Tree Adjoining Grammars

Overview

Laura Kallmeyer & Benjamin Burkhardt
(Slides partly by Timm Lichte and Simon Petitjean)

HHU Düsseldorf

WS 2017/2018

The general setting

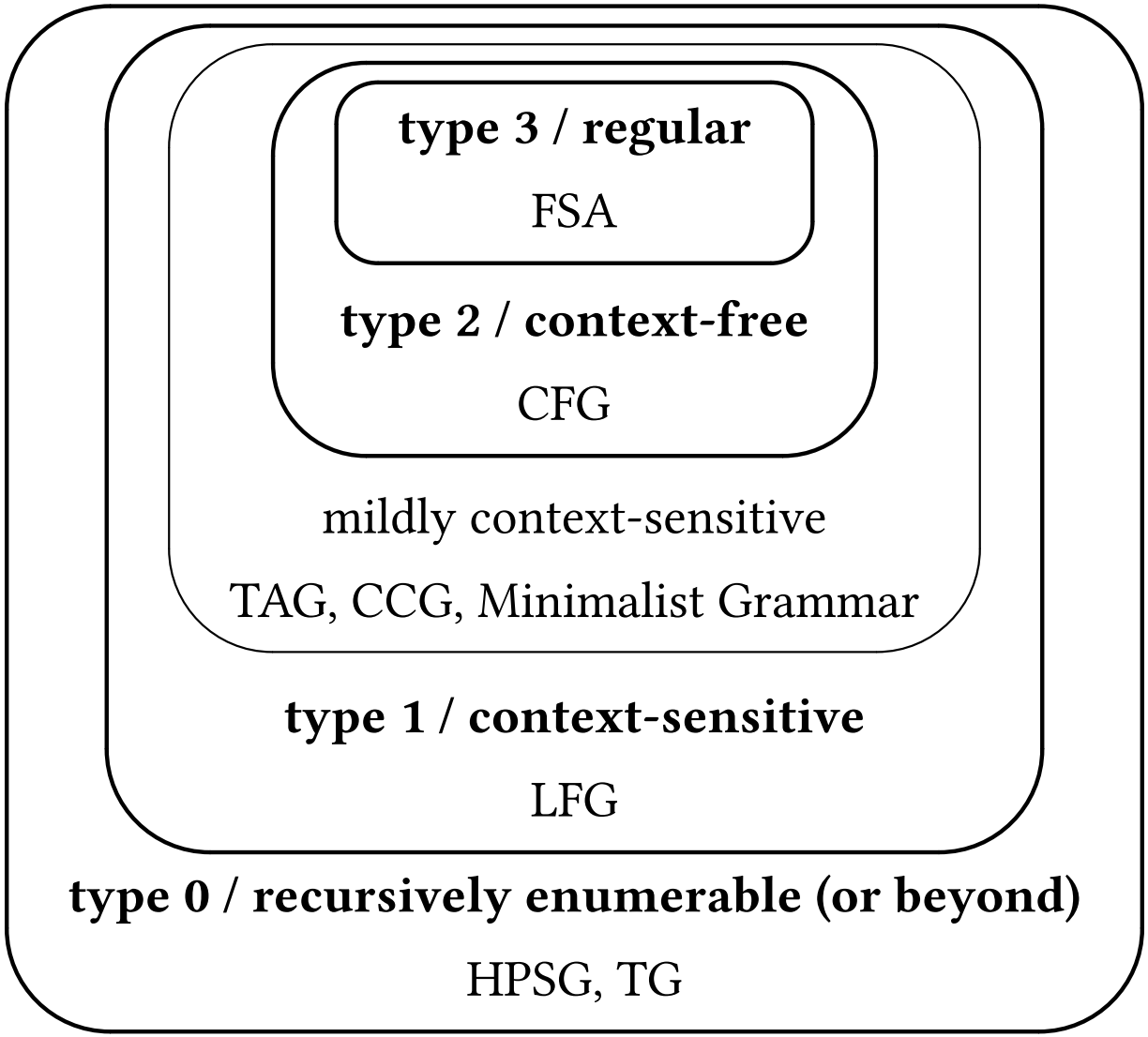
- **Grammar/linguistic theory:** rules for well-formed structures of natural language
- **Grammar formalism:** mathematically concise description language, for instance *Tree Adjoining Grammar (TAG)*
- **Implementation:** (the result of) a process to translate sth.
  - into a specific grammar formalism
  - into a specific input format for a parser
  - into …

The landscape of Grammar Formalisms (1)

- **generative rewriting formalisms:**
  - Context-Free Grammar (CFG)
  - Tree-Adjoining Grammar (TAG)
  - Lexical Functional Grammar (LFG)
  - Transformational Grammar (TG/GB), Minimalism
- **proof-theoretic formalisms:**
  - Combinatorial Categorial Grammar (CCG)
- **model-theoretic/constraint-based formalisms:**
  - Head-Driven Phrase Structure Grammar (HPSG)

The landscape of Grammar Formalisms (2)

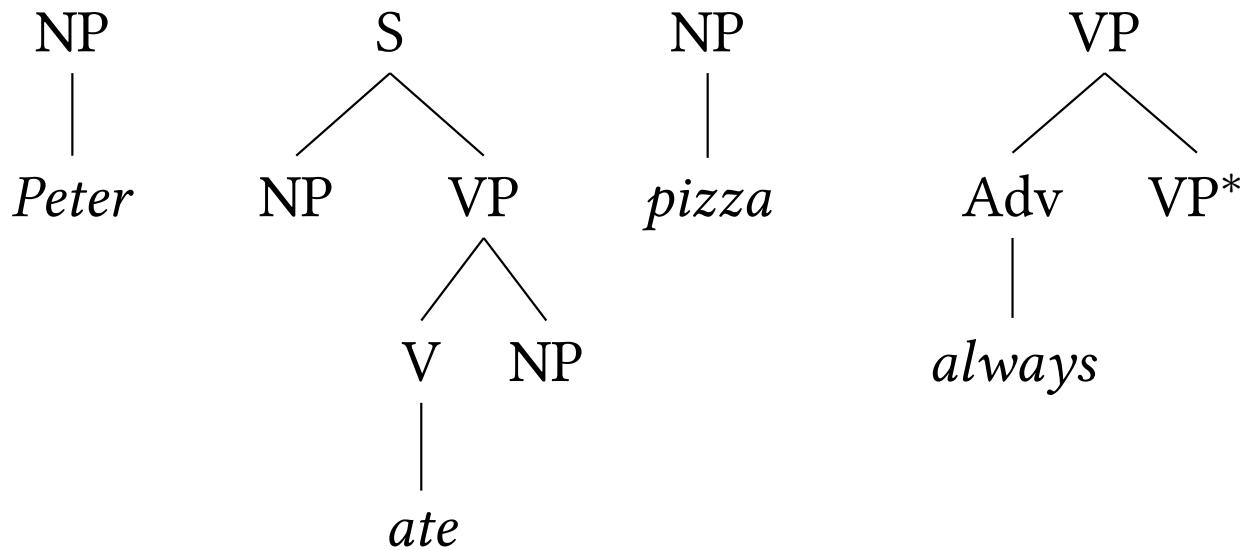
Within **Chomsky hierarchy**:



Tree-Adjoining Grammar - Basics

A **Tree Adjoining Grammar (TAG)** is a set of elementary trees:

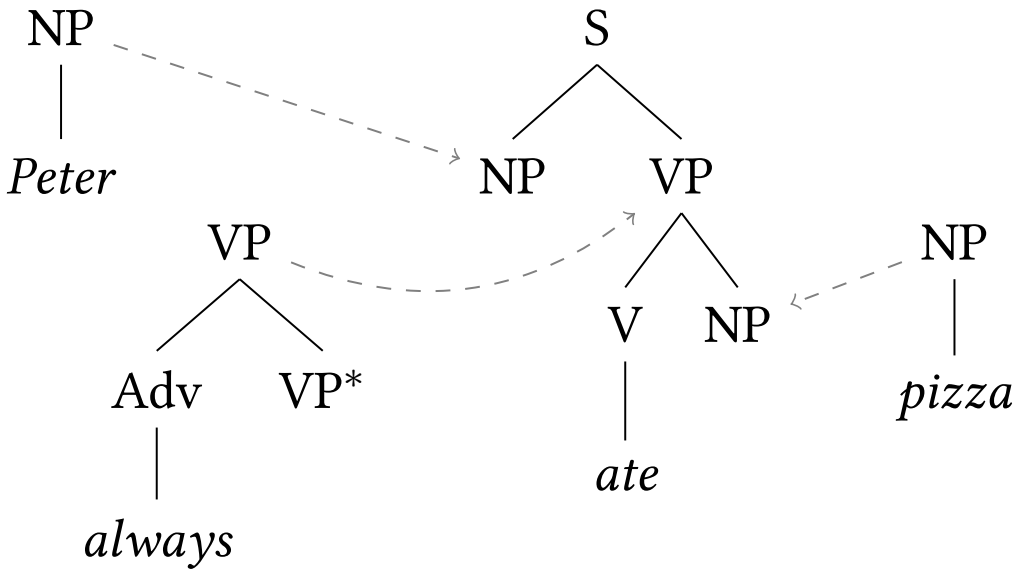
- a finite set of initial trees
- a finite set of auxiliary trees



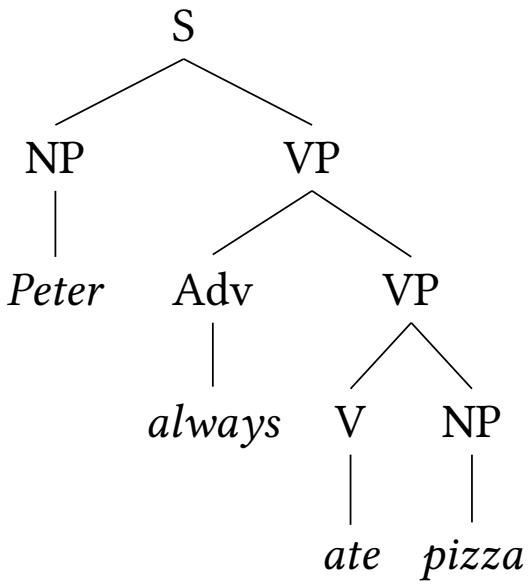
Combinatorial operations:

- substitution: replacing a non-terminal leaf with an initial tree
- adjunction: replacing an internal node with an auxiliary tree

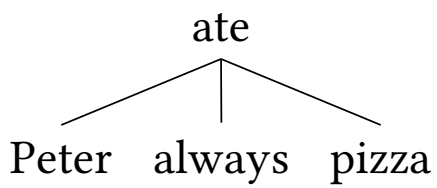
Tree-Adjoining Grammar - Example



derived tree



derivation tree



Tree-Adjoining Grammar - Basics

TAGs are **mildly context-sensitive:**

1) Polynomial time parsing complexity
2) Generation of limited crossing dependencies
3) Constant growth property (semilinearity)

Mild context-sensitivity characterizes the generative capacity needed for the analysis of natural language syntax.

Large TAG grammars:

- English and Korean (XTAG, UPenn)
- French TAG (Benoit Crabbé's PhD-thesis)
- German (GerTT)
- ...

Two ways of grammar implementation with TAG

1) **XTAG tools** (UPenn)
   - parser, editor, viewer, ...
2) **XMG + TuLiPA**
   - XMG: eXtensible MetaGrammar (Duchier et al, 2004)
   - TuLiPA: Tübingen Linguistic Parsing Architecture (Parmentier et al, 2008)

Inside and outside this lecture

- **What we are going to cover:**
  1. Grammar formalism: Tree Adjoining Grammar (TAG)
  2. Phenomena + analysis from the XTAG grammar (syntax)
  3. Implementation: (probably) XTAG tools, XMG + TuLiPA
  4. Analysis and implementation: adding semantic frames (XMG + TuLiPA)

- **What is not part of our concerns in this lecture:**
  - pragmatics, morphology, phonetics/phonology , ...
  - Head Driven Phrase Structure Grammar (HPSG), Combinatorial Categorial Grammar (CCG),
    Lexical Functional Grammar (LFG),
    Transformational Grammar (GB), Minimalism
  - corpus-driven approaches (quantitative linguistics)
  - grammar induction