Introduction to Tree Adjoining Grammar

Natural Language Syntax with TAG

Wolfgang Maier and Timm Lichte
University of Düsseldorf

DGfS-CL Fall School 2011

1st week, 5th session
Sep 2, 2011

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

SFB 991

Outline

1. Feature-Structure Based TAG (FTAG)

2. Multicomponent TAG (MCTAG)

FTAG (1)

Feature-structure based TAG (FTAG)

- Each node has a top and a bottom feature structure (except substitution nodes that have only a top).
- Nodes in the same elementary tree can share features (extended domain of locality).

FTAG (1)

Feature-structure based TAG (FTAG)

- Each node has a top and a bottom feature structure (except substitution nodes that have only a top).
- Nodes in the same elementary tree can share features (extended domain of locality).

Intuition:

- The top feature structure tells us something about what the node presents within the surrounding structure, and
- the bottom feature structure tells us something about what the tree below the node represents.

FTAG (1)

Feature-structure based TAG (FTAG)

- Each node has a top and a bottom feature structure (except substitution nodes that have only a top).
- Nodes in the same elementary tree can share features (extended domain of locality).

Intuition:

- The top feature structure tells us something about what the node presents within the surrounding structure, and
- the bottom feature structure tells us something about what the tree below the node represents.

In the final derived tree, both must be the same.

FTAG (1)

Feature-structure based TAG (FTAG)

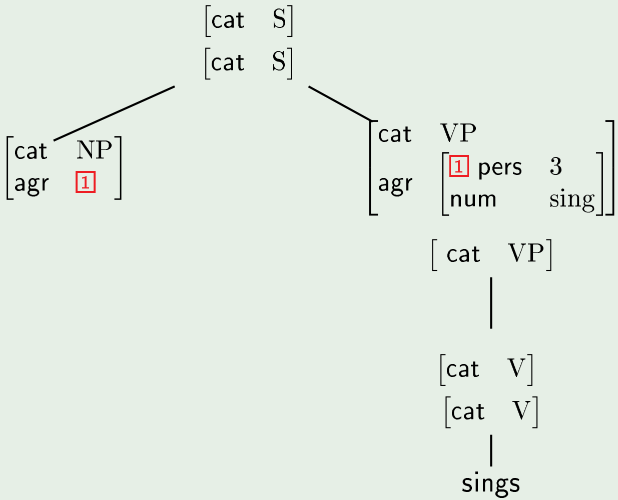
- Each node has a top and a bottom feature structure (except substitution nodes that have only a top).
- Nodes in the same elementary tree can share features (extended domain of locality).

Intuition:

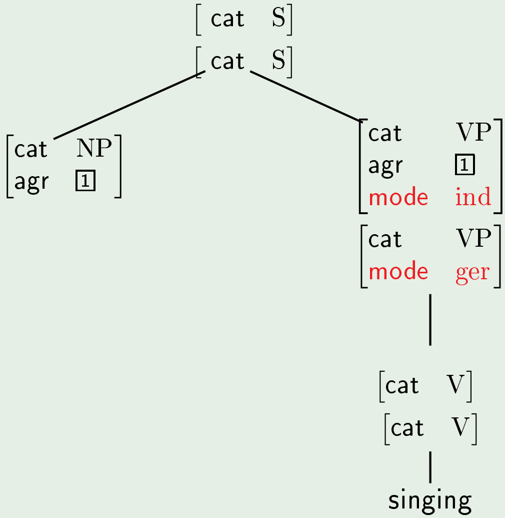
- The top feature structure tells us something about what the node presents within the surrounding structure, and
- the bottom feature structure tells us something about what the tree below the node represents.

In the final derived tree, both must be the same.
[Vijay-Shanker and Joshi, 1988]

FTAG (2)


[cat S]
[cat S]
[cat NP, agr 1]
[cat VP, agr [1 pers 3, num sing]]
[cat VP]
[cat V]
[cat V]
sings

FTAG (3)



cat S
cat S
cat NP
agr 1
cat VP
agr 1
mode ind
cat VP
mode ger
cat V
cat V
singing

FTAG (4)

Unification during derivation:

FTAG (4)

Unification during derivation:

- **Substitution**: the top of the root of the new initial tree unifies with the top of the substitution node

FTAG (4)

Unification during derivation:

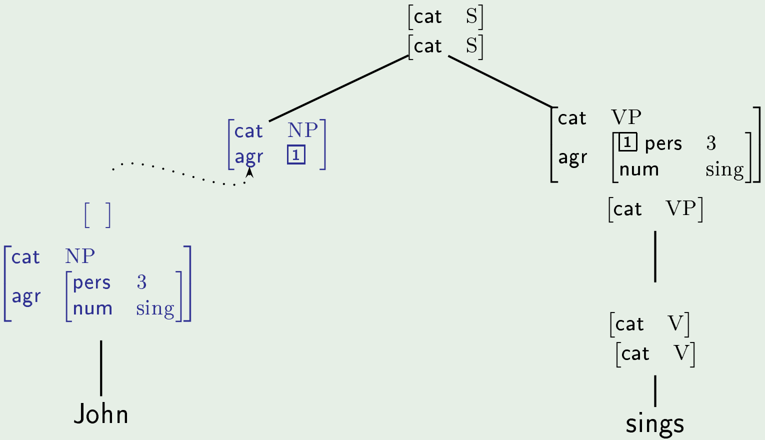
- **Substitution**: the top of the root of the new initial tree unifies with the top of the substitution node
- **Adjunction**: the top of the root of the new auxiliary tree unifies with the top of the adjunction site,
  and the bottom of the foot of the new tree unifies with the bottom of the adjunction site.

FTAG (4)

Unification during derivation:

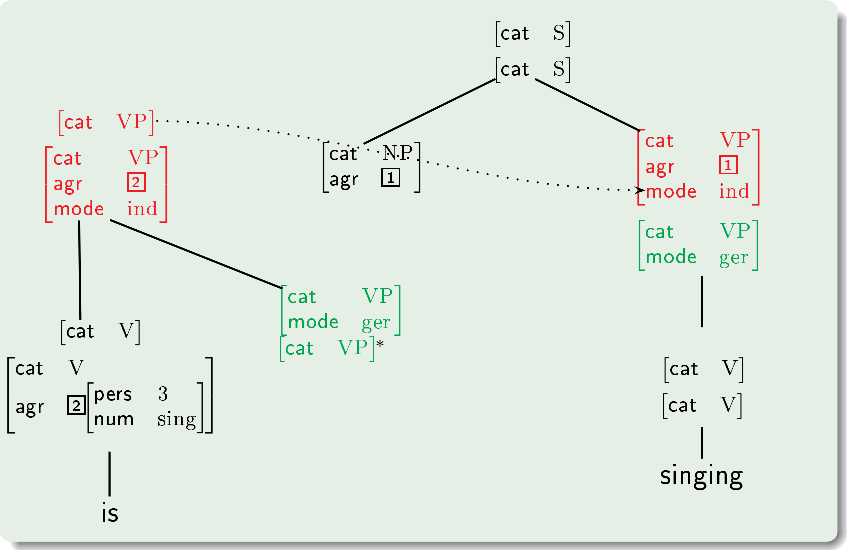
- **Substitution**: the top of the root of the new initial tree unifies with the top of the substitution node
- **Adjunction**: the top of the root of the new auxiliary tree unifies with the top of the adjunction site,
  and the bottom of the foot of the new tree unifies with the bottom of the adjunction site.
- In the final derived tree, top and bottom unify for all nodes.

FTAG (5)



[cat S]
[cat S]
[cat NP
agr 1]
[cat VP
agr [1 pers 3
num sing]]
[cat VP]
[ ]
[cat NP
agr [pers 3
num sing]]
John
[cat V]
[cat V]
sings

FTAG (6)



[cat S]
[cat S]
[cat VP]
[cat VP, agr 2, mode ind]
[cat NP, agr 1]
[cat VP, agr 1, mode ind]
[cat VP, mode ger]
[cat V]
[cat V, agr 2 [pers 3, num sing]]
[cat VP, mode ger]
[cat VP]*
[cat V]
[cat V]
is
singing

FTAG (7)

- In FTAG, there are no explicit adjunction constraints. Instead, adjunction constraints are expressed via feature unification requirements.

FTAG (7)

- In FTAG, there are no explicit adjunction constraints. Instead, adjunction constraints are expressed via feature unification requirements.
- Important: LTAG feature structures are restricted; there is only a finite set of possible feature structures.

FTAG (7)

- In FTAG, there are no explicit adjunction constraints. Instead, adjunction constraints are expressed via feature unification requirements.
- Important: LTAG feature structures are restricted; there is only a finite set of possible feature structures.

Therefore, the following can be shown:

For each FTAG there exists a weakly equivalent TAG with adjunction constraints and vice versa. The two TAGs generate even the same sets of trees, only with different node labels.

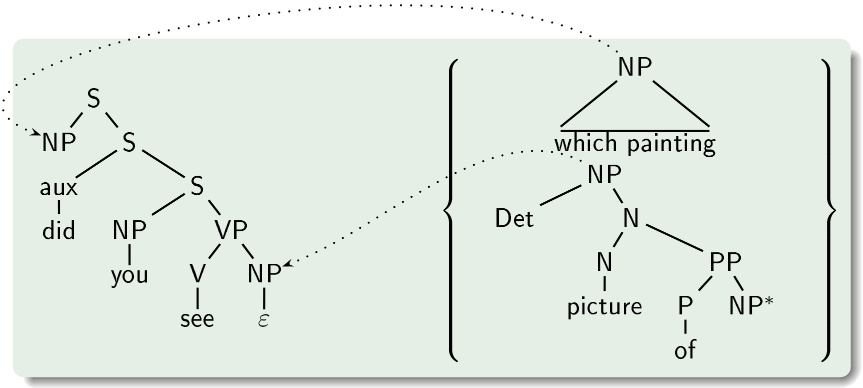
MCTAG: Motivation (1)

Multicomponent Tree Adjoining Grammars (MCTAGs)

- First introduced in [Joshi et al., 1975] as *simultaneous TAGs*, later redefined as *multicomponent TAGs (MCTAGs)* in [Weir, 1988, Joshi, 1985]
- Linguistic motivation: Separate the contribution of a lexical item into several components
- In each derivation step, a new set is picked and all trees from the set are added simultaneously, i.e., they are attached (by substitution or adjunction) to different nodes in the already derived tree.

MCTAG: Motivation (2)

(1) which painting$_i$ did you see a picture of t_i



MCTAG: Motivation (3)

Constructions that require multicomponents:

- Extraction out of complex NPs [Kroch, 1989], stranding phenomena, in particular "picture-NPs":

  (1) which castle did you paint a picture of?

- Subject-aux inversion in raising questions [Frank, 2008]

  (2) Does John seem to annoy you?

- Scrambling in German [Rambow, 1994]

  (3) dass den Kühlschrank niemand zu reparieren versprochen hat

Definition (MCTAG)

An MCTAG is a tuple $G = \langle N, T, S, I, A, f_{OA}, f_{SA}, \mathcal{A} \rangle$ *such that:*

- $G_{TAG} := \langle N, T, S, I, A, f_{OA}, f_{SA} \rangle$ *is a TAG with adjunction constraints, and*
- $\mathcal{A} \subseteq P(I \cup A)$ *is a set of subsets of* $I \cup A$*, the set of elementary tree sets.*[a]

[a] $P(X)$ is the set of subsets of some set X.

Without loss of generality, we can assume that $\mathcal{A}$ is a partition of $I \cup A$.

Definition (MCTAG derivation)

$\gamma \Rightarrow \gamma'$ is a derivation step in G iff there is an instance $\{\gamma_1, \ldots, \gamma_n\}$ of an elementary tree set in $\mathcal{A}$ and there are pairwise different nodes $v_1, \ldots, v_n$ in γ such that $\gamma' = \gamma[v_1, \gamma_1] \ldots [v_n, \gamma_n]$.

As in TAG, a derivation starts from an initial tree and in the end, in the final derived tree, all leaves must have terminal labels (or the empty word) and there must not be any *OA* constraints left.

Different types of MCTAG

An MCTAG is called

Different types of MCTAG

An MCTAG is called

- tree-local iff in each derivation step, the nodes the new trees attach to belong to the same elementary tree.

Different types of MCTAG

An MCTAG is called

- tree-local iff in each derivation step, the nodes the new trees attach to belong to the same elementary tree.
- set-local iff in each derivation step, the nodes the new trees attach to belong to the same elementary tree set.

Different types of MCTAG

An MCTAG is called

- tree-local iff in each derivation step, the nodes the new trees attach to belong to the same elementary tree.
- set-local iff in each derivation step, the nodes the new trees attach to belong to the same elementary tree set.
- non-local otherwise.

Different types of MCTAG

An MCTAG is called

- tree-local iff in each derivation step, the nodes the new trees attach to belong to the same elementary tree.
- set-local iff in each derivation step, the nodes the new trees attach to belong to the same elementary tree set.
- non-local otherwise.

Usually, the term "MCTAG" without specification of the locality means "set-local MCTAG".

Tree-local MCTAG

Tree-local MCTAG and TAG are equivalent since we can precompile the possible adjunctions and substitutions in an elementary tree:

Proposition

Tree-local MCTAG are strongly equivalent to TAG.

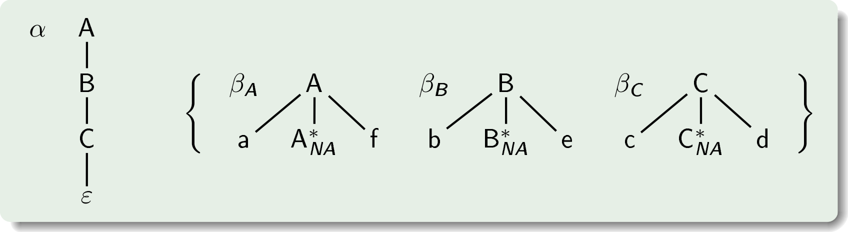
Tree-local MCTAG

Tree-local MCTAG $\Rightarrow$ TAG

To construct strongly equivalent TAG from given tree-local MCTAG, adopt corresponding adjunction constraints that enforce the simultaneous adjunctions of all elementary trees from a tree set.

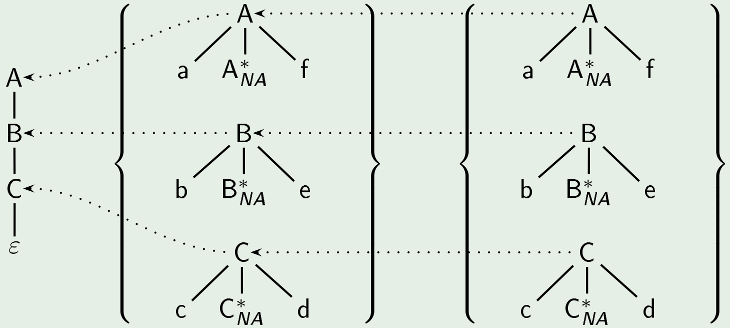
But: the number of elementary trees in the grammar can increase in an exponential way in this construction ($\Rightarrow$ rather a bad strategy for tree-local MCTAG parsing).

Set-local MCTAG for $L_6 = \{a^n b^n c^n d^n e^n f^n \mid n \geq 0\}$:



Set-local MCTAG

Derivation for *aabbccddeeff* :



Non-local MCTAG

Proposition

Unrestricted non-local MCTAG is NP-hard [Rambow and Satta, 1992]. This also hold for lexicalized non-local MCTAG and for non-local MCTAG with dominance links [Champollion, 2007].

Why not TAG for German?

The order of complements (and adjuncts) of a verb is flexible.

(2) *Peter liebt Susi.*
1: `Peter loves Susi`
2: `Susi loves Peter`

(3) *dass Peter heute den Kühlschrank repariert hat*
dass den Kühlschrank heute Peter repariert hat
. . .
('that Peter has repaired the fridge today')

TAG is inappropriate for German, because it is:

- not powerful enough for some constructions (i.e., coherent constructions)
- not descriptively adequat (i.e., one elementary tree for each permutation)

Why not TAG for German?

The order of complements (and adjuncts) of a verb is flexible.

(2) *Peter liebt Susi.*
1: `Peter loves Susi`
2: `Susi loves Peter`

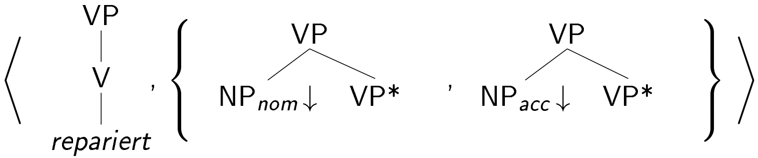
(3) *dass Peter heute den Kühlschrank repariert hat*
dass den Kühlschrank heute Peter repariert hat
...
('that Peter has repaired the fridge today')

TAG is inappropriate for German, because it is:

- not powerful enough for some constructions (i.e., coherent constructions)
- not descriptively adequat (i.e., one elementary tree for each permutation)

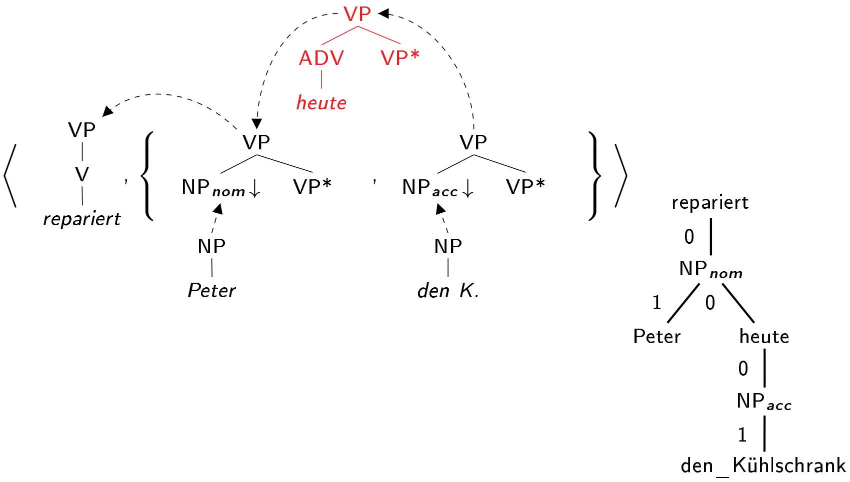
TT-MCTAG: a TAG-extension for German

- Multi-Component TAG (MCTAG) with shared-nodes locality
- Elementary structures are **tuples** $\langle \gamma, \{\beta_1, \ldots, \beta_n\} \rangle$:
  - a lexicalized elementary tree γ (the head tree)
  - a tree set $\{\beta_1, \ldots, \beta_n\}$ (the complement trees)
- **Meaning of tree tuples:** During derivation, the β-trees have to attach to the γ-tree (via node sharing).
- **Node sharing:** In the derivation tree,
  1. a β-tree must either be the immediate daughter of its γ-tree,
  2. or the β-tree must be connected to the daughter of the γ-tree via a chain of root adjunctions.



TT-MCTAG example

(4) *dass den Kühlschrank* *heute* *Peter repariert*
("that Peter repairs the fridge today")

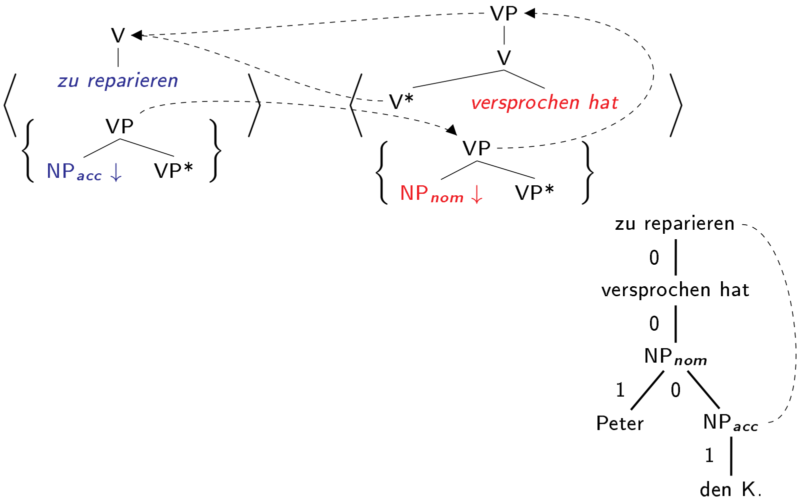


The linguistic application: Principles of complementation

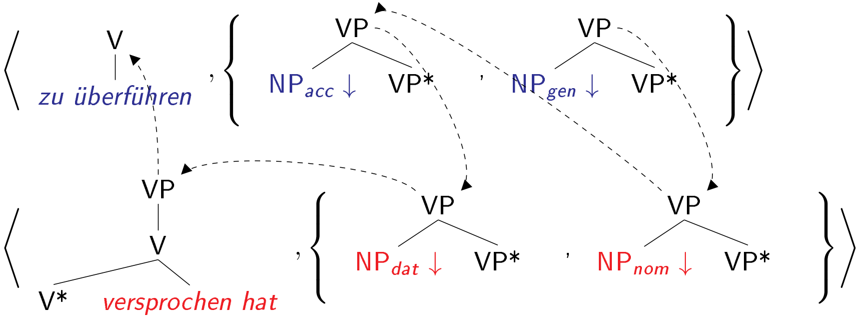
- A tuple = a subcategorization frame, i.e. a head and its complements (as substitution slots and footnodes)
- Substitution = strong islands
- no empty elements (traces, PRO)
- no base order of complements

⇒ **less elementary structures than in a German TAG**

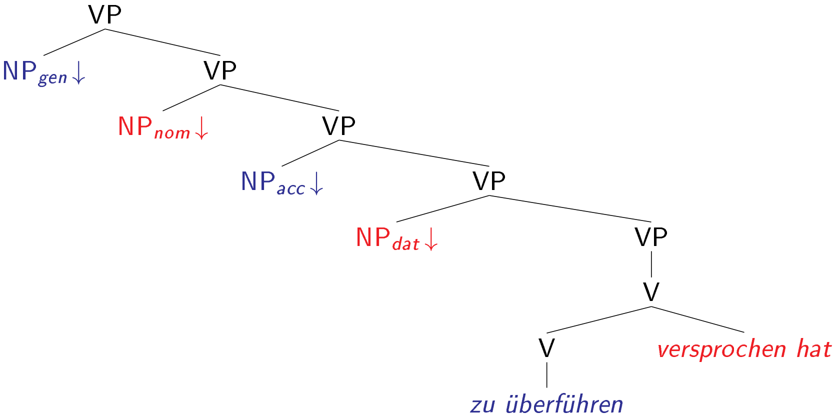
(5) *dass den Kühlschrank Peter zu reparieren versprochen hat*
('that Peter has promised to repair the fridge today')



(6) *dass des Verbrechens der Detektiv den Verdächtigen dem Klienten* ‖ *zu überführen versprochen hat*
('that the detective has promised the client to indict the suspect of the crime')



(6) *dass des Verbrechens der Detektiv den Verdächtigen dem Klienten ‖ zu überführen versprochen hat*
('that the detective has promised the client to indict the suspect of the crime')



Champollion, L. (2007).
Lexicalized non-local MCTAG with dominance links is NP-complete.
In Penn, G. and Stabler, E., editors, Proceedings of Mathematics of Language (MOL) 10, CSLI On-Line Publications.

Frank, R. (2008).
Syntax and ltag.
Slides of a tutorial at TAG+9.

Joshi, A. K. (1985).
Tree adjoining grammars: How much contextsensitivity is required ro provide reasonable structural descriptions?
In Dowty, D., Karttunen, L., and Zwicky, A., editors, Natural Language Parsing, pages 206–250. Cambridge University Press.

Joshi, A. K., Levy, L. S., and Takahashi, M. (1975).
Tree Adjunct Grammars.
Journal of Computer and System Science, 10:136–163.

Kroch, A. (1989).
Asymmetries in long-distance extraction in a Tree Adjoining Grammar.
In Baltin and Kroch, editors, Alternative Conceptions of Phrase Structure. University of Chicago.

Rambow, O. (1994).
Formal and Computational Aspects of Natural Language Syntax.
PhD thesis, University of Pennsylvania.

Rambow, O. and Satta, G. (1992).
Formal Properties of Non-Locality.
In Proceedings of 1st International Workshop on Tree Adjoining Grammars, Philadelphia.

Vijay-Shanker, K. and Joshi, A. K. (1988).
Feature structures based tree adjoining grammar.
In Proceedings of COLING, pages 714–719, Budapest.

Weir, D. J. (1988).
Characterizing mildly context-sensitive grammar formalisms.
PhD thesis, University of Pennsylvania.