Dependency Grammars: Dependency and Non-Dependency Relations

Markéta Lopatková

Institute of Formal and Applied Linguistics, MFF UK

lopatkova@ufal.mff.cuni.cz
Dependency Grammar (DG)

*notion of DG in a nutshell:*

The dependency grammar is

- a model developed by Lucien Tesnière (1893-1954) and
- based on structuralism
- to describe the syntax of natural languages.

The main concern of the dependency grammar is

- the description of the dependency structure of a sentence,
  i.e. the structure of dependency relations between the elements of a sentence.
Dependency Grammar (DG)

*notion of DG in a nutshell:*

The dependency grammar is

- a model developed by Lucien Tesnière (1893-1954) and
- based on structuralism
- to describe the syntax of natural languages.

The main concern of the dependency grammar is

- the description of the dependency structure of a sentence, i.e. the structure of dependency relations between the elements of a sentence.

- dependency as an asymmetric binary relations between language units
- governing/modified unit (head) – dependent/modifying unit (modifier)
  - word (morph) grammar … *lexicalization*
  - no phrase nodes
- dependency trees, with edges ~ *dependency relations* (mostly)
Dependency Relations

structural linguistics:
(based on Ferdinand de Saussure: Course in General Linguistics, 1916)
• synchronic approach (vs. diachronic)
• “sign” : “signified” (idea, concept) – „signifier“ (means of expressing)
• examining language as a (static) system of interconnected units
• stress on structure (signs cannot be examined in isolation)
• syntagmatic vs. paradigmatic relations
• langue (idealized abstraction of language) vs. parole (language as actually used)

structuralist schools:
• Genova School (course 1909-12): Ferdinand de Saussure, Albert Sechehaye, Charles Bally
• Prague School (1926–1939): Vilém Mathesius, Bohumil Trnka, Bohuslav Havránek, Jan Mukařovský, Roman Jakobson, Nikolai Trubeckoj, Sergej Karcevskij
• Copenhagen School (1930-1950): Louis Hjelmslev, glossematics
• “American structuralism” (1920-50): Leonard Bloomfield, Charles Hockett
Dependency-based Approaches

- Pāṇini (6th-4th century BC; India); Ibn Maḍāʾ (12th century AD; Andalusia) … term dependency
- Franz Kern and others ((† 1894, esp. pedagogy) … sentence diagrams
- Lucien Tesnière († 1954; France) … valency, “stemma” (unordered)

Motivation for current computational linguistics / NLP:
- David Hays (1950-60, machine translation ru→eng)
- Harris Zellig (since 1930, † 1992; linguistics as applied mathematics; methodology of linguistic analysis)
- Dependenzgrammatik … esp. Jürgen Kunze (from 1960s, 1975)
  Valenzgrammatik … esp. Gerhard Helbig (from 1960s)
- Richard Hudson (from 1970s, 1984) … Word Grammar
- Michael Halliday … Systemic Functional Grammar
Dependency-based Approaches (cont.)

- **Meaning-Text Theory (MTT)** … applied esp. in machine translation, lexicography; Igor Meľčuk, Aleksandr Žolkovski (1980-)

- **Functional Generative Description (FGD)** ... applied in treebanks from the Prague dependency family, used esp. for machine translation; Petr Sgall and his school (1967-)

- **Universal Dependencies (UDs)** … since 2013, Joachim Nivre et al.
Dependency Relations

types of dependencies:

- **semantic dependencies** … semantic predicates and their arguments
  cf. *Sam likes Sally.* (like(Sam, Sally))
  vs. *new car* (= car being new New(car))

- **syntactic dependencies**

- **morphological dependencies** (agreement)
  cf. *this house* vs. *these houses*
  cf. *strom je zelený* ‘house is green-sg-inan’
  vs. *stromy jsou zelené* ‘houses are green-pl-inan’
  vs. *mužíci jsou zelení* ‘men are green-pl-anim’

- **intra-word dependencies** (→ derivational morphology)

- **prosodic dependencies**
  cf. *clitic* (a syntactically autonomous unit prosodically dependent on its host)
  *H*e’ll stop. *T*here’s a problem. *P*eter’s *h*at. …
  -li; jsem, jsi …, bych, bys, …. ; se, si, mi, ti, mu, mě, tě, ho, …(tam, tu; však, tak)
Syntactic Dependency Relations

• dependency as an asymmetric binary relations between language units
  → detecting heads: not commonly agreed criteria
• number of linguistic criteria
  • e.g., verb as a syntactic center of a sentence
• BUT treebanks: annotation schemata reflect **technical considerations**
  • tree-based data format
  • 1:1 correspondence between nodes and tokens
Syntactic Dependency Relations I

possible reduction criterion ... FGD (Sgall et al., 1986), and thus PDT

- “dependent member of the pair may be deleted
  while the distributional properties are preserved”
  (→ correctness is preserved)

- endocentric constructions

  e.g.  malý stůl → stůl  
  přišel včas → přišel  
  (přišel) velmi brzo → (přišel) brzo

  small table → table  
  he came in time → he came  
  (he came) very soon → (he came) soon
possible reduction criterion … FGD (Sgall et al., 1986), and thus PDT

- “dependent member of the pair may be deleted
  while the distributional properties are preserved”
  (→ correctness is preserved)

- endocentric constructions

- exocentric constructions … principle of analogy (delexicalization)

  *Prši.* [(It) rains.] … ∃ subjectless verbs
  ⇒ *Král zemřel.* [The king died.] … a verb rather than a noun is the head

  *The girl painted a bag.* → *The girl painted.* … ∃ objectless verbs
  ⇒ *The girl carried a bag* … an object is considered as depending on a verb

- plus technical considerations (compare also with “the school grammar”)
  e.g.: prepositions are below nouns;
  auxiliary verbs are (typically) below content verbs
Syntactic Dependency Relations II

**constituent-based** criteria (Osborne, 2019)

- each complete subtree must be a “constituent”, based of formal tests, esp:
  - topicalization
  - clefting and pseudoclefting
  - proform substitution (replacement)
  - answer fragments
  - coordination

**Topicalization:**

… but *tennis* Fred did play this spring.  

This spring Fred played tennis.

**Clefting:**

*It was* *Fred* who played tennis this spring.  

*It was* *tennis* that Fred played this spring.  

*It was* *this spring* that Fred played tennis.
Syntactic Dependency Relations II

**constituent-based** criteria (Osborne, 2019)

- BUT: applied also for (more-or-less) technical solutions

Mary *will eat* bread.

Topicalization:
… and *eat* Mary certainly will.

Proform substitution:
Mary will do so. (do=eat)

Answer fragment:
What will Mary do? Eat.

VP-ellipsis:
Peter will eat and Mary will, too.

⇒ lexical verb should be a dependent
Syntactic Dependency Relations III

criterion of \textit{maximal parallelism} between languages

... Universal Dependencies

\begin{tabular}{llllllll}
\textbf{English:} & \textbf{Bulgarian:} & \textbf{Czech:} & \textbf{Swedish:} \\
\end{tabular}
Syntactic Dependency Relations III

criterion of \textit{maximal parallelism} between languages

... Universal Dependencies

• the upper levels of UD trees should be as similar as possible across languages
  - dependency relations hold primarily between content words
    (rather than being indirect relations mediated by function words)
Syntactic Dependency Relations III

criterion of \textit{maximal parallelism} between languages

- the upper levels of UD trees should be as similar as possible across languages
  - dependency relations hold primarily between content words
    (rather than being indirect relations mediated by function words)
  - function words attach as direct dependents of the most closely related content word
Syntactic Dependency Relations III

criterion of *maximal parallelism* between languages

... Universal Dependencies

- the upper levels of UD trees should be as similar as possible across languages
  - dependency relations hold primarily between content words (rather than being indirect relations mediated by function words)
  - function words attach as direct dependents of the most closely related content word
  - punctuation attach to head of phrase or clause

Dependency Grammars and Treebanks – Dependency and Non-dependency
Non-Dependency Relations

**coordination** … "multiplication" of a single syntactic position

- **different** referents
- coordination of sentence members / sentences

> My sister Mary and John came late.
> Mary came in time but John was late.
> I can't leave since it hasn't stopped raining yet.
> Nemohu odejít, neboť ještě nepřestalo pršet.

- **coordination may be embedded**

  > nice and romantic towers and castles
  > krásné a romantické hrady a zámky
Non-Dependency Relations

**coordination** … "multiplication" of a single syntactic position
- different referents
- coordination of sentence members / sentences
  
  *My sister Mary and John came late.*
  *Mary came in time but John was late.*
  *I can't leave since it hasn't stopped raining yet.*
  *Nemohu odejít, neboť ještě nepřestalo pršet.*

- coordination may be embedded
  *nice and romantic towers and castles*
  *krásné a romantické hrady a zámky*

**apposition** … "multiplication" of a single syntactic position
- identical referent
  *
  *Charles IV, Holy Roman Emperor*
  *The Hobbit, or There and Back Again*
Non-Dependency Relations

**coordination** … "multiplication" of a single syntactic position
- different referents
- coordination of sentence members / sentences
- coordination may be embedded

**apposition** … "multiplication" of a single syntactic position
- identical referent

cannot be represented by dependency edges necessary to enrich the data structure
Coordination in Dependency Trees I

Mel'čuk (1988):

• ‘grouping’ (G) … treating the first conjunct as the head

• problem:
  shared modification
  vs. modification of a single member

Hubení ((mladí muži), vojáci a starci)
[Thin young men, soldiers and old-men]
Coordination in Dependency Trees II

*Universal Dependencies*: version 2 (2016):

- the **first conjunct** ~ the head of all following conjuncts

- attach coordinating conjunctions and punctuation to the immediately succeeding conjunct

Diagram:

```
I love apples and bananas
```

```
We have apples, pears, oranges, and bananas
```
Universal Dependencies:
version 2 (2016):

• the \textit{first conjunct} ~ the head of all following conjuncts

• attach coordinating conjunctions and punctuation to the immediately succeeding conjunct

• left-headed … e.g., \textit{Mary won gold and Sue bronze.}
BUT: right-headed constructions
e.g., \textit{one green and two red cars}
green as a (promoted) head (and \textit{cars} as dependent)
Petkevič (1995) … formal representation of FGD two types of brackets for tree linearization:
• 〈 〉 for dependencies
• [ ] for coordination

John and Mary, who live in Boston, are good people.
PDT 2.0:

*connecting' constructions* ~ coordination, apposition (, OPER)

specific types of nodes and edges:

- **connecting node** = node for coordinating / appositing conjunction

- **effective parent** = node for governing node, i.e. node modified by the whole construction, 'linguistic parent'

- **members of a connecting construction** = nodes that are coordinated / are in apposition
  - `is_member`

- **effective child(ren)** … modification(s) of the individual member of the connecting construction + common/shared modifier(s)

- ‘pass-through’ nodes
The center will gather and distribute the information on tenders and state commissions in this country as well as in abroad.

Centrum bude shromažďovat a distribuovat informace o tendrech a státních zakázkách doma i v zahraničí.
Coordination in Dependency Trees IV

PDT 2.0:
- embedded connecting constructions \rightarrow recursivity

- **TrEd** (Tree Editor, Pajas):
  functions GetEChildren, GetEParents
References


• Universal Dependencies  [https://universaldependencies.org](https://universaldependencies.org)

• Prague Dependency Treebank  [http://ufal.mff.cuni.cz/pdt3.5](http://ufal.mff.cuni.cz/pdt3.5)
Dependency and non-dependency relations

other non-dependency relations in PDT:
• technical root – effective root of a sentence
• syntactically unclear expressions
  rhematizers; sentence, linking and modal adverbial expressions, conjunction modifiers
• list structures
  names, foreign expressions
• phrasemes