

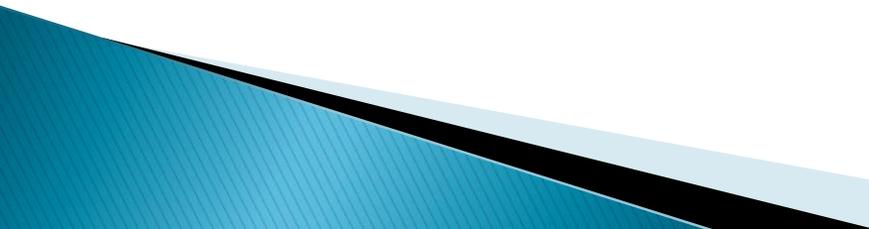
Left-Associative Grammar

(anhand von Hausser, 1988)

18.12.2013

Thorsten Beckers, Maximilian Neitzel, Kai Wittig

Agenda

- ▶ Klärung der Begriffe
 - ▶ Prinzip der möglichen Fortsetzungen
 - ▶ Analyse und Konstituenten-Strukturen
 - ▶ KS-Paradoxon und Lösungen
 - ▶ LA-Grammar im Detail mit Beispielen
 - ▶ NEWCAT-Parser
 - ▶ Motivation und Begründung für LAG durch Hausser
- 

Linksassoziativ

$$a + b + c$$

$$((a + b) + c) \quad \text{vs.} \quad (a + (b + c))$$

linksassoziativ vs. rechtsassoziativ



Inkrementelle Ableitung

linksassoziativ:

a
(a + b)
((a + b) + c)
(((a + b) + c) + d)

rechtsassoziativ:

a
(b + a)
(c + (b + a))
(d + (c + (b + a)))

Prinzip der möglichen Fortsetzungen

- ▶ Left associative-grammar
- ▶ Beginn: Erstes Wort des Satzes
- ▶ Regeln für die nächste, grammatische Komposition

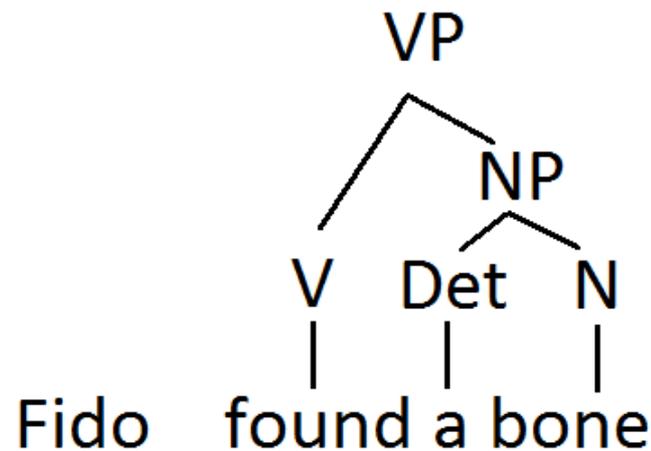
Die richtige Analyse

- ▶ Für einen geg. Satz können eine Vielzahl von verschiedenen Strukturanalysen innerhalb eines formalen Regelsystems definiert werden
 - ▶ Diese sind in der Regel semantischer Natur
 - ▶ Sog. Konstituenten Strukturen
- 

Konstituenten Struktur (1)

- ▶ Definition/Regeln:

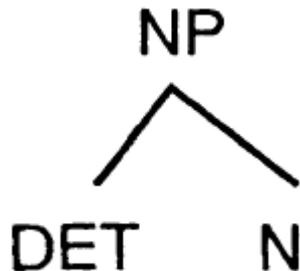
1. Wörter/Konstituenten die semantisch zusammen gehören werden von einem Knoten dominiert:



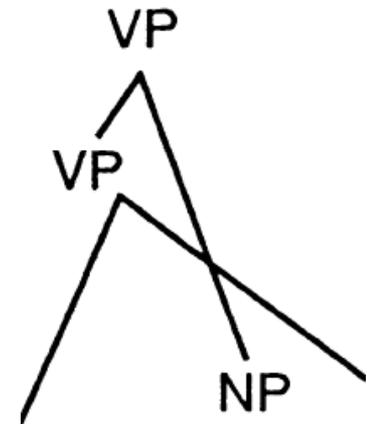
Konstituenten Struktur (2)

- Die Äste des Baumes dürfen sich nicht überschneiden („non-tangling condition“)

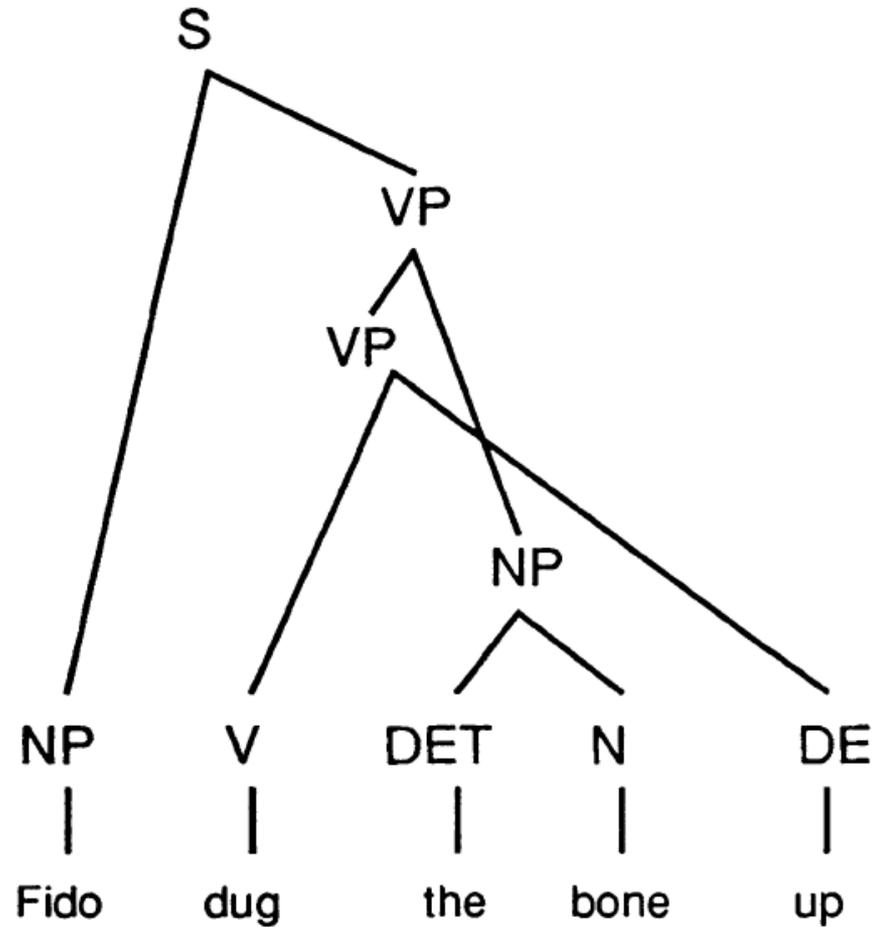
gültig



ungültig



Konstituenten-Struktur-Paradoxon

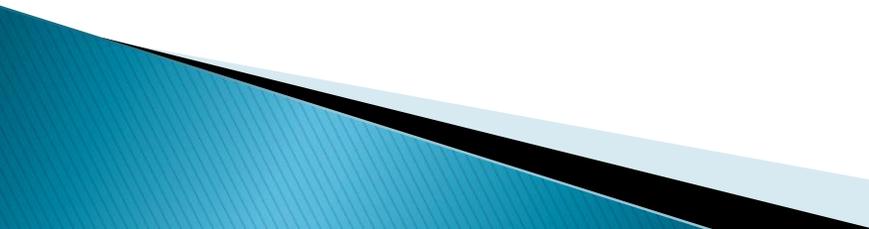


Struktur-Paradoxon-Mögliche Lösungen

1. Schwächung der ersten Regel

- a) Transformational approach
- b) Semantic feature component

2. Schwächung der zweiten Regel

- a) Tree-linking grammar (3D-Baum)
 - b) Orthogonal grammar (Oberfläche u. Grammatik getrennt)
- 

Schlussfolgerung

- ▶ Warum sollten semantische Hierarchien die KS-Regeln befolgen?
 - ▶ Semantische Hierarchie \neq Konstituenten Struktur
 - ▶ Semantische Hierarchien \rightarrow Pragmatische Interpretation
- 

C-grammar



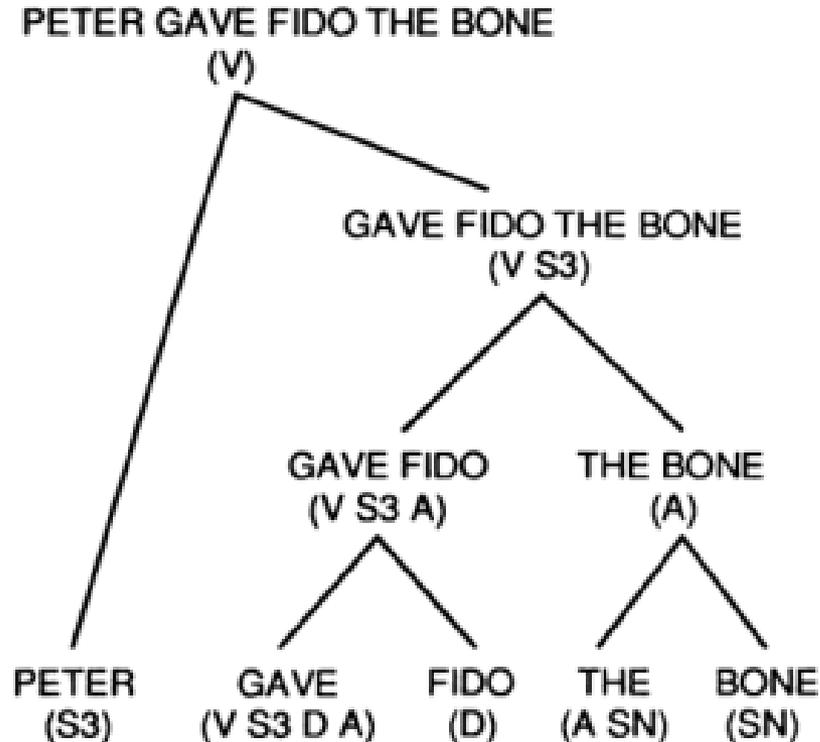
Bottom-up amalgamating

PS-grammar

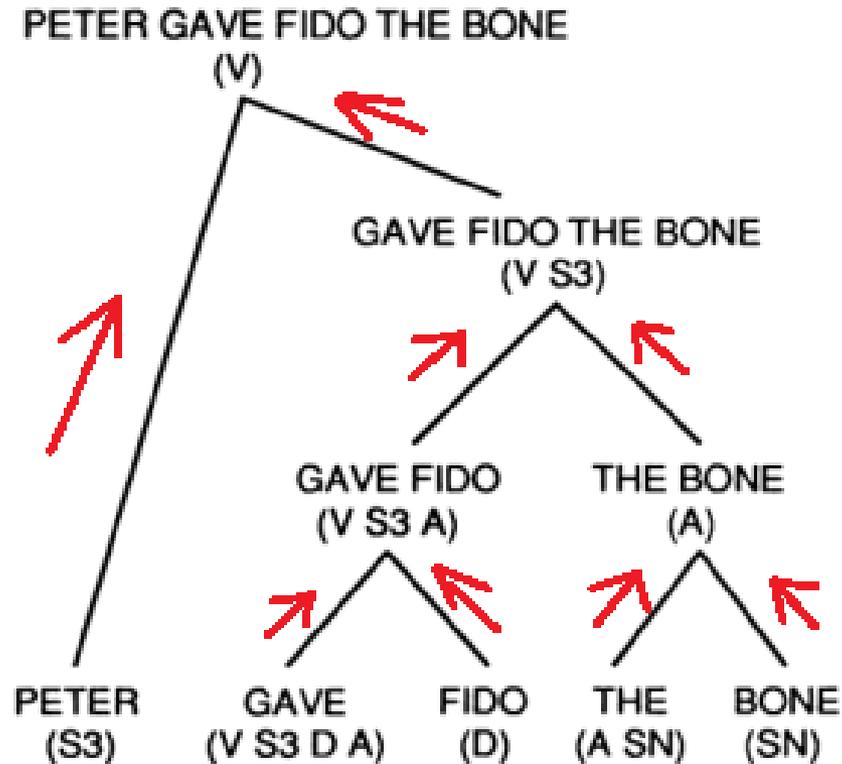


top-down expanding

Traditional analysis in Categorical Grammar

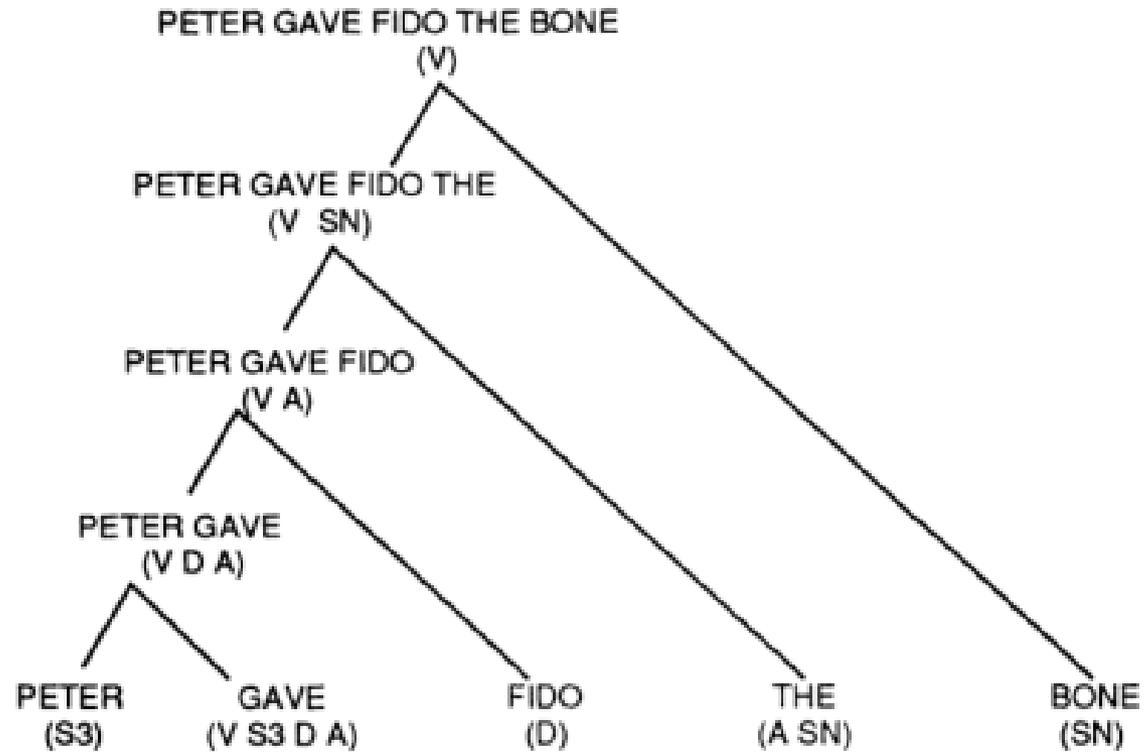


Traditional analysis in Categorical Grammar



- ▶ Kein klarer Start definierbar

LA-grammar

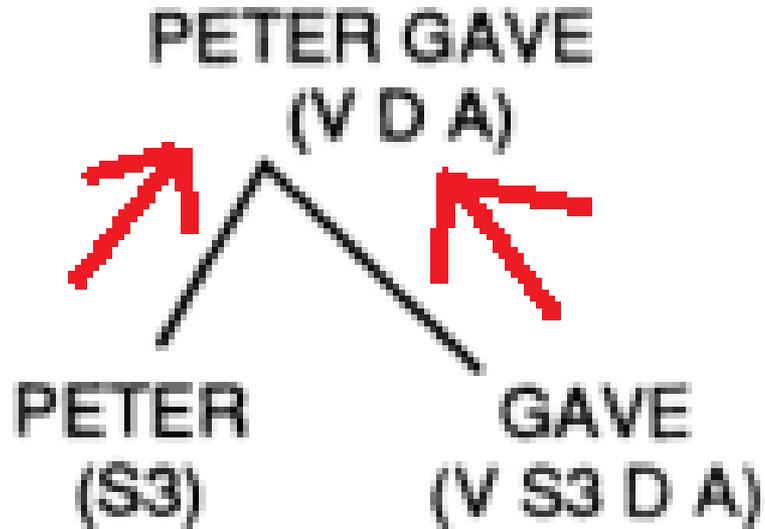


LA-grammar

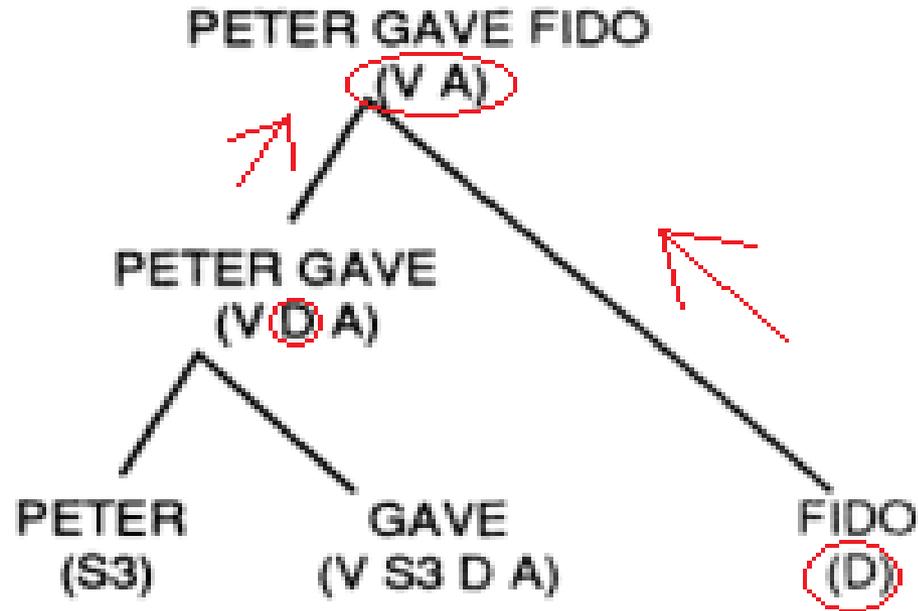
PETER
(S3)

▶ Start

LA-grammar

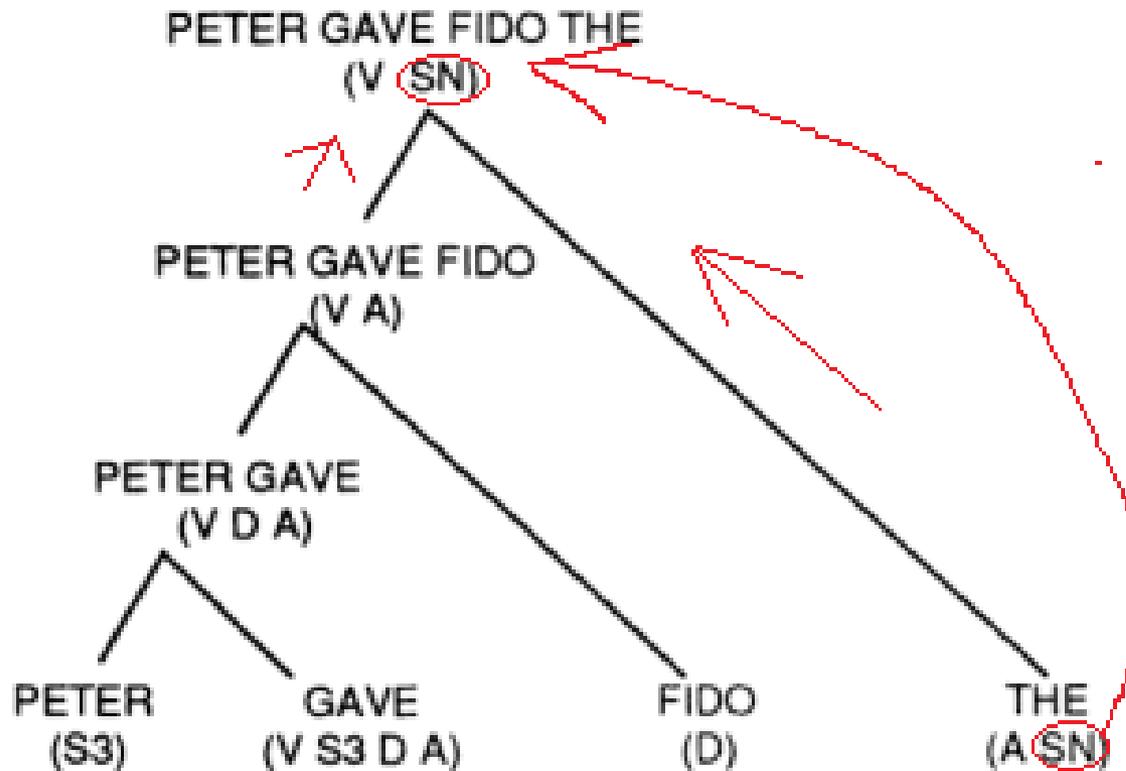


LA-grammar



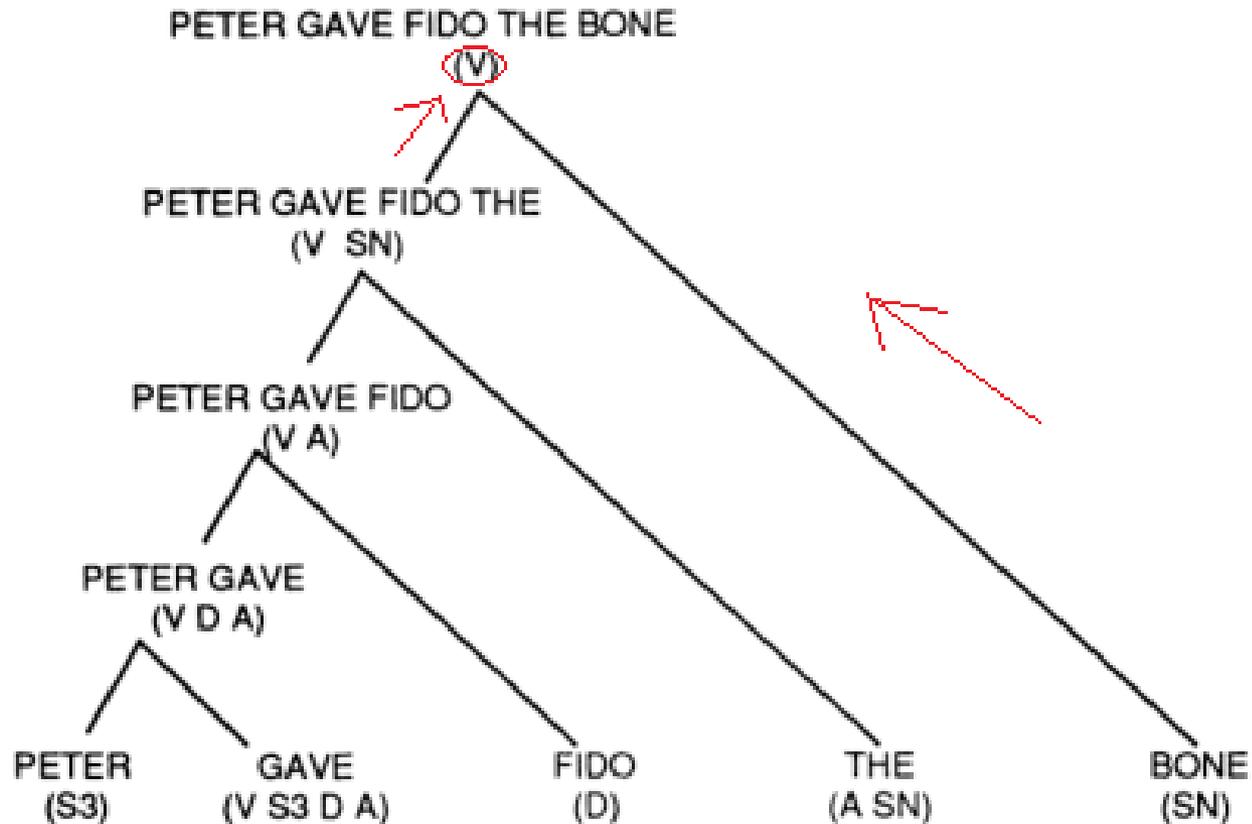
- ▶ V D A werden nach oben weiter gegeben
- ▶ D kommt hinzu
- ▶ Die beiden D heben sich auf, gilt also als erfüllte Bedingung

LA-grammar



- ▶ SN kommt als neue Bedingung hinzu

LA-grammar



- ▶ Alle Bedingungen abgearbeitet
- ▶ Endet eben so wie die traditionelle Variante zuvor auf V

NEWCAT

```
(z Peter gave Fido the bone \.)
```

```
Real time: 0.06 s
```

```
Run time: 0.06 s
```

```
Linear Analysis:
```

```
*START_0
```

```
1
```

```
(NH) PETER
```

```
(N D A V) GAVE
```

NEWCAT

```
(z Peter gave Fido the bone \.)  
Real time: 0.06 s  
Run time: 0.06 s
```

Linear Analysis:

*START_0

1

(NH) PETER

(N D A V) GAVE

*NOM+FVERB_3

2

(D A V) PETER GAVE

(N-H) FIDO

*FVERB+MAIN_4

3

(A V) PETER GAVE FIDO

(GQ) THE

*FVERB+MAIN_4

4

(GQ V) PETER GAVE FIDO THE

(S-H) BONE ?

*DET+NOUN_2

5

(V) PETER GAVE FIDO THE BONE

(V DECL) .

*CMPLT_13

6

(DECL) PETER GAVE FIDO THE BONE .

LA-Grammar

Hausers Motivation

Generell

- ▶ Zentrale Aufgabe der linguistischen Analyse in Syntax, Semantik und Pragmatik ist es, die Funktion der Sprache bei der Kommunikation zu erklären.
- ▶ „The central task of linguistic analysis in the areas of synchronic syntax, semantics, and pragmatics is to explain the functioning of language in communication.”

Lineare vs. hierarchische Strukturen

- ▶ Oberflächenstruktur natürlicher Sprache linear
 - ▶ Alle bisherigen semantischen Sprachmodelle hierarchisch
 - ▶ Es gibt nicht „die eine“ richtige Analyse
 - ▶ Jeder Satz kann in verschiedenen abstrakten Bäumen dargestellt werden
- 

Begründung:

Adäquatheit von Grammatiken

- ▶ ‚surface–meaning–mapping‘ als Teil einer allgemeineren Kommunikationstheorie
- ▶ Hörer–Modus:
 - Unanalysierte Wörter als Input
 - Semantische Repräsentation als Output
- ▶ Sprecher–Modus:
 - Semantische Repräsentation als Input
 - Wörter/Ausdrücke als Output

Criterion of Procedural Adequacy

- ▶ Die Grammatik einer natürlichen Sprache ist „prozedural“ adäquat, wenn:
 - sie explizit ist
 - bidirektional ist (für das surface-meaning-mapping)
 - entscheidbar ist

The Derivational Order Hypothesis

- ▶ Es gibt nur eine richtige Ableitungsreihenfolge (derivational order) für natürliche Sprachen:
 - „time-linear order“ – left-to-right
- ▶ Interpretation und Generierung natürlicher Sprache geschieht in einer zeitliche linearen Reihenfolge

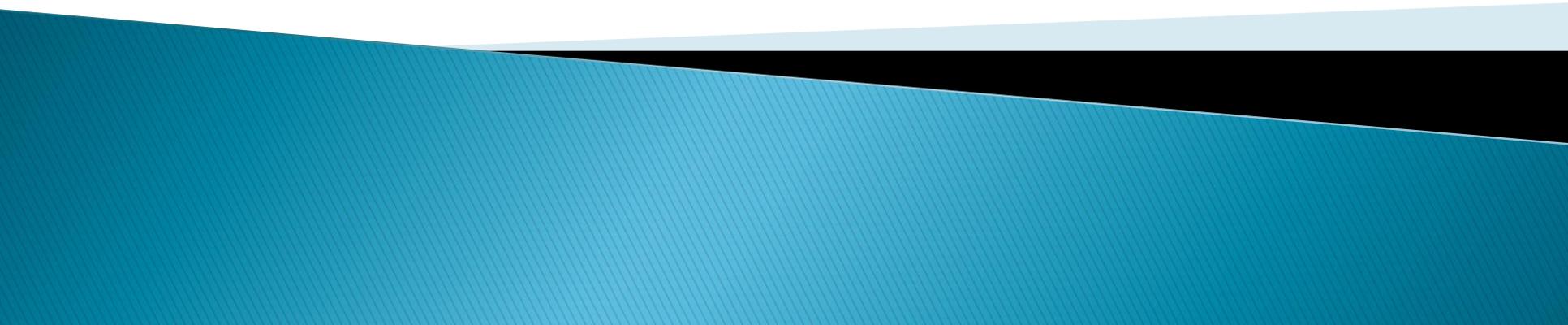
Trugschlüsse der Konstituenten-Struktur

- ▶ Hauser stellt sich gegen Chomskys Kompetenz/Performanz-Unterscheidung
- ▶ „Ausrede“ dafür, Bidirektionalität und Parsingprobleme zu ignorieren
- ▶ Parsing mit Konstituentenstrukturen: „inherently inefficient“, zu viel „trial and error“ (weil sehr nicht-deterministisch)
- ▶ LAG muss nicht viel „raten“ (weil strikt links-assoziatives Vorgehen)

„absolute type transparency“: Vorteile

- ▶ „strongly input–output equivalent“
- ▶ Vorteile der grammar rules im Parser
 - schnelleres Parsen durch bessere Theorie
 - bessere Theorie durch Verbesserungen beim Parsen
- ▶ Vorteile der Verwendung der links–assoziativen Ableitungen
 - einfach erweiterbar
 - leicht auf Fehler zu überprüfen

**Vielen Dank für Ihre
Aufmerksamkeit**



Hausser, Roland. 1988. Left-Associative Grammar: An informal outline . *Computers and Translation*, 3, 23-67.