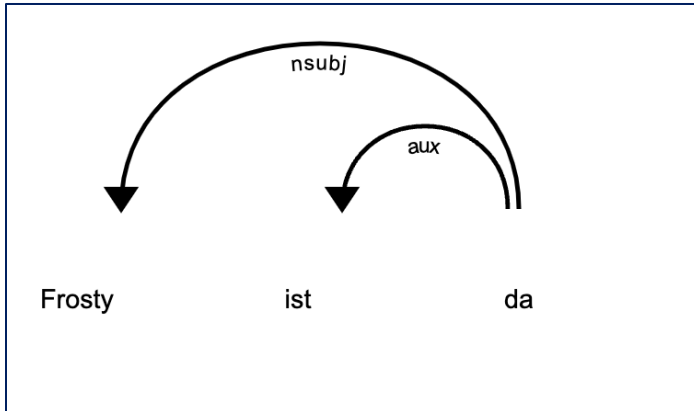
**X ist Y**



```
sent_nr = ""  
1 Frosty 4 nsubj  
2 ist 4 cop  
3 ein 4 det  
4 Schneemann 0 ROOT
```

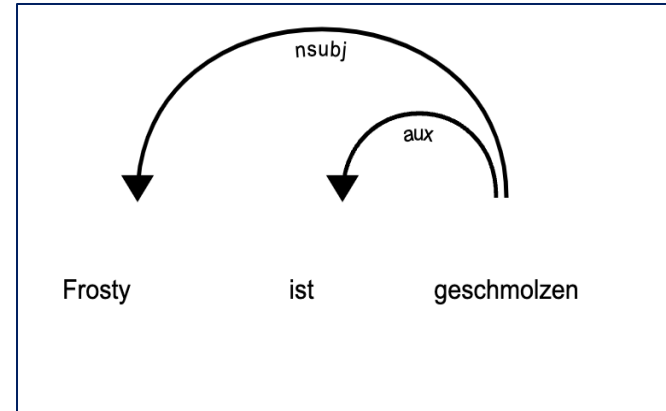
""

# aux



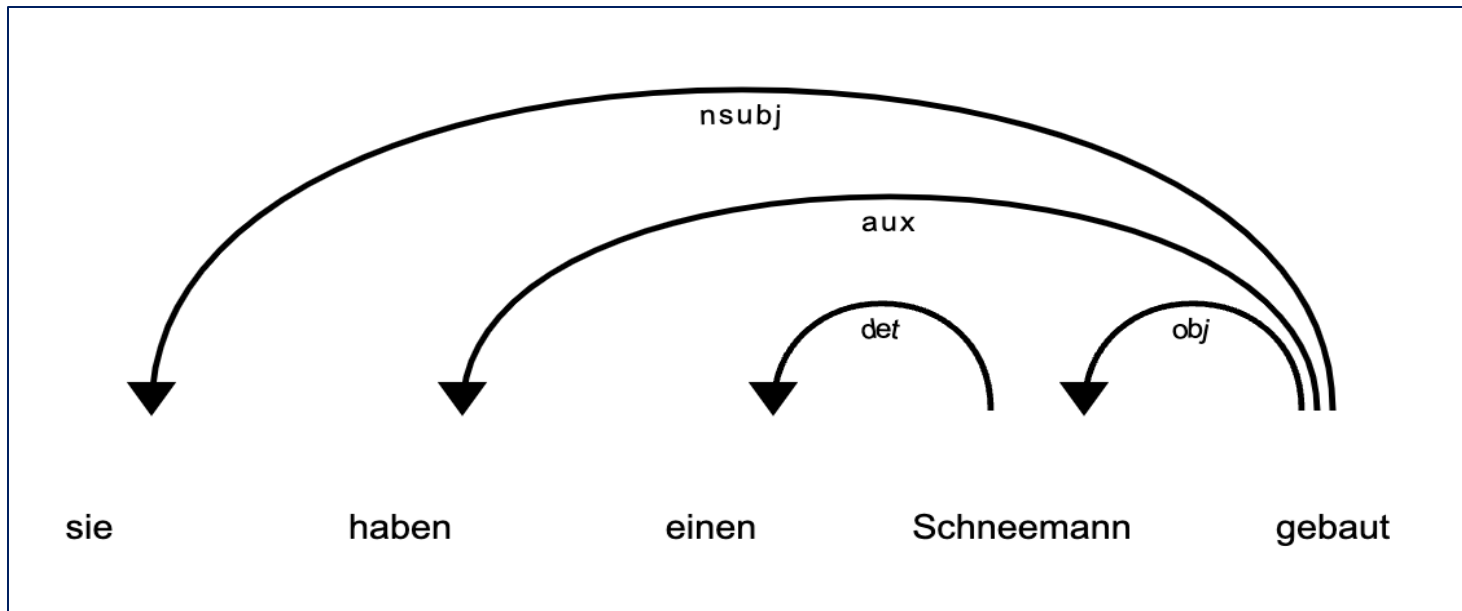
```
sent_nr = ""  
1 Frosty 3 nsubj  
2 ist 3 aux  
3 da 3 ROOT  
""
```

Verb: dasein



```
sent_nr = ""  
1 Frosty 3 nsubj  
2 ist 3 aux  
3 geschmolzen 0 ROOT  
""
```

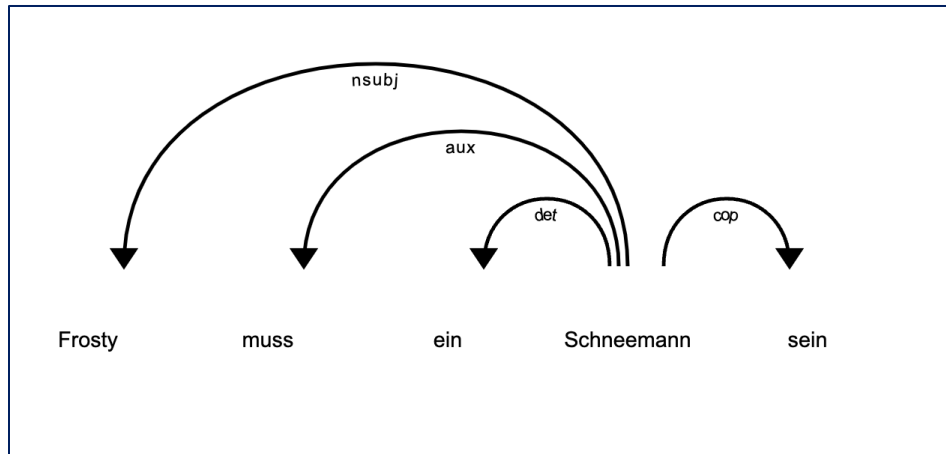
Perfekt von ‚schmelzen‘



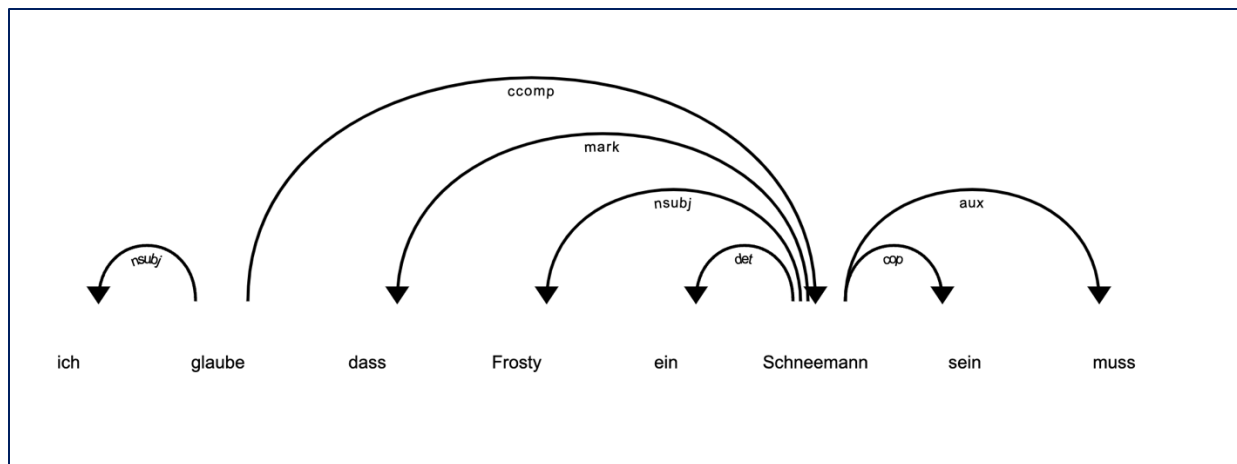
```
sent_nr = ""  
1 sie 5 nsubj  
2 haben 5 aux  
3 einen 4 det  
4 Schneemann 5 obj  
5 gebaut 0 ROOT  
""
```

Perfekt von ‚bauen‘

# cop und aux in einem Satz



```
sent_nr = ""  
1 Frosty 4 nsubj  
2 muss 4 aux  
3 ein 4 det  
4 Schneemann 0 ROOT  
5 sein 4 cop  
""
```



```
sent_nr = ""  
1 ich 2 nsubj  
2 glaube 0 ROOT  
3 dass 6 mark  
4 Frosty 6 nsubj  
5 ein 6 det  
6 Schneemann 2 ccomp  
7 sein 6 cop  
8 muss 6 aux  
""
```



# Grammatische Merkmale

# Problem: Übergenerierung

- X-Bar Struktur kann strukturelle Probleme im Satz lösen
  - Problematische Satz: \*der die das der Mann läuft.
  - Problematische Regeln:  $NP \rightarrow DET\ NP$ ;  $NP \rightarrow N$
  - Lösung:  $NP \rightarrow DET\ NOM$ ;  $NOM \rightarrow ADJ\ NOM$ ;  $NOM \rightarrow N$
- **Feature-Based Grammar** Heutiges Thema löst Probleme über Merkmale und Merkmalswerte.

# Problem: Übergenerierung

- Wir haben die Regeln für diese 2 Sätzen:
  - er sieht den Baum an der Uni
  - ich laufe
- Problematische Satz: \*ich laufe ein Baum an der Uni nicht grammatisch korrekt!
- Aber trotzdem die unterstehenden Regeln lassen diesen Satz zu!

S → NP VP

VP → VERBAL

VERBAL → VERBAL PP

VERBAL → V NP | V

PP → P NP

NP → Det N | N

N → „Baum“ | „er“ | „Uni“ | „ich“

Det → „der“ | „den“ | „ein“

V → „sieht“ | „laufe“

P → „an“

Das Programm weiß nicht, was Nominativ ist, was Akkusativ ist, sowie die richtige Konjugation usw.

**Lösung: Feature-Based Grammar**

# Feature-Based Grammar

Wir lösen dieses Problem, indem wir morphosyntaktische Merkmalsstrukturen (Features) modellieren.

-> **Feature-based Grammar = FCFG** (Feature-based context free grammar)

Wir müssen unsere **Regeln** so **einschränken**, dass ungrammatische Sätze nicht mehr abgeleitet werden können. Dabei müssen wir vor allem die folgenden drei zentralen **morphosyntaktischen Constraints** berücksichtigen:

- **Kasusreaktion (Case Government)** z.B. für + Akkusativ; mit + Dativ
- **Kongruenz (Agreement)** Übereinstimmung, z.B. ich bin; er ist; \*er bin; \*ich ist
- **Subkategorisierung** Anzahl und Art der verbalen Argumente

# Morphosyntaktische Constraints

## Kasusreaktion

Bestimmte Wörter (hauptsätzlich Verb, Präposition) erfordern (regieren) einen bestimmten Fall (Kasus).

### **Beispiele:**

- “warten auf“ + Akkusativ
- “folgen“ + Dativ
- „an etwas glauben“ + Akkusativ
- „gedenken“ + Genitiv
- „mit“ + Dativ
- „für“ + Akkusativ

# Morphosyntaktische Constraints

## Kongruenz

Übereinstimmung von Wörtern oder Satzteilen hinsichtlich grammatischer Merkmale

### Beispiele:

„**Ein** schöner grüner **Baum** steht im Wald.“ (Nom. Sg. mask. --- 3. Pers. Sg.)

„**Der** schöne grüne **Baum** ist eine Tanne.“ (Nom. Sg. mask. --- 3. Pers. Sg.)

„Im Wald stehen viele große **Tannen**.“ (Nom. Pl. fem. --- 3. Pers. Pl.)

„Siehst **Du** den dunklen **Wald**?“ (Akk. Sg. mask. --- 2. Pers. Sg.)

-> Übereinstimmung von Subjekt und Verb in Person und Numerus

-> Übereinstimmung von Kasus, Numerus und Genus in NPs

# Morphosyntaktische Constraints

## Subkategorisierung

Art und Anzahl der Argumente (Valenz)

### Beispiele:

- **intransitive Verben**: Subjekt im Nominativ  
wachsen -> „Der Baum wächst.“
- **transitive Verben**: Subjekt im Nominativ + Objekt (meistens im Akk., aber nicht immer)  
beantworten -> „Meine Freundin beantwortet den Brief.“ (Akkusativobjekt)  
gehören -> „Der Hund gehört ihm.“ (Dativobjekt, in UD auch als obj markiert)
- **ditransitive Verben**: Subjekt im Nominativ + Akkusativobjekt + Dativobjekt  
bringen -> „Der Postbote bringt ihr ein Paket.“

# 2 Typen von Flexion

- **Deklination** (Nomen, Pronomen, Adjektiven und Artikeln)
  - Grammatische Kategorien: Kasus, Numerus, Person, Genus, Definitheit
- **Konjugation** (Verben)
  - Grammatische Kategorien: Person, Numerus, Tempus, Modus, Genus Verbi



# Flexion: Deklination

Nomen: **Hauptwort** bzw. Kern des Nominalkomplexes

- Kasus
  - Nominativ, Genitiv, Dativ, Akkusativ
- Numerus
  - Singular, Plural
- Genus
  - Maskulin, Feminin, Neutrum
- Person
  - 1, 2, 3
- Definitheit
  - Bestimmter Artikel (der, die, das, dem, ...), Unbestimmter Artikel (ein, einer, eines, ...)

# Flexion: Deklination, Beispiele

- Das Nomen Hund dekliniert mit dem bestimmten Artikel
  - der Hund, des Hundes, dem Hund, den Hund
- Das Adjektiv klein dekliniert je nach Fall, Numerus und Genus.
  - der kleine Hund, des kleinen Hundes, dem kleinen Hund, den kleinen Hund
- Deklination eines Pronomens
  - ich; meiner; mir; mich
- Deklination eines Plurals
  - die Hunde, der Hunde, den Hunden, die Hunde

# Flexion: Konjugation

- Person
  - 1, 2, 3
- Numerus
  - Singular, Plural
- Tempus
  - Präsens, Präteritum, Perfekt, Plusquamperfekt, Futur I, Futur II
- Modus
  - Indikativ, Konjunktiv, Imperativ
- Genus Verbi
  - Aktiv, Passiv

# Flexion: Konjugation, Beispiele

- Wort: Machen, Präsens: Person und Numerus ändern sich:
  - Singular: Ich mache, Sie Höflichkeitsform machen, du machst, er/sie/es macht
  - Plural: wir machen, Sie Höflichkeitsform machen, ihr macht, sie machen
- Wort: Machen: Tempus
  - ich mache, ich machte, ich habe gemacht, ich werde machen
- Modus:
  - er macht (Indikativ), er mache (Konjunktiv I), er machte (Konjunktiv II), mach! (Imperativ), machen Sie! (Imperativ), macht (Imperativ)
- Genus Verbi:
  - Er macht die Arbeit →
    - Die Arbeit wird gemacht
    - Die Arbeit ist gemacht worden
    - Die Arbeit ist gemacht

# Grammatische Merkmale

- Jedes Merkmal hat seine Merkmalsausprägung (Wert)

Grammatisches Merkmal	Werte (Merkmalsausprägungen)
Numerus	Singular, Plural
Person	1, 2, 3
Kasus	Nominativ, Genitiv, Akkusativ, Dativ

# Kongurenz (Agreement)

- Übereinstimmung: bestimmte Satzteile stimmen hinsichtlich ihrer Merkmale überein
- Beispiel:
  - Ich laufe
  - Der Baum steht im Wald

# Merkmalsstrukturen

**Merkmalsstrukturen** (= **Feature Structures**) sind Merkmal-Wert-Paare:

$$\text{Merkmalsstruktur} = \begin{bmatrix} \text{MERKMAL1} & \text{WERT1} \\ \text{MERKMAL2} & \text{WERT2} \end{bmatrix}$$

Grammatische Merkmale lassen sich als Merkmalsstruktur darstellen:

$$N \begin{bmatrix} \text{NUM} & \text{SG} \\ \text{GEN} & \text{MASK} \\ \text{CASE} & \text{NOM} \end{bmatrix} \quad \text{oder} \quad \begin{bmatrix} \text{CAT} & N \\ \text{NUM} & \text{SG} \\ \text{GEN} & \text{MASK} \\ \text{CASE} & \text{NOM} \end{bmatrix} \quad \text{CAT} = \text{Category}$$

# Merkmalsstrukturen

Merkmalsstrukturen können selbst Werte einer Merkmalsstruktur sein:

$$\begin{bmatrix} \text{CAT} & N \\ \text{AGR} & \begin{bmatrix} \text{NUM} & SG \\ \text{GEN} & MASK \\ \text{CASE} & ACC \end{bmatrix} \end{bmatrix}$$

Verschiedene Kongruenz-Merkmale können so unter „Agreement“ zusammengefasst werden.

$$\text{Hund} \begin{bmatrix} \text{CAT} & N \\ \text{AGR} & \begin{bmatrix} \text{NUM} & SG \\ \text{GEN} & MASK \end{bmatrix} \end{bmatrix}$$



# Strukturierung Sprachlicher Kategorien

- **Unifikation (Vereinigung)**

- Zwei Merkmalsstrukturen **unifizieren**, wenn sie vereinbar sind.

- **Subsumption (Unterordnung)**

- Eine Merkmalsstruktur **f1** **subsumiert** eine andere Merkmalsstruktur **f2** genau dann, wenn die **in f1 enthaltene Information auch in f2** enthalten ist.

# Unifikation

- Zwei Merkmalsstrukturen unifizieren, wenn sie vereinbar sind.
- Das bedeutet, dass es **keine widersprüchlichen Merkmal-Wert-Paare** in den Merkmalsstrukturen geben darf.

(Voraussetzung, **sonst schlägt die Unifikation fehl**)

- Wenn diese Voraussetzung erfüllt wird, dann ist diese Operation ähnlich wie A U B in Mengenoperationen.

$$\left| \begin{array}{cc} \text{NUM} & \text{SG} \\ \text{GEN} & \text{MASK} \\ \text{CASE} & \text{NOM} \end{array} \right| \sqcup \left| \begin{array}{cc} \text{NUM} & \text{SG} \\ \text{GEN} & \text{MASK} \end{array} \right| = \left| \begin{array}{cc} \text{NUM} & \text{SG} \\ \text{GEN} & \text{MASK} \\ \text{CASE} & \text{NOM} \end{array} \right|$$

# Unifikation: Festlegung Fehlender Informationen

In diesem Beispiel ist die zweite Merkmalsstruktur **unterspezifiziert**, weil kein Merkmal-Wert-Paar für den Kasus angegeben wird:

$$\begin{bmatrix} \text{NUM} & \text{SG} \\ \text{GEN} & \text{MASK} \\ \text{CASE} & \text{NOM} \end{bmatrix} \sqcup \begin{bmatrix} \text{NUM} & \text{SG} \\ \text{GEN} & \text{MASK} \end{bmatrix} = \begin{bmatrix} \text{NUM} & \text{SG} \\ \text{GEN} & \text{MASK} \\ \text{CASE} & \text{NOM} \end{bmatrix}$$

**Beide Merkmalsstrukturen unifizieren, da sie nicht widersprüchlich sind.**

Zum Beispiel könnte die erste Merkmalsstruktur den Artikel „der“ repräsentieren, die zweite das Substantiv „Baum“. „Baum“ kann aber nicht nur Nominativ sondern auch Akkusativ und Dativ sein. Deshalb legen wir in der Merkmalsstruktur den Kasus nicht fest. In dem Beispiel unifiziert „Baum“ mit dem Artikel „der“ zu der NP „der Baum“. Dadurch wird auch der Kasus von „Baum“ (Nominativ) in diesem konkreten Fall festgelegt.

# Subsumption

- Eine Struktur A subsumiert eine andere Struktur B, wenn alle Merkmale von A in B enthalten sind
- Wenn keine spezifischen Merkmale angegeben sind, enthält die Struktur alle möglichen Informationen.

- Z.B.:

A: [CASE: Nominativ] ==

A: [NUM: -, CASE: Nominativ, GEN: -, andere Merkmale (muss man nicht schreiben)]

wenn man anderen unspezifizierten Merkmalen schreiben möchte, kann man einfach wie „GEN: -“ schreiben.

B: [NUM: Singular, CASE: Nominativ, GEN: Maskulin] → A subsumiert B

- D.h. B ist spezifischer als A; A ist allgemeiner als B.

# Unifikation und Subsumption in NLTK

**Methode `unify()`:**

**`f1.unify(f2)`**

-> gibt bei erfolgreicher Unifikation die unifizierte Merkmalsstruktur zurück, ansonsten NONE

**Methode `subsumes()`:**

**`f1.subsumes(f2)`**

-> gibt bei erfolgreicher Subsumption TRUE zurück, ansonsten FALSE

# Unifikation und Subsumption in NLTK

Beispiele:

```
f0 = FeatStruct("[AGR=[NUM=pl]]")  
f1 = FeatStruct("[AGR=[NUM=sg, PERS=3]]")
```

```
#Subsumption (f0  $\sqsubseteq$  f1)  
f0.subsumes(f1)
```

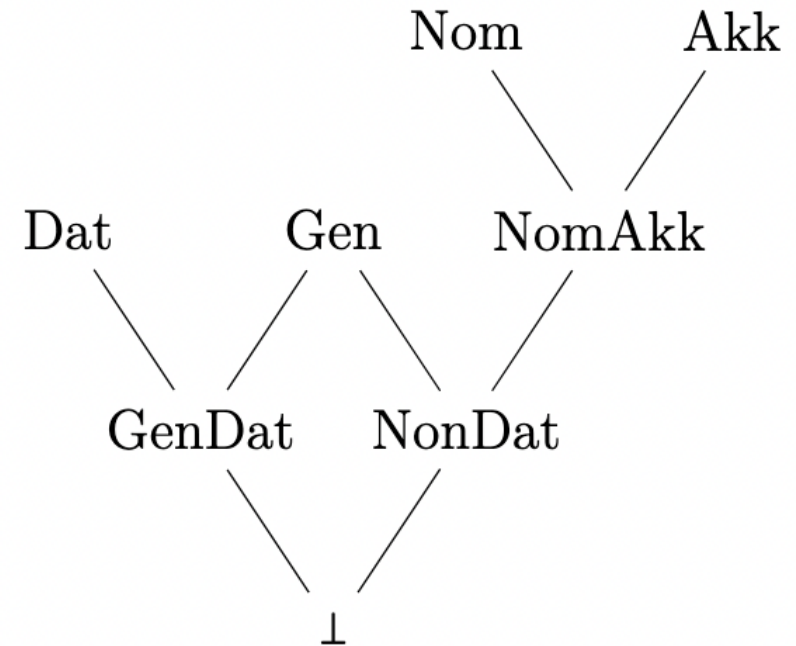
FALSE

```
#Unification (f0  $\sqcup$  f1)  
print(f0.unify(f1))
```

NONE

# Typenhierarchie

- Die Typenhierarchie ist als ein Baumdiagramm dargestellt, das verschiedene Typen und deren Beziehungen zueinander beschreibt.
- **Allgemeine Typen** befinden sich weiter unten in der Hierarchie, während **spezifischere Typen** weiter oben angeordnet sind.
- Allgemeine Typen subsumiert ihre Tochterknoten. Wie z. B. NonDat subsumiert NomAkk und Nom.



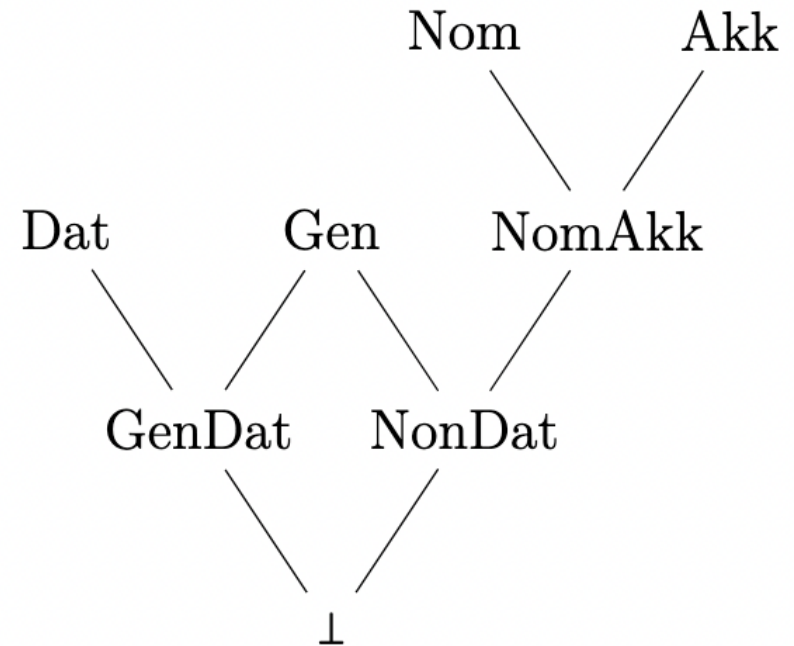
# Typenhierarchie

- $\perp$  (Bottom-Typ)

Am unteren Ende der Hierarchie befindet sich der Typ  $\perp$ , der als der **allgemeinste Typ** gilt.

- Eigenschaften von  $\perp$ :

- Er ist die Basis für alle anderen Typen und enthält keine spezifischen Merkmale.
- Jeder andere Typ der Hierarchie ist eine Spezialisierung von  $\perp$ .
- $\perp$  **subsumiert alle anderen Typen**, da er die allgemeinste Kategorie darstellt.





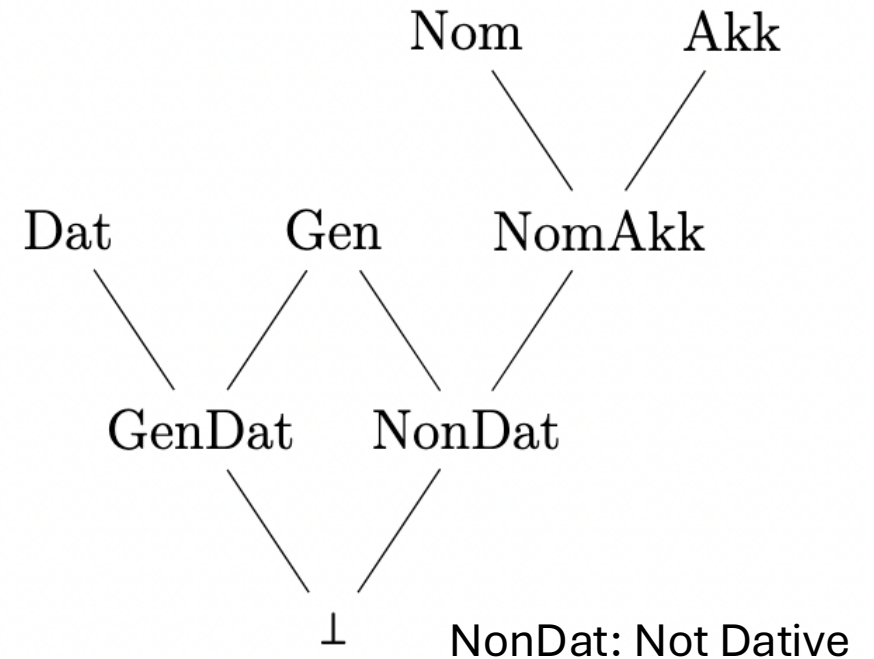
# Typenhierarchie

- **Spezifische Typen:**

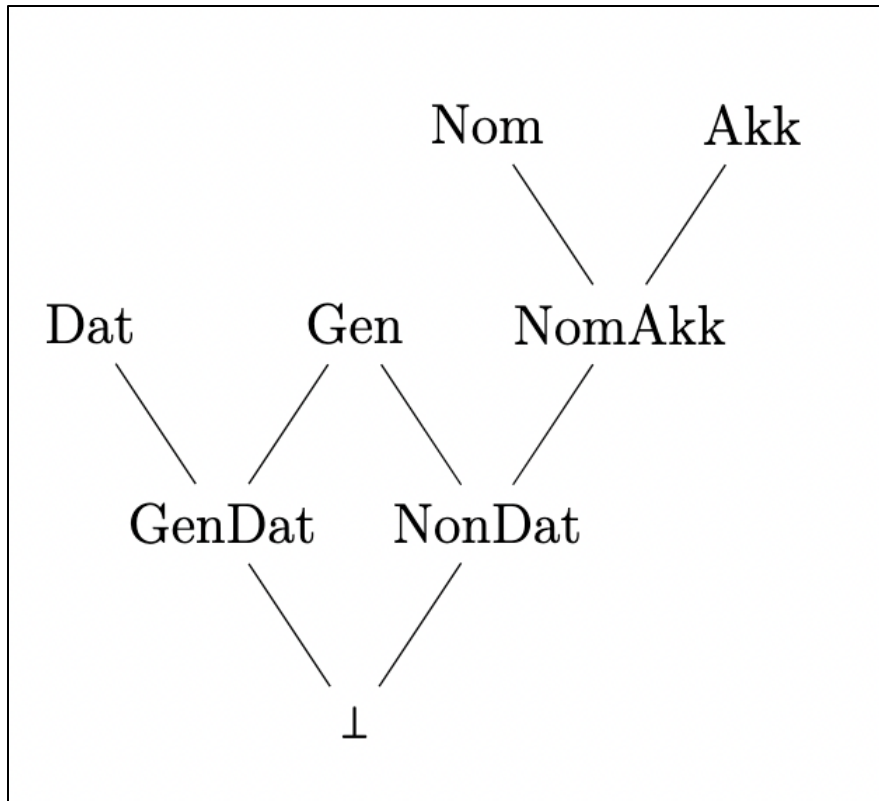
Höher in der Hierarchie sind spezifischere Typen zu finden, die jeweils besondere Merkmale besitzen. In diesem Fall zeigt die Typenhierarchie verschiedene **Kasus-Typen** wie **Nom** (Nominativ), **Akk** (Akkusativ), **Dat** (Dativ) und **Gen** (Genitiv).

- Beispiel:

- **Nom** ist ein Typ, der sich nur auf den Nominativ bezieht (spezifisch).
- **NomAkk** ist ein Typ, der sowohl den Nominativ als auch den Akkusativ umfasst (allgemeiner).



## Beispiel einer Typenhierarchie:



In NLTK definieren wir die Typenhierarchie als Dictionary:

```
case_hierarchy = {  
    "nondat": ["gen", "nomakk"],  
    "gendat": ["gen", "dat"],  
    "nomakk": ["nom", "akk"],  
    "nom": [],  
    "gen": [],  
    "dat": [],  
    "akk": []  
}
```

**Übung**

# Unifikation

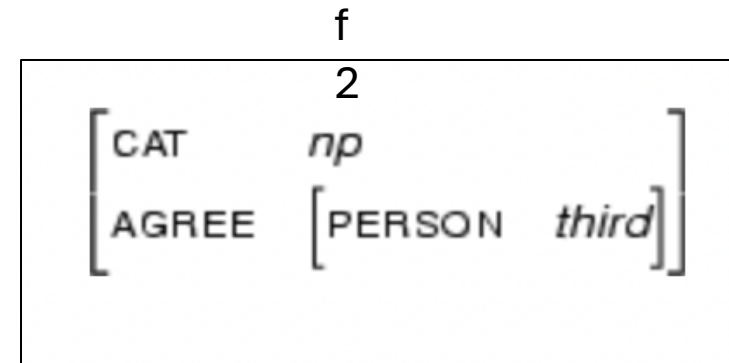
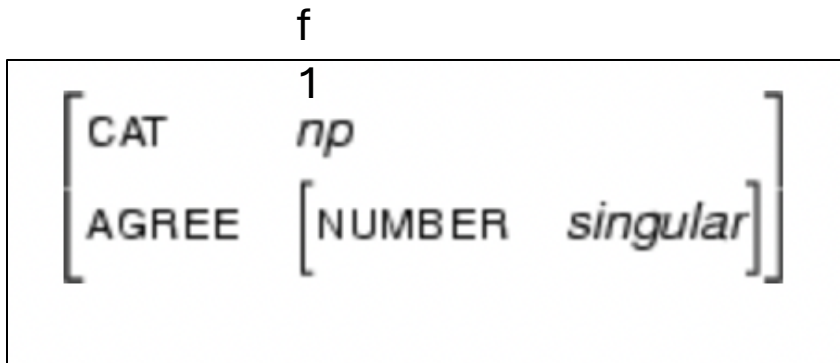
Unifizieren diese beide Merkmalsstrukturen?

$$\begin{bmatrix} \text{NUM} & \text{SG} \\ \text{GEN} & \text{FEM} \end{bmatrix} \sqcup \begin{bmatrix} \text{NUM} & \text{SG} \\ \text{GEN} & \text{MASK} \end{bmatrix}$$

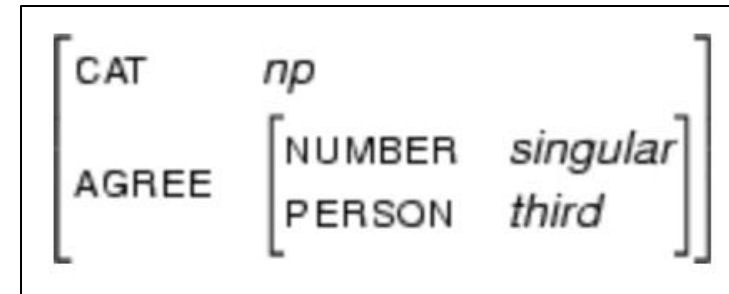
-> Nein, denn es besteht ein Widerspruch zwischen Gen: Fem und Gen: Mask

# Unifikation

Unifizieren diese Merkmalsstrukturen?



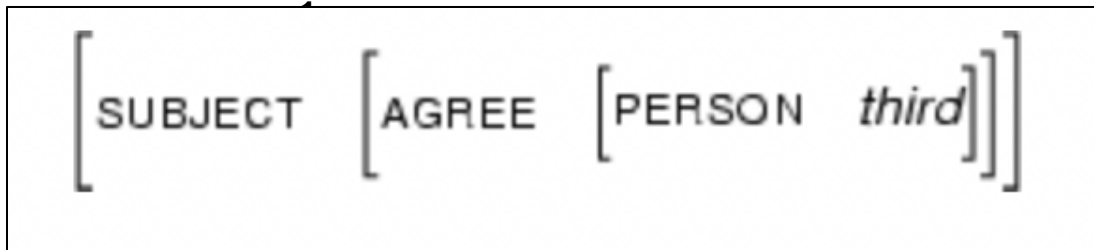
**Ja.** Das Ergebnis der Unifikation ist:



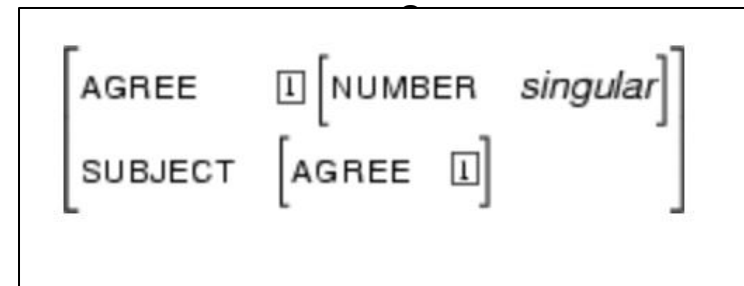
# Unifikation

Unifizieren diese Merkmalsstrukturen?

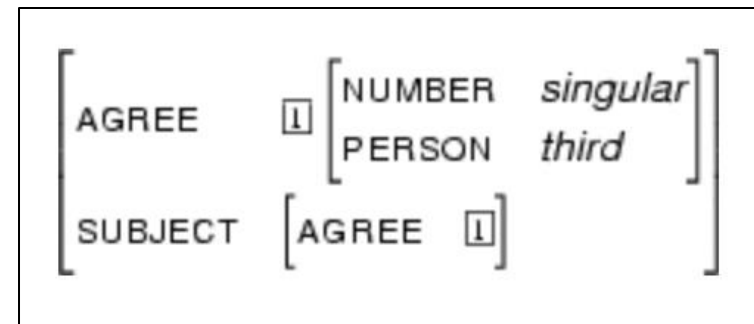
f



f



**Ja.** Das Ergebnis der Unifikation ist:



# Unifikation

Unifizieren diese Merkmalsstrukturen?



**Ja.**

[ Number                      sg  
Gender                      masc  
Person                      3]

# Subsumption

Subsumiert eine Merkmalsstruktur die andere?

f1

f2

[NUMBER *sg*]

[PERSON 3  
NUMBER *sg*]

**Ja, f1 subsumiert f2.**

f2 subsumiert f1 nicht, da f2 Informationen enthält, die nicht in f1 enthalten sind.

Quelle: <https://cs.union.edu/~striegnk/courses/nlp-with-prolog/html/node82.html#l111.sec.subsumption>



# Subsumption

Subsumiert eine Merkmalsstruktur die andere?

f1

f2

NUMBER	<i>sg</i>
GENDER	<i>masc</i>

PERSON	3
NUMBER	<i>sg</i>

**Nein.**

f1 enthält Informationen, die nicht in f2 enthalten sind, und vice versa.

Quelle: <https://cs.union.edu/~striegnk/courses/nlp-with-prolog/html/node82.html#l111.sec.subsumption>