



T-Layer of PDT: Topic Focus Articulation

Markéta Lopatková

Institute of Formal and Applied Linguistics, MFF UK

lopatkova@ufal.mff.cuni.cz

Basic Terms



- topic focus articulation = information structure
~ dynamism of the discourse
cz: aktuální členění, informační struktura věty;
funkční větná perspektiva

- **topic**

cz: východisko, téma,
eng: topic, theme, T

vs.

- **focus**

cz: jádro, réma,
eng: focus, comment, F

about what we speak

what is said about that topic

Basic Terms



- topic focus articulation = information structure
~ dynamism of the discourse
cz: aktuální členění, informační struktura věty;
funkční větná perspektiva

- **topic**

cz: východisko, téma,
eng: topic, theme, T

vs.

- **focus**

jádro, réma,
eng: focus, comment, F

about what we speak

what is said about that topic

Táta přišel. vs. Přišel táta.

*Papa came. It was **papa** who came.*



word order

intonation (intonation centre)

- global characteristics of the sentence

FGD and TFA



- FGD: TFA as a part of a (grammatical) meaning of a sentence
 - expressed by grammatical means (as word order)
 - sentences with different TFA may have different meaning:

1. quantification:

Na Moravě se mluví česky.
In Moravia, one speaks Czech
Mnoho lidí čte málo knih.

vs. *Česky se mluví na Moravě.*
Czech is spoken in Moravia
Málo knih čte mnoho lidí.

FGD and TFA



- FGD: TFA as a part of a (grammatical) meaning of a sentence
 - expressed by grammatical means (as word order)
 - sentences with different TFA may have different meaning:

1. quantification:

Na Moravě se mluví česky.
In Moravia, one speaks Czech
Mnoho lidí čte málo knih.

vs. *Česky se mluví na Moravě.*
Czech is spoken in Moravia
Málo knih čte mnoho lidí.

2. negation

Maminka nemá klíče.
Mamma has no keys.

Klíče nemá maminka.
These keys, mamma doesn't have.

FGD and TFA



- FGD: TFA as a part of a (grammatical) meaning of a sentence
 - expressed by grammatical means (as word order)
 - sentences with different TFA may have different meaning:

1. quantification:

Na Moravě se mluví česky.
In Moravia, one speaks Czech
Mnoho lidí čte málo knih.

vs. *Česky se mluví na Moravě.*
Czech is spoken in Moravia
Málo knih čte mnoho lidí.

2. negation

Maminka nemá klíče.
Mamma has no keys.

Klíče nemá maminka.
These keys, mamma doesn't have.

*Dogs must be **carried**.*

***Dogs** must be carried.*

*Dobrá zpráva Češi udělali **revoluci**.*
Good news: Czechs made a revolution.

*Špatná zpráva: revoluci udělali **Češi**.*
Bad news: The revolution was made
by Czechs.

Contextual Boundness in FGD



- **local characteristic of nodes** ... based on context
 - CB = contextually bound ... refer to previous context
Tu knihu [tfa=t] ti přinesu zítra. *I will bring you the book tomorrow.*
pronominalization, coreference

Contextual Boundness in FGD



- **local characteristic of nodes** ... based on context

- CB = contextually bound ... refer to previous context

Tu knihu [tfa=t] ti přinesu zítra. I will bring you the book tomorrow.

pronominalization, coreference

- CN = contextually non-bound ... new information

Ještě jsem to nenapsal [tfa=f]} , ale zítra to napíšu [tfa=f]

I have not written it yet, but tomorrow I will write it.

Contextual Boundness in FGD – NEW



- **local characteristic of nodes** ... based on context

- CB = contextually bound ... refer to previous context

Tu knihu [tfa=t] ti přinesu zítra. I will bring you the book tomorrow.

pronominalization, coreference

- CN = contextually non-bound ... new information

Ještě jsem to nenapsal [tfa=f] , ale zítra to napíšu [tfa=f]

I have not written it yet, but tomorrow I will write it.

- contrastive topic ... bound; a choice from a set of alternatives
contrastive stress (intonation)

(Dnes nesu jen dopis.) Tu knihu [tfa=c] ti přinesu zítra [tfa=f]

(Today I am bringing just (a) letter.) The book, I will bring you tomorrow.

Jedině [tfa=f] s úspěšnými [tfa=f] vzory [tfa=c] se můžeme poměřovat [tfa=f]

*Only with **successful models**, we can **compare**.*

*Jemu [tfa=c] to [tfa=t] Martin [tfa=t] nedal [tfa=f] Martin **did not give** it only to him.*

*Janu [tfa=c] Marie [tfa=t] neviděla [tfa=f] Only Jane, Mary **did not see**.*

PDT: Two Phenomena wrt TFA



- contextual boundness ... attribute tfa
- communicative dynamism ... underlying order of nodes

documentation: <http://ufal.mff.cuni.cz/pdt2.0/doc/manuals/en/t-layer/html/index.html>



1. Contextual Boundness in PDT

- **local characteristic of nodes** ... based on context
 - CB = contextually bound ... refer to previous context
 - CN = contextually non-bound ... new information
 - contrastive topic ... bound; a choice from a set of alternatives
- PDT ... attribute tfa

c	the node represents a contrastive contextually bound expression
f	the node represents a contextually non-bound expression
t	the node represents a non-contrastive contextually bound expression

- PDT nodetypes with tfa:
 - complex, qcomplex
 - list, dphr, atom
- PDT nodetypes without tfa:
 - root, coap, CM, fphr (\rightarrow #Forn)

2. Word Order and Communicative Dynamism



- **communicative dynamism** (CD) ... relative degree of importance of an expression in comparison with other expressions in the sentence
- hypothesis:
CD is expressed by word order
wrt **the governing node and to siblings** of a given node
 - CB nodes precede CN nodes, contrastive topic being the first
 - CB nodes ... deep word order ~ surface word order (speaker's strategy)
 - CN nodes ... deep word order ~ SO (next slide) and CD

2. Word Order and Communicative Dynamism



- **system ordering** (SO) ... the default order of items depending on a verb; SO is considered to be fixed for a given language

Rád vypravuje dětem pohádky.

He likes to tell them fairy tales.

*Rád vypravuje pohádky.CB **dětem**.CN.*

He likes to tell fairy tales to children.

Přenesl křeslo z ložnice do jídelny.

He moved the chair from the bedroom to the living room.

*Přenesl křeslo do jídelny.CB **z ložnice**.CN*

He moved the chair to the living room from the bedroom.

2. Word Order and Communicative Dynamism



- **system ordering** (SO) ... the default order of items depending on a verb; SO is considered to be fixed for a given language

Rád vypravuje dětem pohádky.

He likes to tell them fairy tales.

Rád vypravuje pohádky.CB dětem.CN.

He likes to tell fairy tales to children.

Přenesl křeslo z ložnice do jídelny.

He moved the chair from the bedroom to the living room.

Přenesl křeslo do jídelny.CB z ložnice.CN

He moved the chair to the living room from the bedroom.

- SO for Czech (for CN expressions)

ACT – temporal m. – m. of manner – local and direct. m. – ADDR – ORIG – PAT – EFF

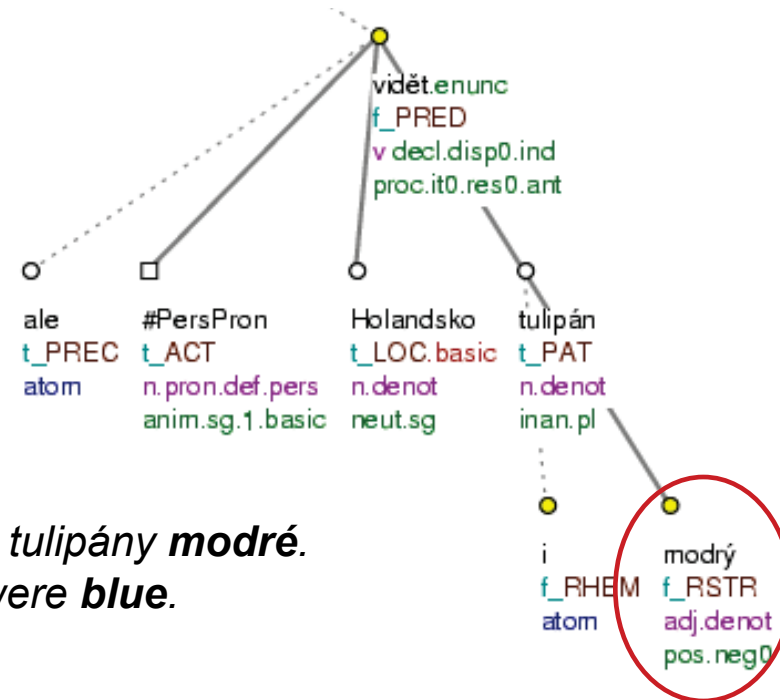
- SO for Czech (for CB expressions, incl. contrastive)

VOCAT – PREC – ATT – (RHEM) – contrastive topic – expressed CB expressions – elided expressions – #PersPron – background

Intonation



- an important indicator of TFA:
 - contextual boundness, communicative dynamism
- the **intonation centre** of a sentence:
 - the prosodic unit carrying the "sentential" stress (in the spoken form)
 - hypothesis:
 - Cz: the intonation centre signals a focus proper
 - the most dynamic expression
 - (usually the last sentence unit(s))



(Mám rád červené tulipány.) Ale v Holandsku jsem viděl i tulipány **modré**.
(I like red tulips.) But in Holland, I saw also tulips which were **blue**.



Intonation

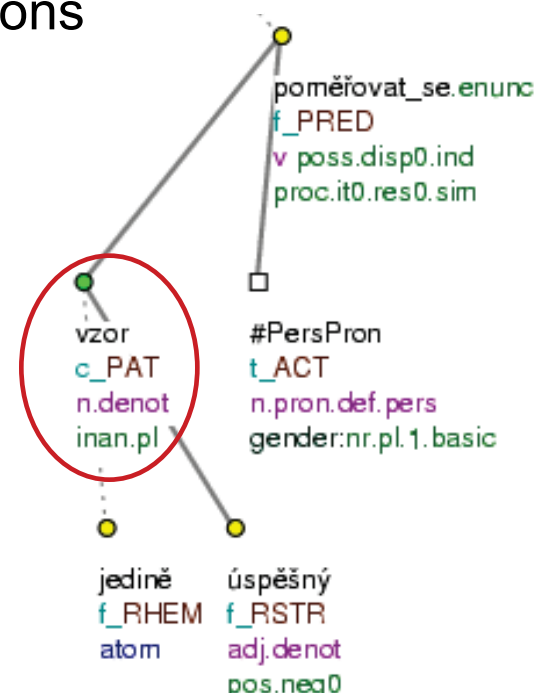
- an important indicator of TFA:
 - contextual boundness, communicative dynamism
- the **intonation centre** of a sentence
- the **contrastive stress**

a specific stress characterized by a rising pitch contour

it signals contrastive contextually bound expressions

(optional)

beginning of a sentence



Jedině s úspěšnými vzory se můžeme poměřovat.
Only with successful models, we can compare.

Topic and Focus



- global characteristic of a sentence
- **topic** ~ that part of the sentence that connects it to the preceding context
 - effective root with $tfa=t$
 - subtrees rooted in child nodes of the effective root with
 - (i) $tfa=t$
 - (ii) that do not contain "focus proper" among their descendants
 - subtrees rooted in the child node of the effective root with $tfa=c$



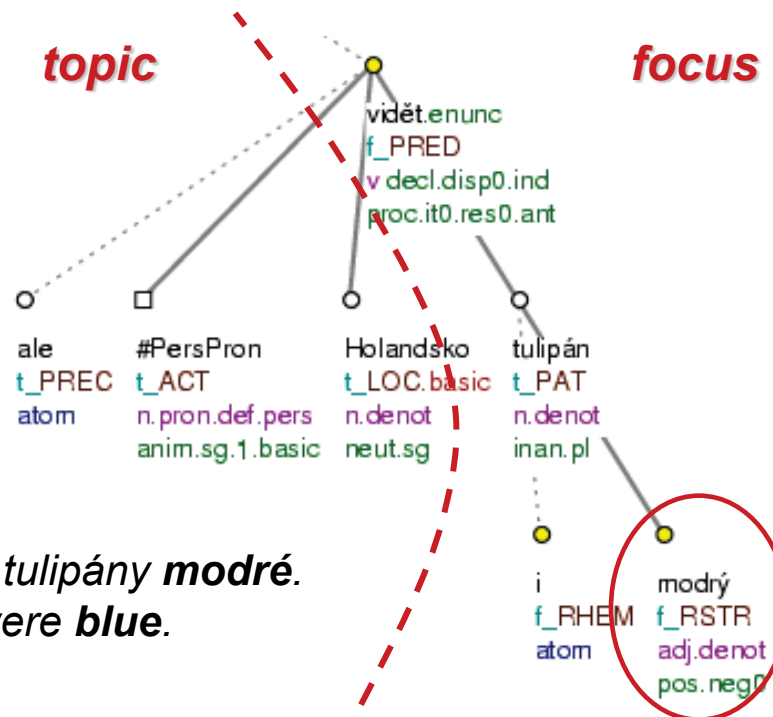
Topic and Focus

- global characteristic of a sentence
- **topic** ~ that part of the sentence that connects it to the preceding context
 - effective root with $tfa=t$
 - subtrees rooted in child nodes of the effective root with
 - (i) $tfa=t$
 - (ii) that do not contain "focus proper" among their descendants
 - subtrees rooted in the child node of the effective root with $tfa=c$
- **focus** ~ that part of the sentence that introduces new info (not deducible from the context); more communicatively important
 - effective root with $tfa=f$
 - subtrees rooted in the child nodes of the effective root with $tfa=f$
 - all deeply embedded nodes $tfa=f$ depending on a node with $tfa=t$ provided that one of them is a "focus proper"

Topic and Focus



- **focus proper** ... the most dynamic and communicatively significant CN part of the sentence
 - intonation centre



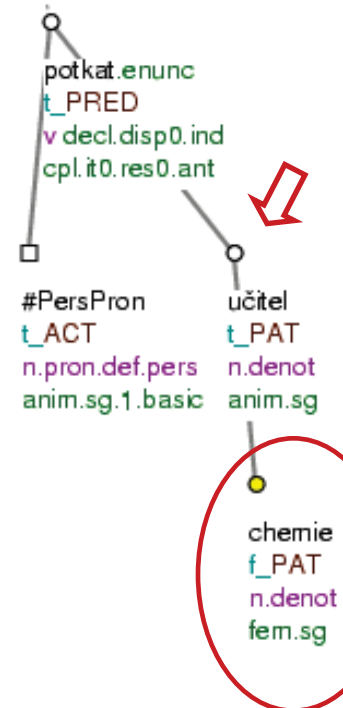
(Mám rád červené tulipány.) Ale v Holandsku jsem viděl i tulipány **modré**.
(I like red tulips.) But in Holland, I saw also tulips which were **blue**.

PDT: valency II.

Topic and Focus



- **focus proper** ... the most dynamic and communicatively significant CN part of the sentence
 - intonation centre
- **quasi-focus** ... (both contrastive and non-contrastive) contextually bound expressions, on which the focus proper is dependent (may be also more deeply embedded)
 - to the right from their governing node !!!

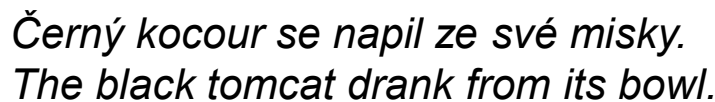


(Kterého učitele jsi potkal?) Potkal jsem učitele **chemie**.

(Which teacher did you meet?) I met the teacher who teaches **chemistry**.



- nodes for CB expressions **to the left** from their governing node
- nodes for CN expressions **to the right** from their governing node
- focus proper: on the **rightmost path** leading from the effective root of the tectogrammatical tree
- a tectogrammatical tree is **projective**



Question Test



- criterion for distinguishing a boundary between topic and focus:
(Hajičová, 2002)

Soused přinesl babičce dárek.

A/The neighbor has brought a gift to the grandma.

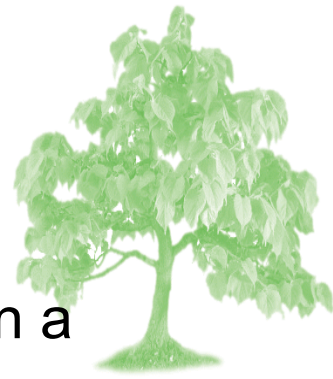
- *What has happened?*
- *What has the neighbor done?*
- *What has the neighbor brought to the grandma?*
- *What and to whom has the neighbor brought?*
- *Who has ~~brought~~ a gift to the grandma?*

Rhematizers

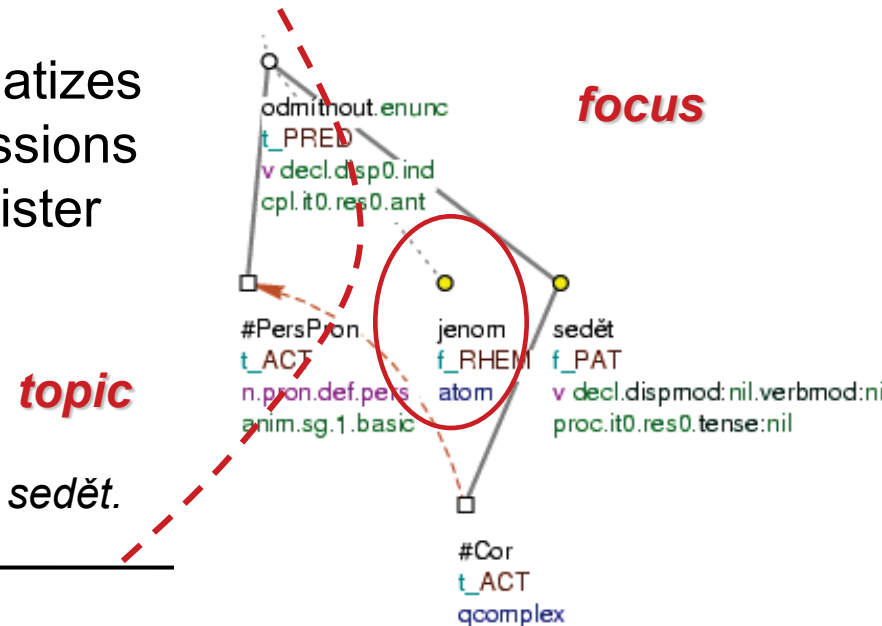


- ***rhematizers*** ... expressions that signal TFA categories in a sentence; they put 'stress' on a part of a sentence
 - typical rhematizers in Cz:
pouze, jen, jenom, jedině, toliko výlučně, výhradně, zejména, zvláště, především, obzvlášť, hlavně, například, ...
only, just, merely, purely, entirely, especially, primarily, mainly, firstly, for example, ...
PLUS negation / affirmation expressions ... *ne, ano, #Neg*
-

Rhematizers



- ***rhematizers*** ... expressions that signal TFA categories in a sentence; they put 'stress' on a part of a sentence
- typical rhematizers in Cz:
pouze, jen, jenom, jedině, toliko výlučně, výhradně, zejména, zvláště, především, obzvláště, hlavně, například, ...
only, just, merely, purely, entirely, especially, primarily, mainly, firstly, for example, ...
PLUS negation / affirmation expressions ... *ne, ano, #Neg*
- position of a rhematizer
 - right before the expressions it rhematizes
 - scope: it affects one or more expressions
 - annotation: prototypically as a left sister



Odmítl jsem jenom sedět.
I refused just to sit.

Rhematizers

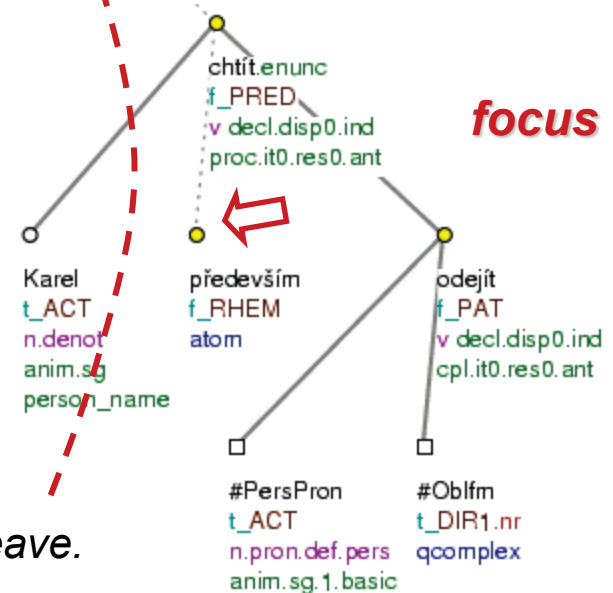


- ***rhematizers*** ... expressions that signal TFA categories in a sentence; they put 'stress' on a part of a sentence
- typical rhematizers in Cz:
pouze, jen, jenom, jedině, toliko výlučně, výhradně, zejména, zvláště, především, obzvláště, hlavně, například, ...
only, just, merely, purely, entirely, especially, primarily, mainly, firstly, for example, ...
PLUS negation / affirmation expressions ... *ne, ano, #Neg*

- position of a rhematizer
 - right before the expressions it rhematizes
 - scope: it affects one or more expressions
 - annotation: prototypically as a left sister
BUT ... a main predicate, coordination

topic

Karel především chtěl, abych odešel.
Charles primarily wanted "so that" I leave.



References



- Manual for Tectogrammatical Annotation
<http://ufal.mff.cuni.cz/pdt2.0/doc/manuals/en/t-layer/html/index.html>
- Hajičová et al. (2002). *Úvod do teoretické a počítačové lingvistiky. I. sv. - Teoretická lingvistika*. Karolinum, Praha.
- Kučová, L., Hajičová, E., Veselá, K. Havelka, J. (2005) Topic-focus articulation and anaphoric relations: A corpus based probe. In *PBML 84*, pp. 5-12.
- Hajičová, E. (1999) The Prague Dependency Treebank: Crossing the Sentence Boundary". In *Proceedings of TSD'99*, Springer-Verlag Berlin Heidelberg, pp. 20-27