

Tree Unification with Constraints (TUCO) as a model of Flat Syntax

Timm Lichte

Heinrich-Heine-University of Düsseldorf, Germany

TreeGrasp Kick-off Meeting, February 21 2018



SFB 991



HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

What this talk is about

How to design a syntax framework/formalism?

Number one advice

Try to avoid idealizations! Keep it flexible and unbiased!

Some daunting examples:

- headedness as a primary category

This talk is about trying to do better with **TUCO!**

But first: What is syntactic structure anyway??

What this talk is about

How to design a syntax framework/formalism?

Number one advice

Try to avoid idealizations! Keep it flexible and unbiased!

Some daunting examples:

- headedness as a primary category
- valency as a primary category

This talk is about trying to do better with **TUCO!**

But first: What is syntactic structure anyway??

What this talk is about

How to design a syntax framework/formalism?

Number one advice

Try to avoid idealizations! Keep it flexible and unbiased!

Some daunting examples:

- headedness as a primary category
- valency as a primary category
- strictly binary branching phrase structures

This talk is about trying to do better with **TUCO!**

But first: What is syntactic structure anyway??

What this talk is about

How to design a syntax framework/formalism?

Number one advice

Try to avoid idealizations! Keep it flexible and unbiased!

Some daunting examples:

- headedness as a primary category
- valency as a primary category
- strictly binary branching phrase structures
- global locality restrictions

This talk is about trying to do better with **TUCO!**

But first: What is syntactic structure anyway??

What this talk is about

How to design a syntax framework/formalism?

Number one advice

Try to avoid idealizations! Keep it flexible and unbiased!

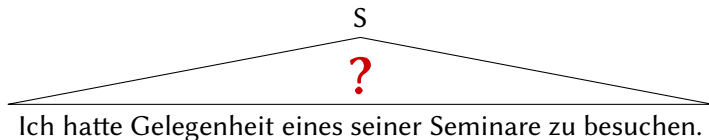
Some daunting examples:

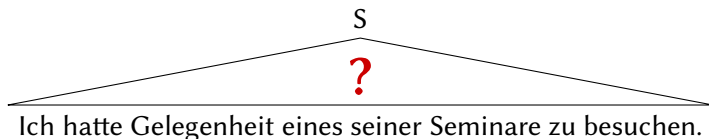
- headedness as a primary category
- valency as a primary category
- strictly binary branching phrase structures
- global locality restrictions
- strong lexicalism / constructionism

This talk is about trying to do better with **TUCO!**

But first: What is syntactic structure anyway??

In search of the syntactic structure





What tree am I supposed to grasp???

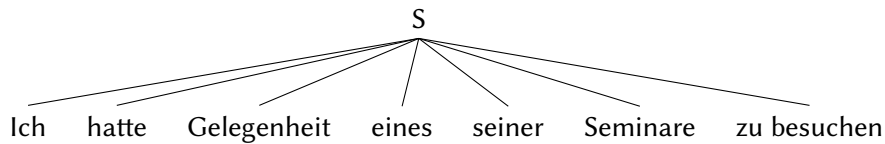
In search of the syntactic structure (cont.)

In search of the syntactic structure (cont.) (cont.)

S

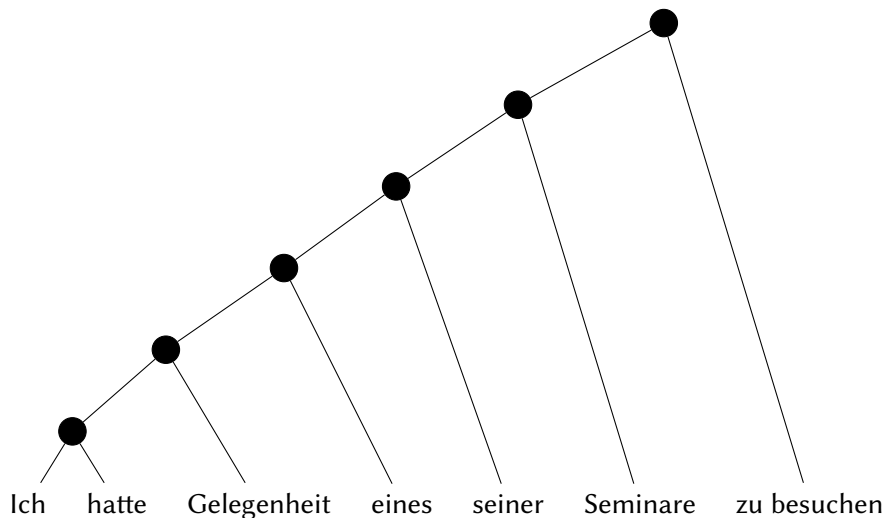
Ich hatte Gelegenheit eines seiner Seminare zu besuchen

In search of the syntactic structure (cont.) (cont.)

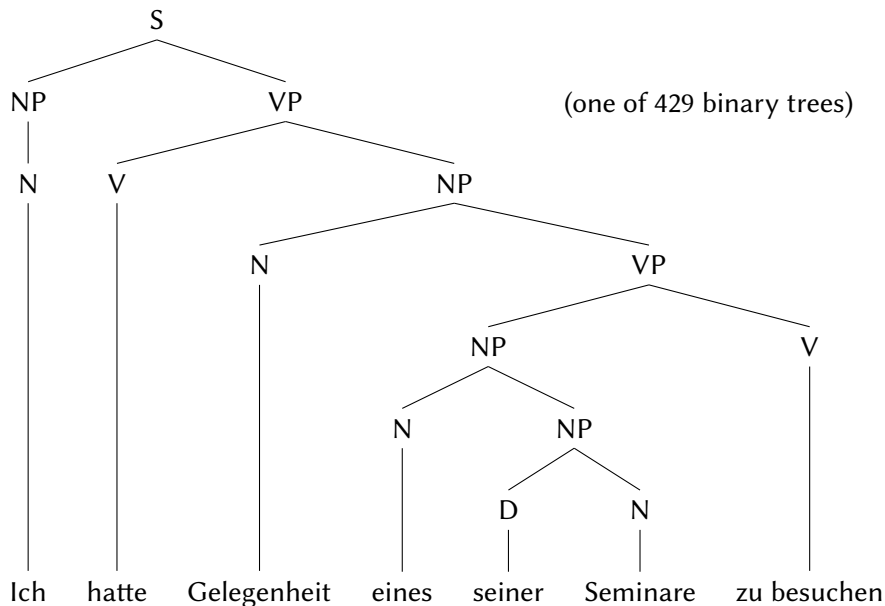


In search of the syntactic structure (cont.) (cont.)

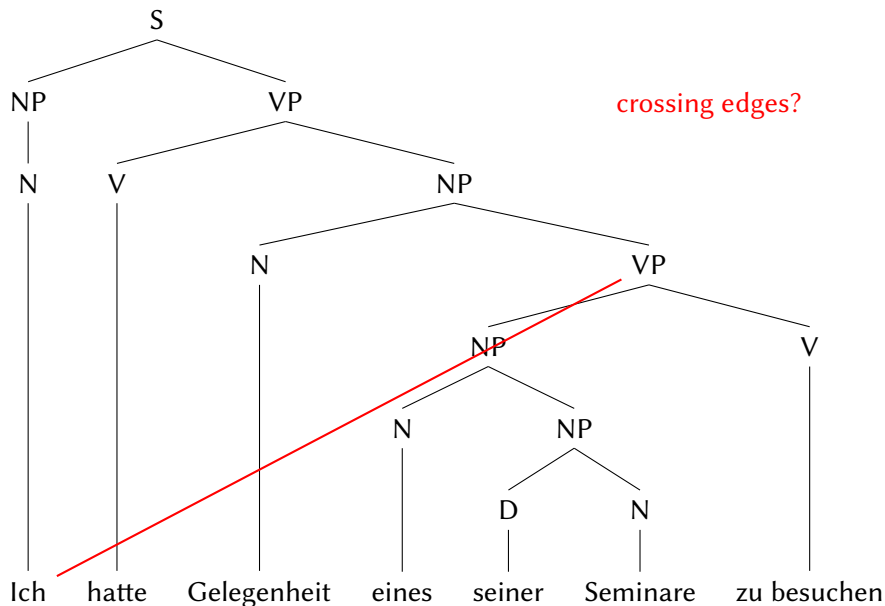
(à la Left-Associative Grammar)



In search of the syntactic structure (cont.) (cont.)

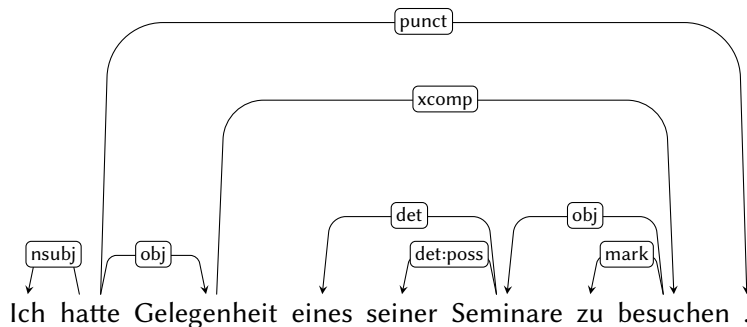


In search of the syntactic structure (cont.) (cont.)



In search of the syntactic structure (cont.) (cont.)

(from UD_German)



In search of the syntactic structure (cont.)

A fundamental issue:

- There is no commonly agreed upon syntactic structure.

Lets try:

Is there a theory-neutral syntactic structure that we could agree upon when guided only/mainly by intuition?

- What is the **meaning** of structural distinction (vertically and horizontally)?
- Do we want abstract structural entities (other than the words)?

A fundamental issue:

- There is no commonly agreed upon syntactic structure.
- Syntactic structures are largely dependent on specific grammar formalisms/theories.

Lets try:

Is there a theory-neutral syntactic structure that we could agree upon when guided only/mainly by intuition?

- What is the **meaning** of structural distinction (vertically and horizontally)?
- Do we want abstract structural entities (other than the words)?

A fundamental issue:

- There is no commonly agreed upon syntactic structure.
- Syntactic structures are largely dependent on specific grammar formalisms/theories.
- I think this is bad: **uniformity facilitates comparison of theories and sharing of data.**

Lets try:

Is there a theory-neutral syntactic structure that we could agree upon when guided only/mainly by intuition?

- What is the **meaning** of structural distinction (vertically and horizontally)?
- Do we want abstract structural entities (other than the words)?

A uniform, theory-neutral syntactic structure?

First: **Derived Structure (DedS)** or **Dervation Structure (DonS)**?

- DonS seems to depend more on the theory/framework.
- DedS directly includes the “surface”, and this is what we can actually observe best (other than semantics).

⇒ We set “syntactic structure” = DedS

Why do we need DedS at all?

- There is no purely syntactic DedS – semantics is **always** involved.
- Hence Dynamic Syntax sais “No!” and relies on DonS ...
- Still it bears aspects that are alien to semantics proper: a “**surface**”.

⇒ DedS is a product of the **syntax-semantics interface**.

If we seek a uniform, theory neutral syntactic structure, we have to look for it in the set of DedS.

What is the “right” DedS?

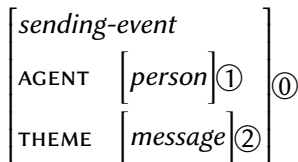
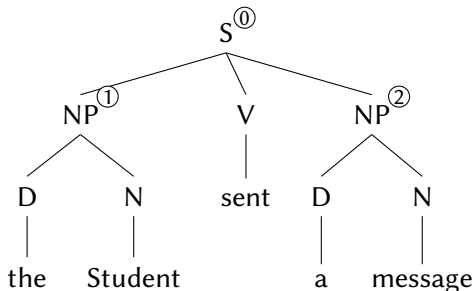
- compatible with strong intuitions
- minimal
- hence uniform and theory-neutral

The unfortunate gaps:

- intuitions may still vary
- the source of “strong intuitions” may be hidden – related to semantics?
- the “surface” is fixed, but the semantics is damn volatile.

What the “right” operationalizable DedS?

- 1 Tokens of a sentence form contiguous **constructs**.
- 2 Constructs correspond to **semantic entities**.
- 3 Adjacent constructs correspond to different semantic entities or different paths to those entities.



Potential problems:

- semantically idiomatic MWE

(1) The student shot the breeze.

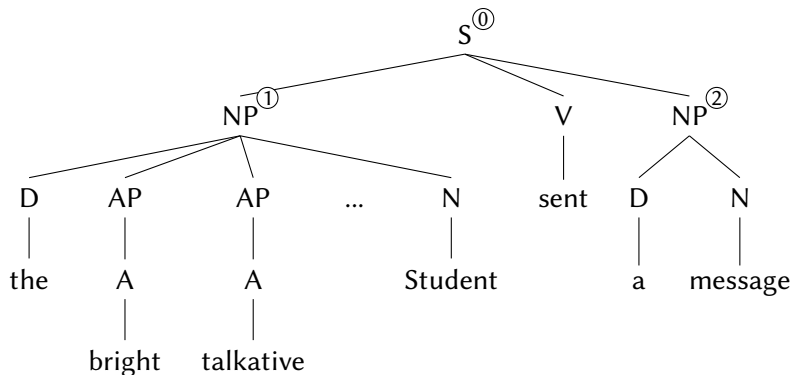
DedS is based on the **literal meaning**.

- unclear semantics: Does *has sent* correspond to one semantic entity or two (reference and event time)?

Choose the **simpler** alternative for DedS!

⇒ We then call this **uniform DedS (uDedS)**.

How to generate uDedS?

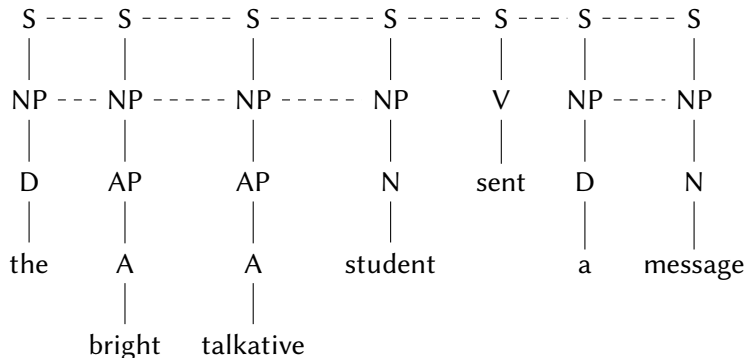


Not with: CFG/TSG, TAG, CG

But maybe with: TAG + sister adjunction

And surely with: **Tree Unification Grammar** (TUG)

An analysis with TUG



Desiderata

- means to express linearization constraints
- means to express agreement constraints
- means to express valency constraints
- means to express phrasal units

Constraints use **tree descriptions**:

- nodes with optional label, properties and polarization marker:
 $! \{p_1, \dots, p_n\}$
- dominance: $\rightarrow, \rightarrow^+, \rightarrow^*$
- precedence: $<, <^+, <^*$
- conjunction and disjunction

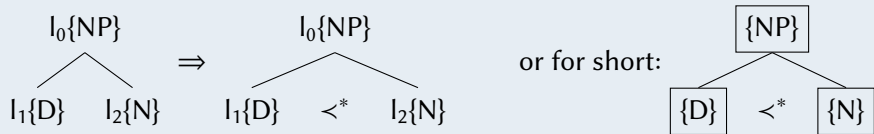
Let D_1 and D_2 be sets of descriptions, constraints have the form

$$D_1 \Rightarrow D_2$$

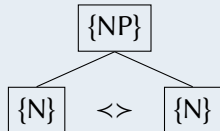
and are true in a Model \mathcal{M} , iff \mathcal{M} satisfies $D_1 \cup D_2$ or it does not satisfy D_1 .

TUCO: linear precedence and uniqueness

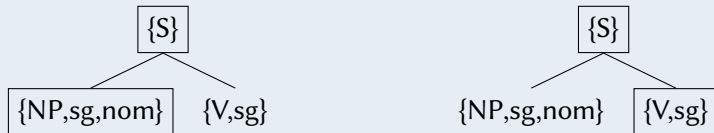
Linear precedence between determiners and nouns:



Uniqueness of the noun:

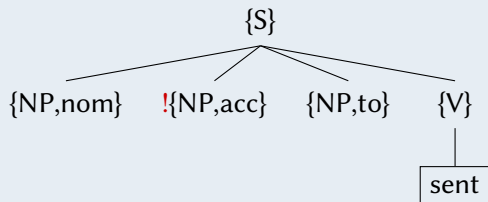


Subject-verb agreement (two constraints):



Agreement can be also dealt with in the parent node.

Valency also triggers the prediction of nodes (potentially without yield):

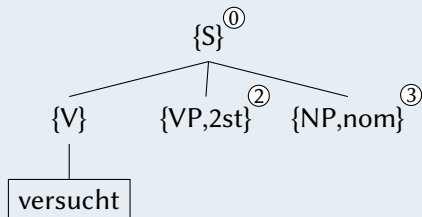


TUCO: valency and coherent constructions

Challenge: coherent constructions in German

(2) Peter hat **ihn** versucht **uns zu geben**.

Peter has him tried us to give
'Peter has tried to give it to us.'



<i>trying-event</i>			
AGENT	[]	③	①②
THEME	[]	②	

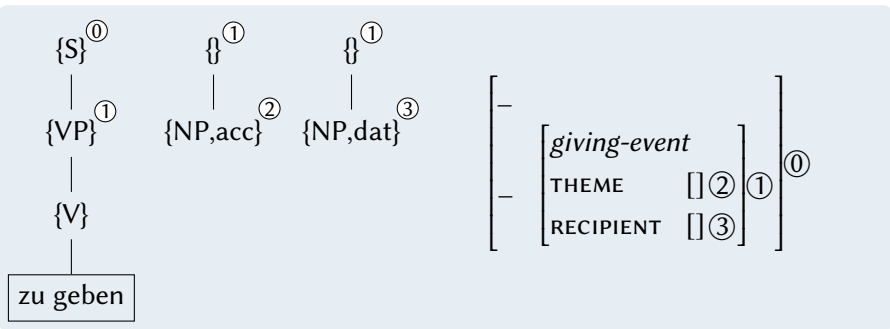
For this, interface links can be used.

Note: Every node has exactly one interface link!

Challenge: coherent constructions in German

(3) Peter hat **ihn** versucht **uns zu geben**.

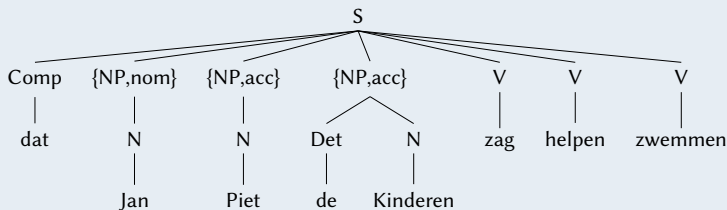
Peter has him tried us to give
'Peter has tried to give it to us.'



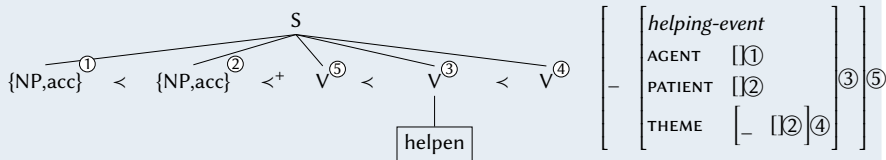
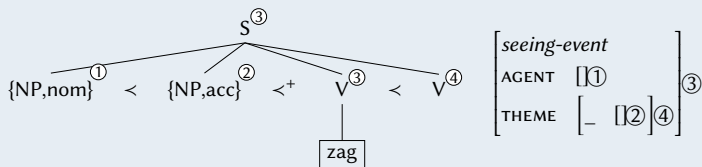
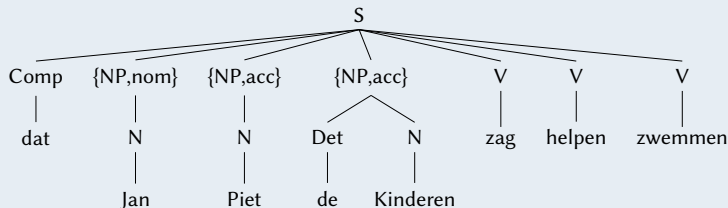
TUCO: valency and cross-serial dependencies

Challenge: cross-serial dependencies

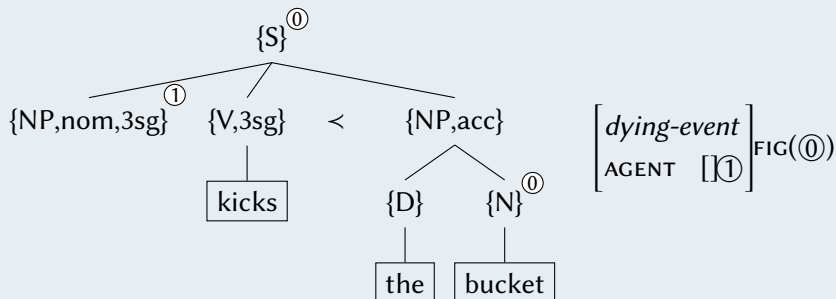
- (4) dat Jan Piet de kinderen zag helpen zwemmen.
that Jan Piet the children saw help swim
'that Jan saw Piet helping the children to swim.'



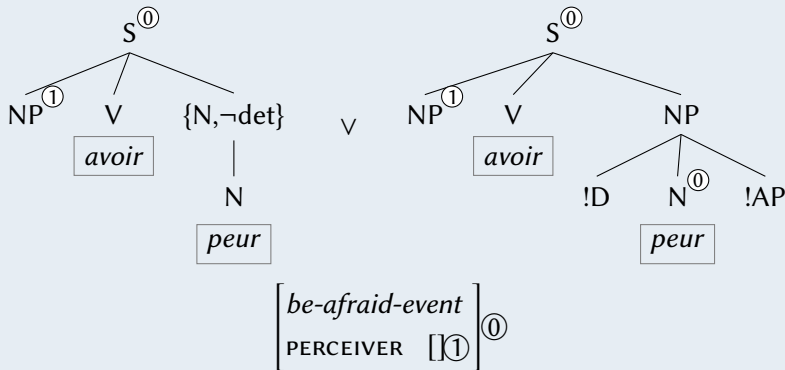
TUCO: valency and cross-serial dependencies



MWEs are treated by single tree constraints thanks to their EDL:



- (4) a. avoir (*une) peur
 b. avoir une peur folle



The general ideas are:

- to generate DedS of Flat Syntax with constrained tree unification (TU+TC = TUCO)
- head-, valency-, locality-, architecture-agnostic
- generation is straightforward (low degree of freedom)
- constraints are proliferating (high degree of freedom)

But the proposal is still rather programmatic:

- many details of the framework remain to be fixed.
 - What constraints need to be true to make a DedS grammatical?
 - How to enforce that “boxed” nodes were previously generated?
- there are two poles: use only TU – use only TC
- ...

References