

Towards an adequate account of parataxis in Universal Dependencies

Lars Ahrenberg

Department of Computer and Information Science
Linköping University

Workshop on Universal Dependencies, Paris
August 30, 2019



Overview

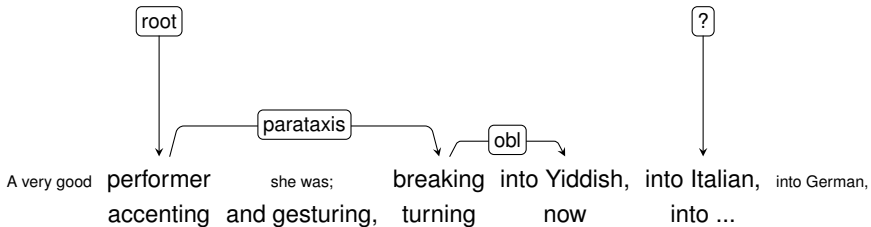
- Introduction: Some problematic examples
- Parataxis and its competitors in UD guidelines
- Annotation diversity in selected PUD treebanks
- Conclusions and suggestions

An example: vivid style

A very good performer she was; breaking into Yiddish, into Italian, into German, accenting and gesturing, turning now into a clique of elderly Jews, now into a frightened small boy.

Source: Jeanette Winterson. *Gut symmetries*, Granta Books, London, 1997.

Vividness



Another example

Her waist was curved, her legs were long, her breasts round, her stomach was flat, her bottom was not.

UD characterization of parataxis

parataxis is a relation between a word (often the main predicate of a sentence) and other elements, such as a sentential parenthetical or a clause after a ':' or a ';', placed side by side without any explicit coordination, subordination, or argument relation with the head word.

Five sub-types are mentioned: side-by-side sentences, reported speech, interjected clauses, tag questions, and news article bylines.

<https://universaldependencies.org/u/dep/parataxis.html>

UD characterization of conj

A *conjunct* is the relation between two elements connected by a coordinating conjunction, such as and, or, etc.

Coordination may be *asyndetic*, which means that the coordinating conjunction is omitted. Commas or other punctuation symbols will delimit the conjuncts in the typical case. Asyndetic coordination may be more frequent in some languages, while in others, conjunction will appear between every two conjuncts.

<https://universaldependencies.org/u/dep/conj.html>

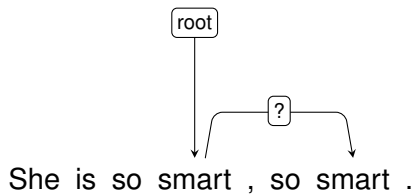
Tesnière on asyndetic conjuncts

CHAPITRE 137. — La jonction sans jonctif.

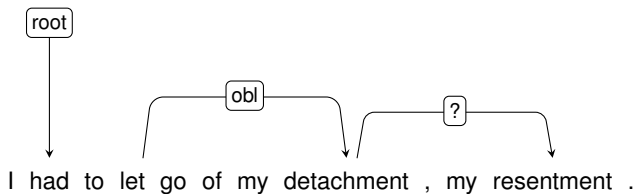
1. — Le moyen le plus simple de joncter deux nœuds est la juxtaposition. Il y a alors absence de jonctif, ce que l'on peut exprimer en terminologie grammaticale en disant que le jonctif est zéro.

2. — Le jonctif zéro existe dans un certain nombre de langues et peut même être employé en français : fr. *Je vois, je sais, je crois, je suis désabusée* (Corneille, *Polyeucte*, V, 5); lat. *Veni, uidi, uici* (César) « Je suis venu, j'ai vu, j'ai vaincu » ;

Example: Repetition



Example: Reformulation



Frequency of *parataxis* in 14 other PUD treebanks

Treebank	en	ar	cs	de	es	fi	fr	hi
Frequency	97	24	20	68	106	108	106	93

Treebank	en	id	it	pt	ru	sv	tu
Frequency	97	116	99	103	195	134	74

Comparison English – other languages

How are *parataxis* relations distributed compared to English?

Treebank	Overlaps	English only	Other only
UD_Arabic-PUD	15	82	9
UD_Czech-PUD	10	87	10
UD_Finnish-PUD	73	24	35
UD_French-PUD	62	35	42
UD_German-PUD	43	54	25
UD_Hindi-PUD	40	57	53
UD_Indonesian-PUD	44	53	70
UD_Italian-PUD	73	24	26
UD_Portuguese-PUD	79	18	24
UD_Russian-PUD	74	23	112
UD_Spanish-PUD	72	25	34
UD_Swedish-PUD	86	11	48
UD_Turkish-PUD	43	54	26

Types of *parataxis* in English PUD

Subtype	Frequency
Side-by-side sentences	44
Reported speech	42
Parentheticals	8
Tag questions	0
Other	3
Total	97

Side-by-side sentences

After all, the internet is not a luxury; it is an essential tool.

He was utterly gracious and took it like a man, I am beyond furious.

Reported speech

“There are 1200 volunteers working for the parade alone,” González Anaya says.

He told the BBC’s Andrew Marr Show: “A general election is frankly the last thing that the Government wants.”

Parentheticals

And, she granted, “you have to look at where she has acknowledged that we need to do something different—*we can do better*—and where she has expressed regret.”

Afterwards, browse at the morning market (6.30-10am): rice lollipops, wasp cocoons (*the pupae are considered a delicacy*), buffalo lung, betel-nut bark and leaves, live toads and catfish.

Side-by-side sentences

parataxis or *conj*?

EN: I'm **going** to jail either way, **hope** it was worth it

(n01011017) *parataxis(going, hope)*

IT: In entrambi i casi **finirò** in prigione, **spero** ne sia valsa la pena

(n01011017) *conj(finirò, spero)*

Observations on parataxis vs. conj

- Hard cases are comma-separated units without explicit conjuncts
- A conjunction can often be inserted without changing the denotation of a construction
- Translators may insert conjunctions in the target language ('explicitation')
- There are sentences with overt conjunctions annotated as parataxis

Reported speech

EN: "This is a **disaster** for pain patients," Mailis **said** in an interview ...
(n01041006) *parataxis(disaster,said)*

DE: Das ist eine **Katastrophe** für Schmerzpatienten, **sagte** Mailis in einem Interview ...
(n01041006) *ccomp(sagte,Katastrophe)*

FR: C'est un **désastre** pour les malades en souffrance, a **déclaré** Mailis dans un entretien ...
(n01041006) *parataxis(déclaré,désastre)*

Observations on reported speech

- The UD guidelines are detailed, clear and with many examples
- The analysis as parataxis has authoritative support (e.g. Halliday, 1985)
- When reported speech is indirect, or topicalized, a ccomp-analysis can be motivated
- With parataxis, a semantic argument of a verb is treated as paratactic
- The head may be small compared to the daughter: No, said Henry with a smile.

Parentheticals

EN: ... and sent so many **species** - not just the **dinosaurs** - into oblivion .
(n01023034) *parataxis(species,dinosaurs)*

ES: ... y que hizo que muchas **especies**, no solo los **dinosaurios**, cayeran en el olvido.
(n01023034) *appos(especies,dinosaurios)*

FR: ... qui ont causé l'extinction de nombreuses **espèces**, et pas uniquement des **dinosaures**.
(n01023034) *conj(espèces,dinosaures)*

Observations on parentheticals

- Hard cases are when the parenthetical is a (long) phrase
- Sometimes a conjunction may be inserted, (ruling out *appos?*)

Paratactic relations in UD

Paratactic relation: a relation between units of roughly equal status, where dependency direction is essentially defined by order.

Level	Relations
Clause	parataxis
(Noun) Phrase	appos
Multiword expressions	fixed, flat
Generic	conj, list

Conclusions and suggestions

- The variation in the use of *parataxis* that can be seen is at least partly caused by differences in annotation practices

Conclusions and suggestions

- The variation in the use of *parataxis* that can be seen is at least partly caused by differences in annotation practices
- Disallow *conj* when no conjunction is present.

Conclusions and suggestions

- The variation in the use of *parataxis* that can be seen is at least partly caused by differences in annotation practices
- Disallow *conj* when no conjunction is present.
- There is no need for more than one general relation of asyndetic sequencing of clauses and phrases, *parataxis*. *list* can be deprecated.

Conclusions and suggestions

- The variation in the use of *parataxis* that can be seen is at least partly caused by differences in annotation practices
- Disallow *conj* when no conjunction is present.
- There is no need for more than one general relation of asyndetic sequencing of clauses and phrases, *parataxis*. *list* can be deprecated.
- *parataxis* may be further subtyped for the genres and treebanks that need it, (as is done for UD_Italian-Postwita and UD_French-Spoken.)

Conclusions and suggestions

- The variation in the use of *parataxis* that can be seen is at least partly caused by differences in annotation practices
- Disallow *conj* when no conjunction is present.
- There is no need for more than one general relation of asyndetic sequencing of clauses and phrases, *parataxis*. *list* can be deprecated.
- *parataxis* may be further subtyped for the genres and treebanks that need it, (as is done for UD_Italian-Postwita and UD_French-Spoken.)
- The choice between *parataxis* and *appos* needs clarification.