

# **Eine nebenläufige Syntax-Semantik-Schnittstelle für Dependenzgrammatik**

Ralph Debusmann

## 1. Diplomarbeit:

- TDG - ein deklarativer Grammatikformalismus für Dependenzgrammatik

## 2. Forschungsvorhaben:

- eine nebenläufige Syntax-Semantik-Schnittstelle für TDG

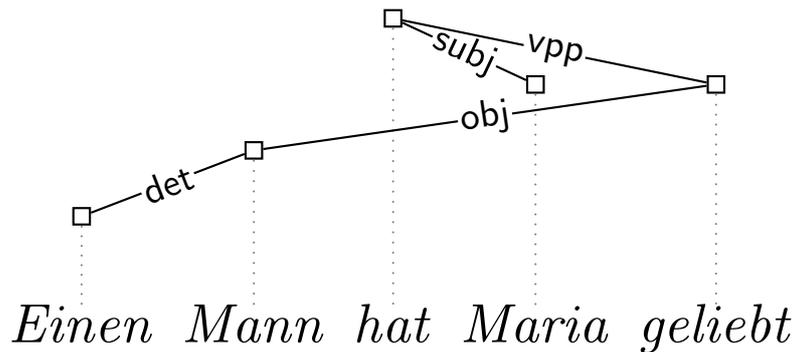
# Die Diplomarbeit

- Entwurf eines deklarativen Grammatikformalismus für Dependenzgrammatik
- wichtigste Frage hierbei: wie lässt sich Wortstellung deklarativ beschreiben
- Parsing mit Constrainttechnologie, Basis: Denys Duchiers Dependenzparser (Duchier 99)
- Hauptergebnis: Topologische Dependenzgrammatik (TDG)

# Geschichte der Abhängigkeitsgrammatik

- Tesnière 59, Lecerf 60
- Moskauer Schule/Meaning Text Theory (Melcuk 87)
- Prager Schule/Functional Generative Description (Sgall et al 86)
- außerdem: Dependency Unification Grammar (Hellwig 86), Lexibase (Starosta 88), Word Grammar (Hudson 90)

# Dependenzbäume



- repräsentieren syntaktische Struktur von Sätzen
- 1:1-Abbildung Knoten-Wörter, d.h. keine phrasalen Knoten
- Kantenbeschriftungen = syntaktischen Rollen
- ungeordnet

## Erlaubte Wortabfolgen

- Beispiel lizenziert  $5! = 120$  Wortabfolgen, davon aber nur 5 erlaubt

- erlaubt:

1. *Maria hat einen Mann geliebt.*
2. *Einen Mann hat Maria geliebt.*
3. *Geliebt hat Maria einen Mann.*
4. *Einen Mann geliebt hat Maria.*
5. *Hat Maria einen Mann geliebt?*

- unerlaubt:

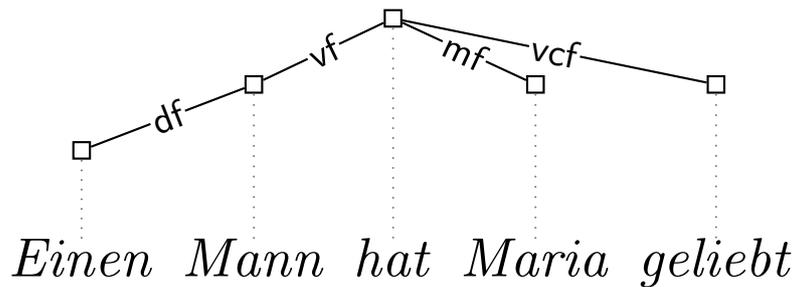
1. *\*Einen Mann geliebt Maria hat*
  2. *\*Einen Mann hat geliebt Maria*
  3. *\*Einen Mann Maria geliebt hat*
  4. *\*Einen Mann Maria hat geliebt*
  5. *\*Geliebt einen Mann hat Maria*
  6. *\*Geliebt einen Mann Maria hat*
  7. *\*Geliebt Maria einen Mann hat*
  8. *\*Geliebt Maria hat einen Mann*
- ... (115)

# Die Theorie der topologischen Felder

- deskriptive Theorie zur Beschreibung der deutschen Wortstellung, Höhle 1986 (Herling 1821, Erdmann 1886)
- wiederentdeckt z.B. von Kathol (1995, 2000), Müller (1999) für HPSG
- Idee: Satz in zusammenhängende Wortketten einteilen, die *topologische Felder* genannt werden:

Vorfeld	(	Mittelfeld	)	Nachfeld
<i>Maria</i>	<i>hat</i>	<i>einen Mann</i>	<i>geliebt.</i>	
<i>Einen Mann</i>	<i>hat</i>	<i>Maria</i>	<i>geliebt.</i>	
<i>Geliebt</i>	<i>hat</i>	<i>Maria einen Mann.</i>		
<i>Einen Mann geliebt</i>	<i>hat</i>	<i>Maria.</i>		
	<i>Hat</i>	<i>Maria einen Mann</i>	<i>geliebt?</i>	

# TDG: Topologiebäume



- Idee: zusätzlich zum ungeordneten Dependenzbaum: geordneter Topologiebaum
- Kantenbeschriftungen = topologische Felder
- repräsentiert geschachtelte topologische Struktur von Sätzen
- Töchter analog zu Ordnung auf Menge der Felder geordnet:

$vf \prec mf \prec vcf$   
 $Mann \prec Maria \prec geliebt$

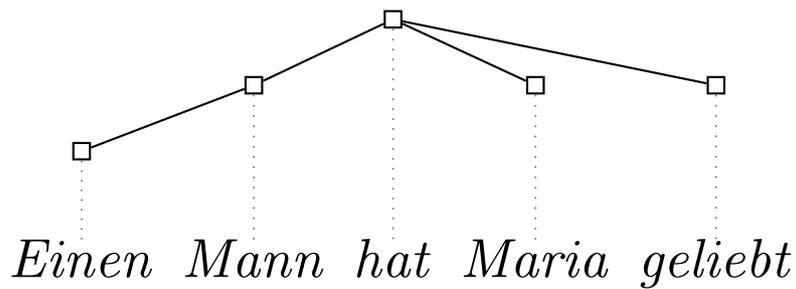
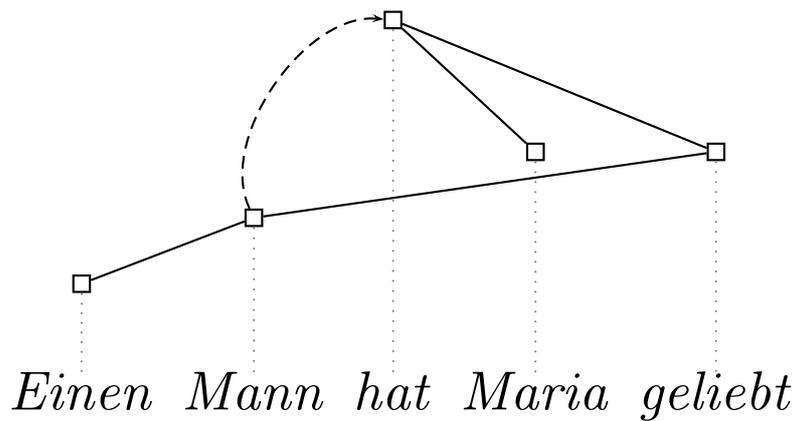
# TDG: Beziehung zwischen Abhängenz- und Topologiebäumen

- TDG-Analyse besteht aus je einem Abhängenz- und einem Topologiebaum
- beide haben die gleiche Knotenmenge aber verschiedene Kantenmengen
- *Kletterconstraint* setzt beide Bäume miteinander in Beziehung

# TDG: Kletterconstraint

- Klettern: Knoten können im Topologiebaum im Vergleich zum Dependenzbaum “hochklettern”
- Topologiebaum dadurch flacher als entsprechender Dependenzbaum
- ähnlich Reapes *Word Order Domains* (Reape 1994) für HPSG

# TDG: Beispiel für Kletterconstraint

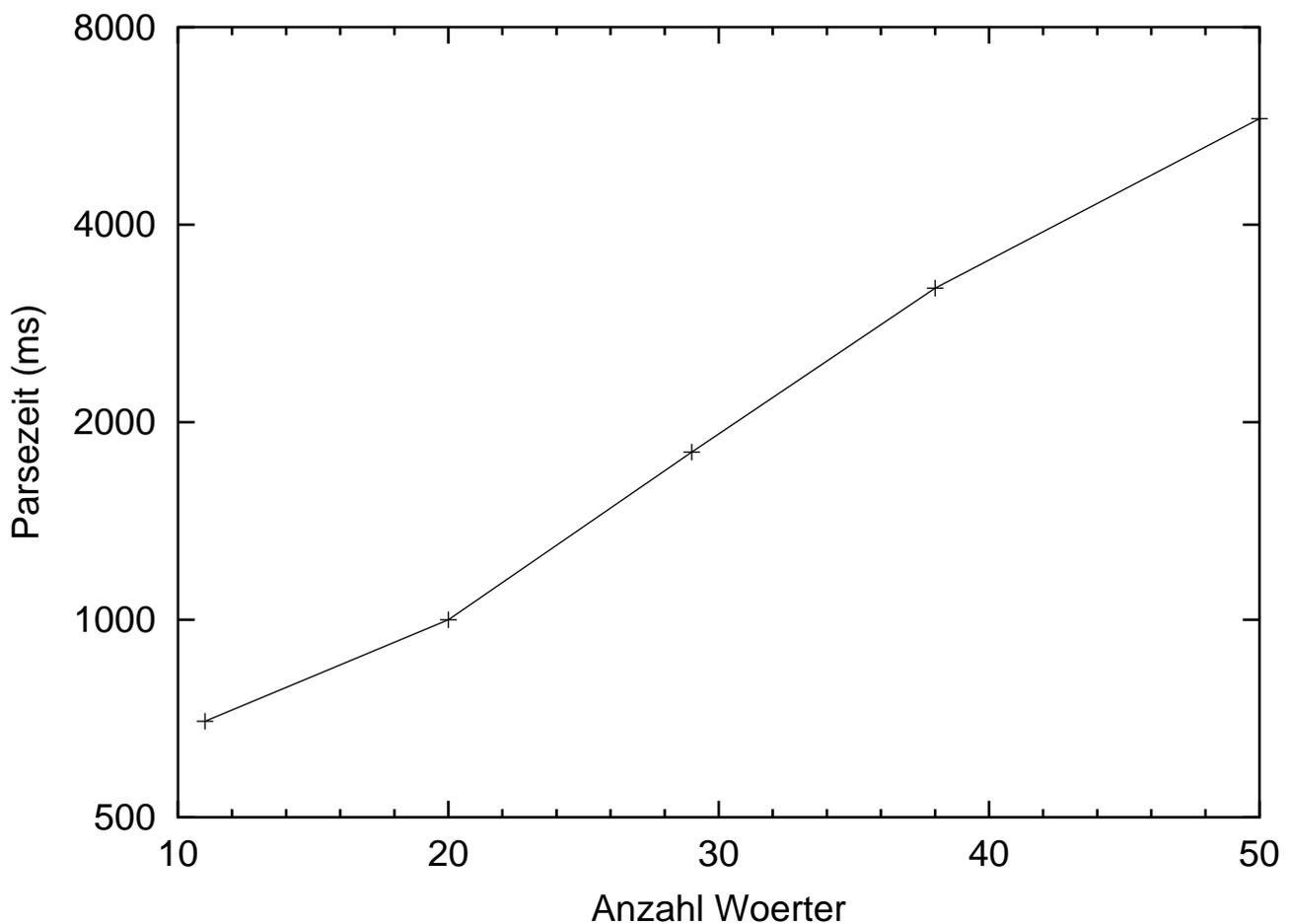


# TDG: Behandelte Phänomene

- fürs Deutsche:
  - Fragen, Aussagesätze, Nebensätze
  - Relativsätze (mit Extraposition, Rattenfänger-Konstruktionen)
  - Scrambling im Mittelfeld
  - (partielles) VP-Fronting, (partielle) VP-Extraposition, VP-Intraposition
  - Verbalkomplex-Phänomene (u.a. Auxiliary Flip, Zwischenstellung) (ACL-Papier 2001)
- fürs Holländische:
  - Cross-Serial Dependencies (CSD)
  - Verbalkomplex-Phänomene: Inversion, Interaktion von CSD und Inversion
  - (partielle) VP-Extraposition

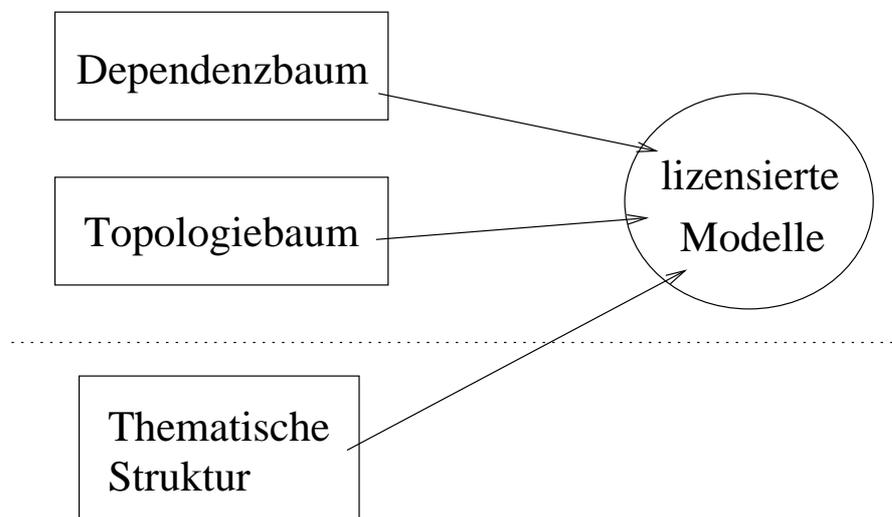
# TDG: Parsingzeiten

- theoretisch: Parsingproblem NP-vollständig (Koller 2001), in Praxis Laufzeit aber annähernd quadratisch
- Beispiel (gemessen auf 466MHz Celeron):

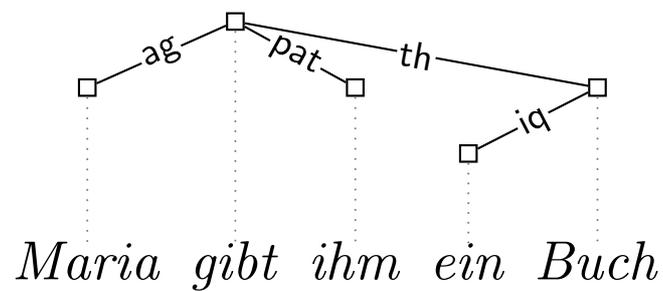
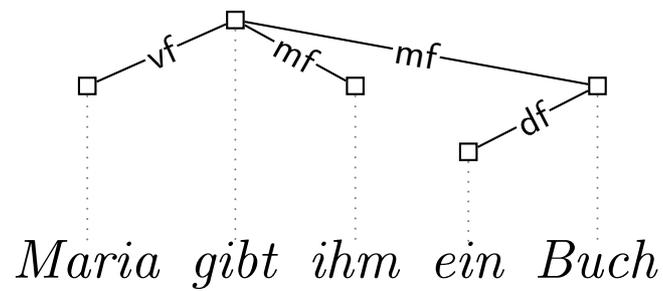
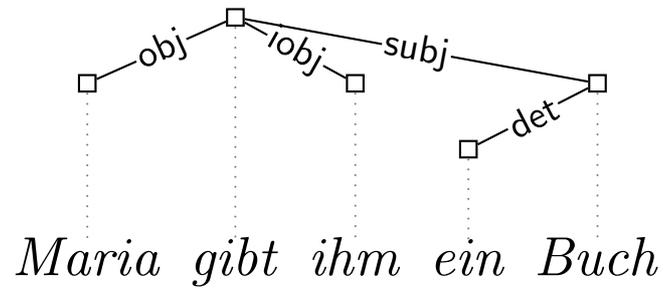


# Forschungsvorhaben: Syntax-Semantik-Schnittstelle

- Idee: neue Struktur zu TDG-Architektur hinzufügen: Thematische Struktur
- Bindeglied zwischen Syntax und modelltheoretischer Semantik
- neue Architektur:

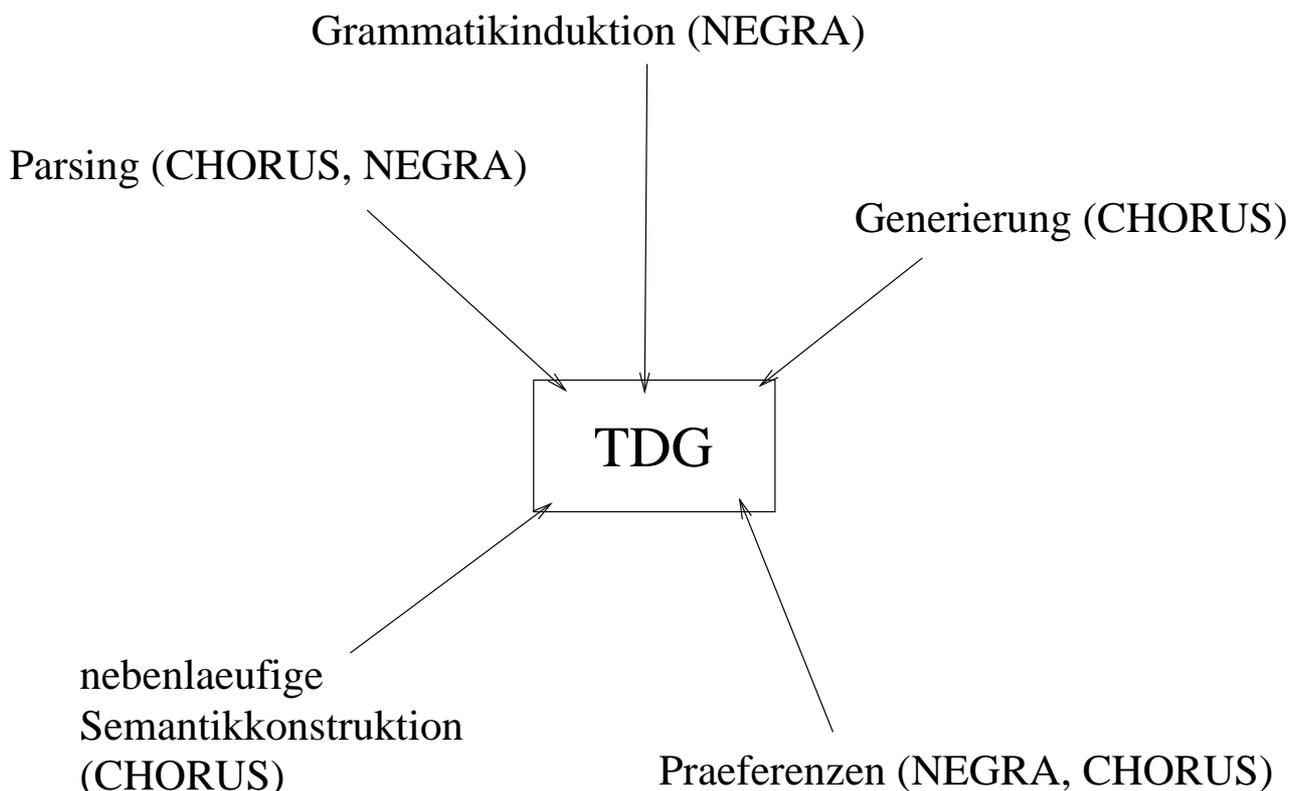


# Forschungsvorhaben: Beispielanalyse

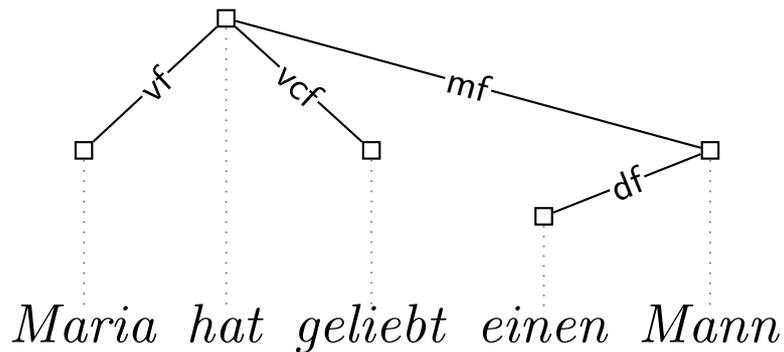


# Forschungsvorhaben: Das Umfeld

- Idee: TDG soll gemeinsames Framework für CL-Forschung und Applikationen in Saarbrücken sein (oben: bisher, unten: Zukunft):



## TDG: Verletzung der Ordnung

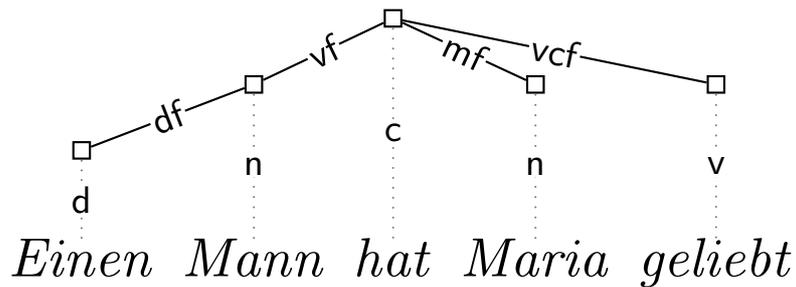


- nicht lizenziert. Ordnung auf der Menge der Felder:

$$vf \prec mf \prec vcf$$

- hier ist *geliebt* aber im *vcf* vor *Mann* im *mf*.

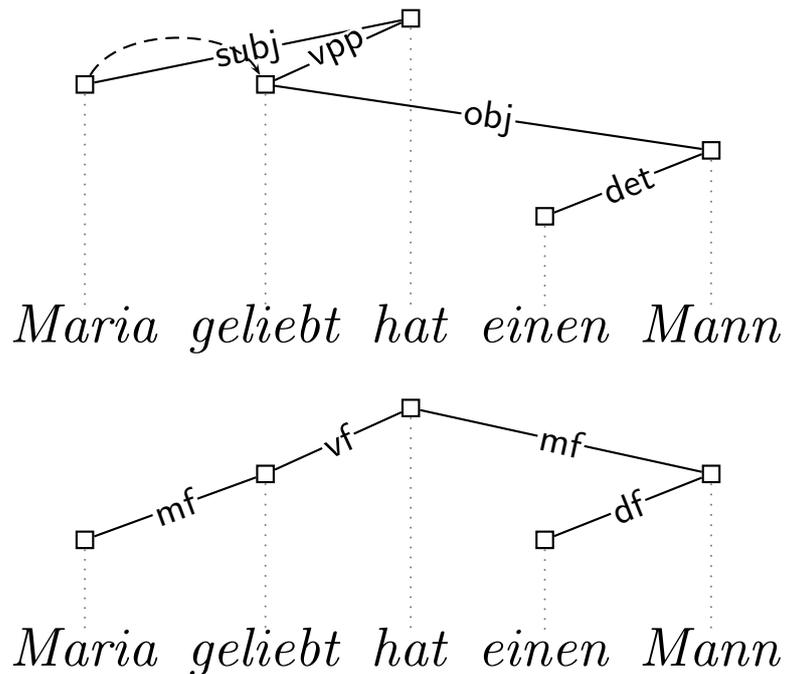
# TDG: Knotenbeschriftungen



- Wie werden Knoten selbst ihren Töchtern gegenüber angeordnet?
- Knoten (topologische Köpfe) und Töchter analog zu Ordnung auf Menge der Felder und *Knotenbeschriftungen* geordnet:

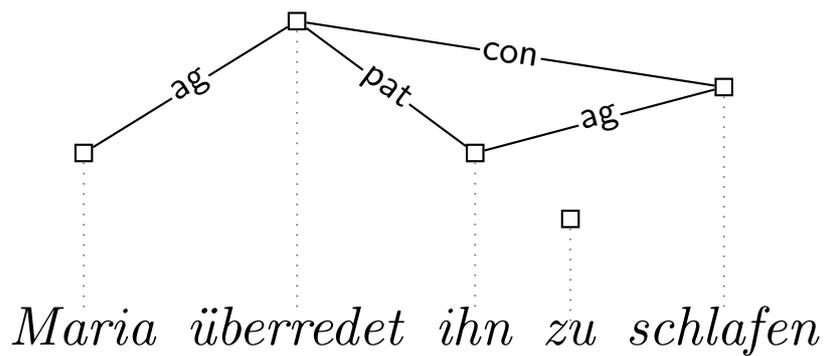
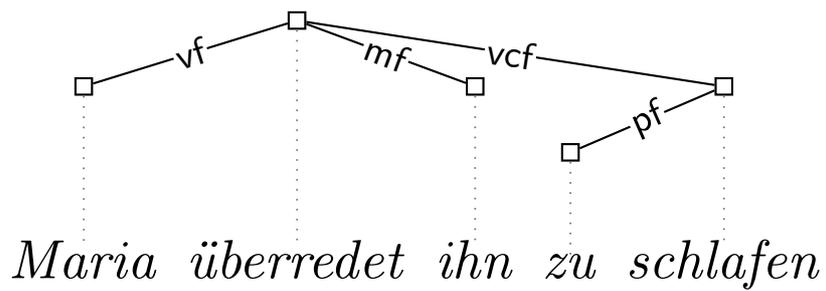
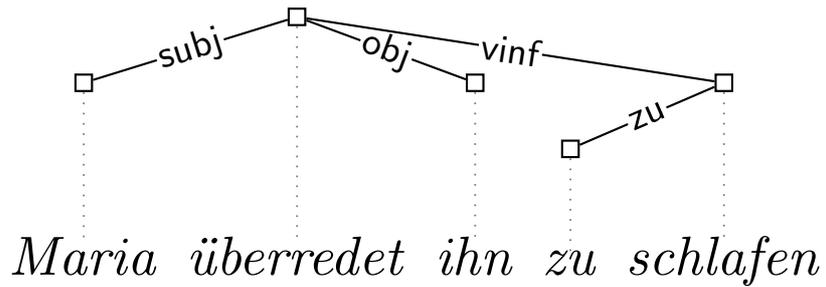
vf < c < mf < vcf  
*Mann* < *hat* < *Maria* < *geliebt*

# TDG: Verletzung des Kletterconstraints



- nicht lizenziert: *Maria* nicht geklettert, sondern zur Schwester *geliebt* bewegt

# Forschungsvorhaben: Kontrolle



# Forschungsvorhaben: Lexikoneinträge

## Kontrolle

*überredet* :

valency <sub>ID</sub>	:	{subj, obj, vinf}									
valency <sub>TH</sub>	:	{ag, pat, con}									
link	:	<table><tr><td>ag</td><td>:</td><td>subj</td></tr><tr><td>pat</td><td>:</td><td>obj</td></tr><tr><td>con</td><td>:</td><td>vinf</td></tr></table>	ag	:	subj	pat	:	obj	con	:	vinf
ag	:	subj									
pat	:	obj									
con	:	vinf									
raisedsubj	:	{obj}									

*schlafen* :

valency <sub>ID</sub>	:	∅			
valency <sub>TH</sub>	:	{ag}			
link	:	<table><tr><td>ag</td><td>:</td><td>subj</td></tr></table>	ag	:	subj
ag	:	subj			

- Wo ist das Subjekt von *schlafen*?
- Idee: Subjekte können auf Vorfahren im Dependenzbaum realisiert sein: z.B. das Subjekt von *schlafen* als Objekt von *überredet*.