Formal & Computational Aspects of

Dependency Grammar

– That depends: Pros & Cons of DG –

Geert-Jan M. Kruijff

Computational linguistics
University of the Saarland
Saarbrücken Germany
⟨gj@coli.uni-sb.de⟩

July 25, 2002

Goal & Overview

- **Goal**:
  1. To discuss the potential advantages and disadvantages of DG.
  2. To review how different DG frameworks have (or have not) tried to exploit the advantages, and overcome possible disadvantages.
- **The big question**: How can we overcome disadvantages and turn them into advantages?

Goal & Overview: Pros and cons at a glance

- **Deep dependency relations**
  - From arbitrary syntax/semantics interface to cross-linguistic generalizations about the realization of meaning
  - Flat representations
  - Interpretative import
- **Word order**
  - From insufficient linearization principles to dependency, flexible constituency, long-distance dependencies, and domains
- **Coordination**
  - From problems with coordination having no head/dependent asymmetry to two dimensions of coordination: constituency, and dependency.
- **Realization of information structure**
  - Necessity of deep relations for explaining word order as realization of IS
  - Usefulness of flexible constituency for intonational phrasing

Contents

Deep dependency relations

- A criticism levelled against "deep" dependency relations (θ-roles,...) like **Actor**, **Effect**, etc. is that the relation between such roles and their realization is arbitrary. But,
  - We can give a typological account of how these dependency relations are realization, using an *abstract theory of case*. This leads to cross-linguistic generalizations about the realization of valency, something that phrase-structure grammar is not capable of.
  - Furthermore, dependency relations can be defined as having *interpretative import*, which not only gives a more detailed description of linguistic meaning, but also helps explaining coercion effects.
- The kinds of structures we build using dependency relations are very *flat*, because heads and dependents are related *immediately* without any intervening nonterminal nodes.

Typological theories of case

- The typological distinction between *morphological categories* and *morphological strategies*:
  - A category can be realized by different strategies, within and across languages; (Croft, 1990).
  - Government & Binding's theory of abstract case, cf. (Haegeman, 1991).
- Examples of morphological strategies (Croft, 1990; Kruijff, 2001):
  - **Case**: The use of bound morphemes or *case markers* indicates the morphological category. E.g. *knih*-**u** (Czech) or *Kind*-**es** (German)
  - **Adposition**: A *function word* affixed to the wordform signals the morphological category. E.g. **aan** *Kathy* (Dutch)
  - **Positioning**: The wordform's position in the clause (usually, relative to the finite verb), indicates the morphological category.
  - **Linker**: Invariant marker that relates the modifer and the modified. E.g. **of** *Kathy, Kathy*'**s** (English).

Kuryłowicz's theory of case

- To relate functional relations and their realization, Kuryłowicz (1964) proposes to distinguish *primary* from *secondary functions* of morphological categories:
  - The primary function of a morphological category is the functional relation it realizes using an unmarked morphological strategy.
  - The secondary function of a morphological category is the functional relation it realizes using a more unmarked morphological strategy.
- Consider the abstract Accusative:

				Czech	**German**
Patient	$\overset{primary}{\longleftarrow}$	*abstract Accusative*	$\overset{unmarked}{\longrightarrow}$	Acc. case	Acc. case
Dir:WhereTo	$\overset{secondary}{\longleftarrow}$		$\overset{marked}{\longrightarrow}$	*na*+Acc	*auf*+Acc

- As (Sgall et al., 1996)(p.72) note, the same is observable for Sanskrit.

Cross-linguistic generalization

- **Universality**: Morphological strategies are language-specific, and can be ordered by markedness. Morphological categories are assumed to be *universals* (universal interfaces), as are functional relations.
- Hence, we can make predictions about how valency frames will be realized in a particular language, given that language's inventory of morphological strategies (Skalička and Sgall, 1994; Kruijff, 2001). Constituency cannot do this!
- This perspective has a long tradition in the Prague School of Linguistics, building on work by Jakobson, Mathesius, Skalička, Trnka, later influencing Daneš, Dokulil, Kuryłowicz, and Sgall and his collaborators.

Interpretative import of dependency relations

- When we interpret wordforms in a sentence, either as heads or as dependents modifying heads along specific kinds of relations, we are not just building a formal structure.
- By interpreting a wordform as a particular kind of dependent, we specify *how* the meaning underlying that wordform contributes to the meaning of the head – and as such, the linguistic meaning of the entire expression.
- We call the specification of this "how" the **interpretative import** of a dependency relation (Kruijff, 2001).

Classes of dependency relations

- We can distinguish different *classes* of dependency relations. These classes arise from regularities in the interpretative imports of (similar) dependency relations (Kruijff, 2001).
- **Dependency relations that relate eventualities in a *causative way***:

  They provide more detail about the underlying causo-temporal structure of an event (giving rise to, preparation, consequence).

  (1) a. **Purpose**: Elijah studies intensively (in order to pass the exam).
  b. **Result**: Kathy lowered the fire (so the food would not burn).
  c. **Cause**: The tracks were blocked (because of sheep on the rails).

  (2) Other relations are Condition, Factual/Counterfactual, and Concession.

Classes of dependency relations (cont'd)

- **Dependency relations reflecting the *stage* of the event**

  These dependency relations specify aspects of time, direction, and location.

  (3) a. **Time:Since When**: Kathy has been working (since last night).
  b. **Time:Till When**: Christopher was in New York (till yesterday).
  c. **Direction:Where To**: Elijah intends to go (to Texas).
  d. **Locative**: (In Amarillo), Christopher and Elijah met (at the local *Starbucks*).

  (4) Other relations include Time:When, Time:HowLong, Time:Contemporary, Dir:FromWhere, Dir:Through.

Classes of dependency relations (cont'd)

- **Dependency relations that attribute qualities to objects.**

  These dependency relations mostly modify nominal heads.

  (5) a. **Partitive**: Elijah bought a bunch (of flowers) for Kathy.
  b. **Appurtenance**: Christoper is (Elijah's) brother.
  c. **Identity**: Christopher and Elijah have both visited the city (of New York).

  (6) Other relations include General Relation, and Descriptive Relation.

Classes of dependency relations (cont'd)

- **Dependency relations that express *modus operandi*.**

  These dependency relations describe the way in which the process denoted by the verbal head happens, by relating it to an external event or object.

  (7) a. **Norm**: Kathy drives (in accordance with the traffic rules).
  b. **Means**: Christopher arrived (by bike).
  c. **Comparison**: Elijah's cowboy-boots are (as red as) herrings.

  (8) Other relations include Criterion, Attitude, Restriction, Accompaniment, Substitution, Regard, and Difference.

Classes of dependency relations (cont'd)

- **Inner participants.**

(9) a. **Actor**: (Elijah) slipped off a cliff.
b. **Objective**: Because he fell asleep, Elijah burnt (the potatoes).
c. **Addressee**: As its official chairperson, Kathy addressed (the committee).
d. **Effect**: Elijah dug (a hole) with a hoe.
e. **Origin**: Christopher cut a horse (out of wood).

Formalizing interpretative import of dependency relations

- We assume an account of verbal tense and aspect on the basis of (Moens and Steedman, 1988; Steedman, 2000b), using an *event nucleus*.
- **Event nucleus**
  - The event is related to a Preparation, an *activity* that brings the event about.
  - It is also related to a Consequent, connecting the event to an ensuing *state*.
  - The event itself is a complex *achievement*.
  - The entire event nucleus is an *accomplishment*.
- We can specify the interpretative import of various dependency relations in terms of their influence on the event nucleus.

- For example a **Result** states more about the consequent state of the event nucleus.
- **Temporal coercion**:

  Modifying a present-tensed verb by a temporal dependency relation that specifies a future point in time, the clause is *not* to be interpreted on the present but on a future timepoint.

  (10) *Kathy ist in Paris.*
  Kathy is in Paris, right now.

  (11) *Kathy ist in Paris* ***morgen****.*
  Kathy will be in Paris tomorrow.

- **Aspectual coercion**:

  The aspectual category of the overall clause is no longer determined by the verbal head.

  (12) Activity+**Time:Till When** → Accomplishment.
  - a. **Activity**: Kathy walked.
  - b. **Accomplishment**: Kathy walked till down.

  (13) Achievement+**Time:How Long** → Activity.
  - a. **Achievement**: The visitors arrived.
  - b. **Activity**: The visitors arrived through the night.

Formalizing interpretative import of dependency relations

- To formalize the interpretative import, we need a model of the event nucleus and its interpretation.
- Because an event nucleus has a *spatial*, a *temporal* and a *causal* extension, we need a model that
  - represents each of these dimensions separately, and that
  - can relate each of these dimensions.
- A formal characterization is given in (Kruijff, 2001), making use of hybrid logic (Blackburn, 1994; Blackburn, 2000).

Coordination

- A coordination is a construction that is hard to describe in terms of dependency.

  (14) Christopher and Kathy went camping in the Rockies.

  (15) Christopher cooked and might have eaten the beans.

- For example, in (14), what is the head?

  1. It cannot be any of the individual nouns, because none has a higher priority than the other.
  2. It cannot be both the individual nouns, because that is incompatible with the notion of dependency.
  3. It cannot be the connective "and" because this cannot be the subject (e.g. it would never be inflected, etc.).

- So, what do we do? There are two possible answers.

Two possible solutions

1. Coordination structures do have heads (Mel'čuk, 1988).

   a. In the majority of cases coordination structures are not symmetrical (i.e. their order cannot be reversed).

      (16) Christopher stood up and passed me the beans
      $\neq$ Christopher passed me the beans and stood up.

      (17) Go to bed, or I'll spank you! (Mel'čuk, 1988)(p.27)
      $\neq$ I'll spank you, or go to bed!

      (18) not only a good worker but also a nice man (Mel'čuk, 1988)(p.27)
      $\neq$ not only a nice man but also a good worker

   b. the right conjunct (and the conjunction) is always omissible, whereas the left-hand conjunct is not.

      (19) John, but not Mary, came. $\Rightarrow$ John came.
      $\not\Rightarrow$ * But not Mary came.

Two possible solutions (cont'd)

2. Coordination structures are *not* dependency structures – they require a different treatment.

   a. Tesnière already distinguished coordination structures from dependency structures (Tesnière, 1959)(p.80).

   b. (Hudson, 1990): A coordination is a word string, i.e. a continuous string of words held together by the *Dependency in Coordination Principle*:

      "The conjuncts of a coordination must share the same dependencies to words outside the coordination" (DiCP); cf. (Hudson, 1990)(p.220ff)

      (20) ((S:e,Ann), (S:e,Betty) and (S:e,Chloe)) had(e) (O:e,supper).

Two possible solutions (cont'd)

c. (Plátek et al., 1984; Sgall et al., 1986; Petkevič, 1995): Coordination adds a new dimension to a dependency structure, that can be rendered in a linearized notation using special brackets. A principle identical to Hudson's (DiCP) is adopted.

(21) Lapierre and Collins wrote and issued a nice book.
$< [(Lapierre, N, T); (Collins, N, T)]COP > 3$
$[(write, V, F); (issue, V, F)]COP$
$<< (nice, ADJ, t) > 36(book, N, F) > 30$

Observations

- Although both Word Grammar and FGD adhere to the same the underlying principle, Word Grammar represents the dependency relation with the dependent, whereas FGD represents the relation with the conjunction node.
- Distinguishing a coordination structure *as such* from the conjuncts makes it in principle possible to account for both collective and disjunctive readings.

(22) Christopher and Kathy went swimming.

a. Collective reading: Christopher and Kathy went swimming together.

b. Disjunctive reading: Christopher went swimming, and Kathy went swimming, but not together.

Observations (cont'd)

- The DiCP makes it impossible to deal with unlike-dependent coordinations (23):

  (23) Při zvýšené opatrnosti v jízě a pro dobrý stav komunikací se nemůže nic stát.

  "With an increased cautiousness in driving and because of the good state of the roads nothing can happen." (Panevová, 1974)(p.14)

  To account for this type of coordination we have to abandon the DiCP.

Two dimensions in coordination?

- When considering coordination from a constituency-based perspective, we can distinguish *constituent* (24) from *non-constituent* coordination (25), cf. also (Sag et al., 1985):

  (24) Christopher cut the onions and fried the beans.

  (25) Christopher is a Democrat and proud of it.

- The differentiation here concerns whether or not the conjuncts form standard phrase-structure constituents.

- As we saw above, we can also consider whether the conjuncts are the same kind of dependent, or are unlike. This cuts across the constituency, as unlike dependents may still be like constituents (26,27) though need not be (28):

  (26) We plunged from the rock and into the sea.

  (27) I'll be in Prague before Xmas, during Xmas, and after Xmas.

  (28) I'll do it because of you and to have some peace at last.

Two dimensions in coordination? (cont'd)

- What does the above data mean for existing accounts of coordination?
  - Most frameworks can handle non-constituent coordinations.
  - However, because dependency relations are not distinguished, ungrammatical combinations of unlike dependency relations cannot be ruled out.
- How can we handle examples like (26) and (28)?
- Panevová explains the well-formedness of (some) unlike-dependent coordinations in terms of their meaning, i.e. their interpretative import.

  If two (unlike) dependents belong to the same class, they can be coordinated.

Handling unlike-dependent coordination?

- Consider again (26):

  (29) We plunged from the rock and into the sea.

- Assume the following informal characterizations of lexical entries:

  - *from* takes a noun to form a **Direction:From Where**, and *into* takes a noun to form a **Direction:Where To**.
  - *plunged* takes an **Actor** as its subject, and a dependent of type **Direction**.
  - **Direction:From Where** and **Direction:Where To** are both **Direction** relations.
  - *and* can coordinate unlike dependents if they are of the same class.

Handling unlike-dependent coordination? (cont'd)

- Informal analysis:

  1. *from the rock* → **Direction:From Where** → **Direction**
  2. *into the sea* → **Direction:Where To** → **Direction**
  3. *from the rock and into the sea* → **Direction**
  4. *we* → **Actor**
  5. *we plunged from the rock and into the sea* → sentence

- A formal analysis has been constructed in Dependency Grammar Logic.

Two dimensions in coordination? (cont'd)

- Do coordinations of unlike dependents always allow for both collective and distributive readings?
- Do coordinations of unlike dependents occur with multiples, alike (30)?

  (30) Christopher gave (Elijah a cowboy-hat) and (Kathy flowers).

- We leave these question for those who want to write term papers ...

Realization of information structure

- What is **information structure**?

  Information structure is an inherent aspect of an expression's linguistic meaning, with which a speaker can indicate how some parts of that meaning are *context-dependent* ("given") and others *context-affecting* ("new").

- How does information structure get realized?

  - *Structural indications of informativity*: Word order, syntactic structures (e.g. *it-clefts*), morphology, intonation are all means that languages use to express the informativity of a bit of meaning.
  - A language may use several of these means to realize information structure, possibly in combination (e.g. word order and intonation in languages like German or Dutch).

- Examples ... and then, why Dependency Grammar is necessary to explain them. At the end, we look in more detail at word order, dependency, and information structure.

Realizing IS in free word order languages

Czech
Elijah koupil knihu. ("Elijah bought a BOOK.")
Knihu koupil Elijah. ("As for the book, ELIJAH bought it.")
Elijah knihu koupil. ("Elijah BOUGHT the book.")

Turkish (Hoffman, 1995)
Esra kitab-ı okuyor. ("Esra is reading the book.")
Kitab-ı Esra okuyor. ("As for the book, it is Esra who is reading it.")

Hungarian (Vallduví and Engdahl, 1996)
Mari JÁNOST látta. ("Mary saw JOHN.")
Mari látta Jánost. (MARY saw John.")

- Use free word order to freely place any element in the canonical focus position, and project the focus. Focus projection depends on canonical ordering of elements: it only projects over elements that are in canonical ordering, not over scrambled elements.

Realizing IS in rigid & mixed word order languages

Dutch
... dat Jan Marie het boek heeft gegeven.
... dat Jan het boek aan Marie heeft gegeven.
... *dat het boek aan Marie Jan heeft gegeven.

English (Sgall et al., 1986)
John flew from Paris to London on Tuesday.
On Tuesday John flew from Paris to London.
On Tuesday John flew to London from Paris.

- Even in rigid word order languages there is a degree of variability, depending on the type of modifier.
- Variability in rigid and mixed word order languages is restricted by well-formedness; markedness may need to be obtained through other means like intonation.

Issues in the realization of IS

- There are several reasons for why we would need a dependency-based explanation of the role of word order in realizing IS.
- The head/dependent-asymmetry is crucial for a cross-linguistic account of basic word order (Greenberg, 1966; Hawkins, 1983) and its variability (Steele, 1978; Kruijff, 2001).
- 'Deep' dependency relations are also needed, for two reasons:

    1. They are useful in explaining why *variation* of word order realizes different information structure, (Sgall et al., 1986), and how word order interacts with intonation in realizing IS (Kruijff, 2001).
    2. They are in fact necessary to explain the realization of IS in languages like Tagalog (Kroeger, 1993).

Issues in the realization of IS (cont'd)

- Similarly, there are several reasons for the usefulness of dependency in explaining the role of intonation in realizing IS.
- Analyzing sentential *structure* in terms of heads and dependents corresponds to a very *flexible notion of constituency*. Flexible constituency is needed for modelling the relation between syntactic structure and intonational phrasing, cf. Steedman (2000a; 2000c).
- Dependency relations are useful in explaining how word order interacts with intonation in realizing IS, e.g. in languages like German or Dutch (Kruijff, 2001).

Word order and systemic ordering

- How *does* word order act as a structural indication of informativity? Why is it that we perceive of a variation to be attributing a different informativity to constituents?
- *Hypothesis*: Differentiation in the order in which ("deep") dependency relation are realized.
  - Languages are attributed a canonical order in which dependency relations (*Actor, Patient*,...) are realized.
  - Focus elements are realized in canonical order (projection!).
  - Variation on that order may indicate a change in informativity.
- This ordering is called *systemic ordering* (Sgall et al., 1986).

What is systemic ordering?

- What is systemic ordering? Systemic ordering is a language-specific order over argument roles that holds 'universally' within the given language.
- Examples of systemic ordering:

  Czech, Russian
  Actor < Temporal < Purpose < Spacelocative < Means < Addressee < Patient < Source < Destination

  English
  Actor < Addressee < Patient < Spacelocative < Temporal < Means < Source < Destination < Purpose

  German
  Actor < Temporal < Spacelocative < Means < Addressee < Patient < Source < Destination < Purpose

- But, is there such a thing as systemic ordering? Experimentally determined and verified – (Sgall et al., 1995).

The need for dependency relations

- Do we really need systemic ordering? Or, to put it differently, do we really need *dependency relations*? We already have canonical surface word order in terms of grammatical relations ...
- Dependency relations are an aspect of linguistic meaning, as is information structure. Compare:

  (31) (Elijah)$_{subj}$ bought (the red cowboy-boots)$_{obj}$.
  (*Actor* < *Patient*)

  (32) (The red cowboy-boots)$_{subj}$ were bought (by Elijah)$_{obj}$.
  (*Patient* < *Actor*)

- Dependency relations are **necessary** to explain well-formed variability in word order – e.g. Tagalog (Kroeger, 1993).

Systemic ordering as interface

- **Proposal**: Use systemic ordering as an abstract *interface* between information structure and its realization.
  - Consider two distinctions: An constituent is realized in systemic ordering (so), or it is not (ns). This abstracts away from a language-specific systemic ordering.
  - Define well-formedness of linearization (using grammatical relations, etc) as well as tune relative to tree-structures constructed with so, ns edges. This enables interaction between word order and tune as structural indications of informativity.
- This is worked out in the formalization presented in (Kruijff, 2001).

In conclusion

- Issues that initially were presented as *problems* of dependency grammar, could be addressed and turned into *advantages* of dependency grammar.
- For some of these issues, in fact, we argued that *only* a *dependency-based* account can give an adequate account, e.g.:
  - Cross-linguistic account of how linguistic meaning (dependency relations) can get realized.
  - Cross-linguistic account of how information structure can get realized.
  - Interpretative import of dependency relations – modifications that e.g. give rise to temporal, aspectual coercion.
  - Dependency as another dimension to coordination, next to constituency.
- So, the saga continues ...

References

Patrick Blackburn. 1994. Tense, temporal reference and tense logic. *Journal of Semantics*, 11:83–101.

Patrick Blackburn. 2000. Representation, reasoning, and relational structures: a hybrid logic manifesto. *Logic Journal of the IGPL*, 8(3):339–365.

William Croft. 1990. *Typology and Universals*. Cambridge Textbooks in Linguistics. Cambridge University Press, Cambridge, United Kingdom.

Joseph H. Greenberg. 1966. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg, editor, *Universals of Language*, pages 73–114. The MIT Press, Cambridge, Massachusetts, second edition edition.

Liliane Haegeman. 1991. *Introduction to Government and Binding Theory*. Basil Blackwell, Oxford.

John A. Hawkins. 1983. *Word Order Universals*. Academic Press, New York, London, etc.

Beryl Hoffman. 1995. *Computational Analysis of the Syntax and Interpretation of "Free" Word-Order in Turkish*. Ph.d. thesis, ircs report 95-17, University of Pennsylvania.

Richard Hudson. 1990. *English Word Grammar*. Basil Blackwell, Oxford, United Kingdom.

Paul Kroeger. 1993. *Phrase Structure and Grammatical Relations in Tagalog*. Dissertations in Linguistics. CSLI Press, Stanford, CA.

Geert-Jan M. Kruijff. 2001. *A Categorial-Modal Logical Architecture of Informativity: Dependency Grammar Logic & Information Structure*. Ph.D. thesis, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic, April.

Jerzy Kuryłowicz. 1964. *The Inflectional Categories of Indo-European*. Carl Winter Universitätsverlag, Heidelberg, Germany.
Igor A. Mel'čuk. 1988. *Dependency Syntax: Theory and Practice*. SUNY Press, Albany NY.
Marc Moens and Mark Steedman. 1988. Temporal ontology and temporal reference. *Journal of Computational Linguistics*, 14:15–28.
Jarmila Panevová. 1974. On verbal frames in functional generative description I. *Prague Bulletin of Mathematical Linguistics*, 22:3–40.
Vladimír Petkevič. 1995. A new formal specification of underlying structures. *Theoretical Linguistics*, 21(1):7–61.
Martin Plátek, Jiři Sgall, and Petr Sgall. 1984. A dependency base for linguistic description. In Petr Sgall, editor, *Contributions to Functional Syntax, Semantics and Language Comprehension*, pages 63–97. Akademia, Prague, Czechoslovakia.
Ivan A. Sag, Gerald Gazdar, Thomas Wasow, and Stephan Weisler. 1985. Coordination and how to distinguish categories. *Natural Language and Linguistic Theory*, 3:117–171.
Petr Sgall, Eva Hajičová, and Jarmila Panevová. 1986. *The Meaning of the Sentence in Its Semantic and Pragmatic Aspects*. D. Reidel Publishing Company, Dordrecht, Boston, London.
Petr Sgall, Oskar E. Pfeiffer, Wolfgang U. Dressler, and M. Půček. 1995. Experimental research on systemic ordering. *Theoretical Linguistics*, 21:197–239.
Petr Sgall, Alla Bémová, and Eva Hajičová. 1996. Remarks on the semantic features of cases and prepositions as related to syntax. In Edna Andrews and Yishai Tobin, editors, *Towards a Calculus of Meaning*, volume 43 of *Studies in Functional and Structural Linguistics*, pages 71–82. John Benjamins, Amsterdam, The Netherlands.
Vladimír Skalička and Petr Sgall. 1994. Praguian typology of languages. In Philip A. Luelsdorff, editor, *Prague School of Structural and Functional Linguistics*, number 41 in Linguistics and Literary Studies in Eastern Europe, pages 333–357. John Benjamins, Amsterdam, The Netherlands.
Mark Steedman. 2000a. Information structure and the syntax-phonology interface. *Linguistic Inquiry*, 31(4):649–689.
Mark Steedman. 2000b. The productions of time. Draft 4.0, July 24.
Mark Steedman. 2000c. *The Syntactic Process*. The MIT Press, Cambridge Massachusetts.
Susan Steele. 1978. Word order variation: A typological study. In Joseph H. Greenberg, editor, *Universals of Language. Volume 4: Syntax*, pages 585–624. Stanford University Press, Stanford, California.
Lucien Tesnière. 1959. *Éléments de Syntaxe Structurale*. Klinsieck, Paris, France.
Enric Vallduví and Elisabet Engdahl. 1996. The linguistic realization of information packaging. *Linguistics*, 34:459–519.