



# PDT – Tectogrammatical Layer Grammatemes

Markéta Lopatková

Institute of Formal and Applied Linguistics, MFF UK

[lopatkova@ufal.mff.cuni.cz](mailto:lopatkova@ufal.mff.cuni.cz)

---



---

# PDT: t-layer

## Goal:

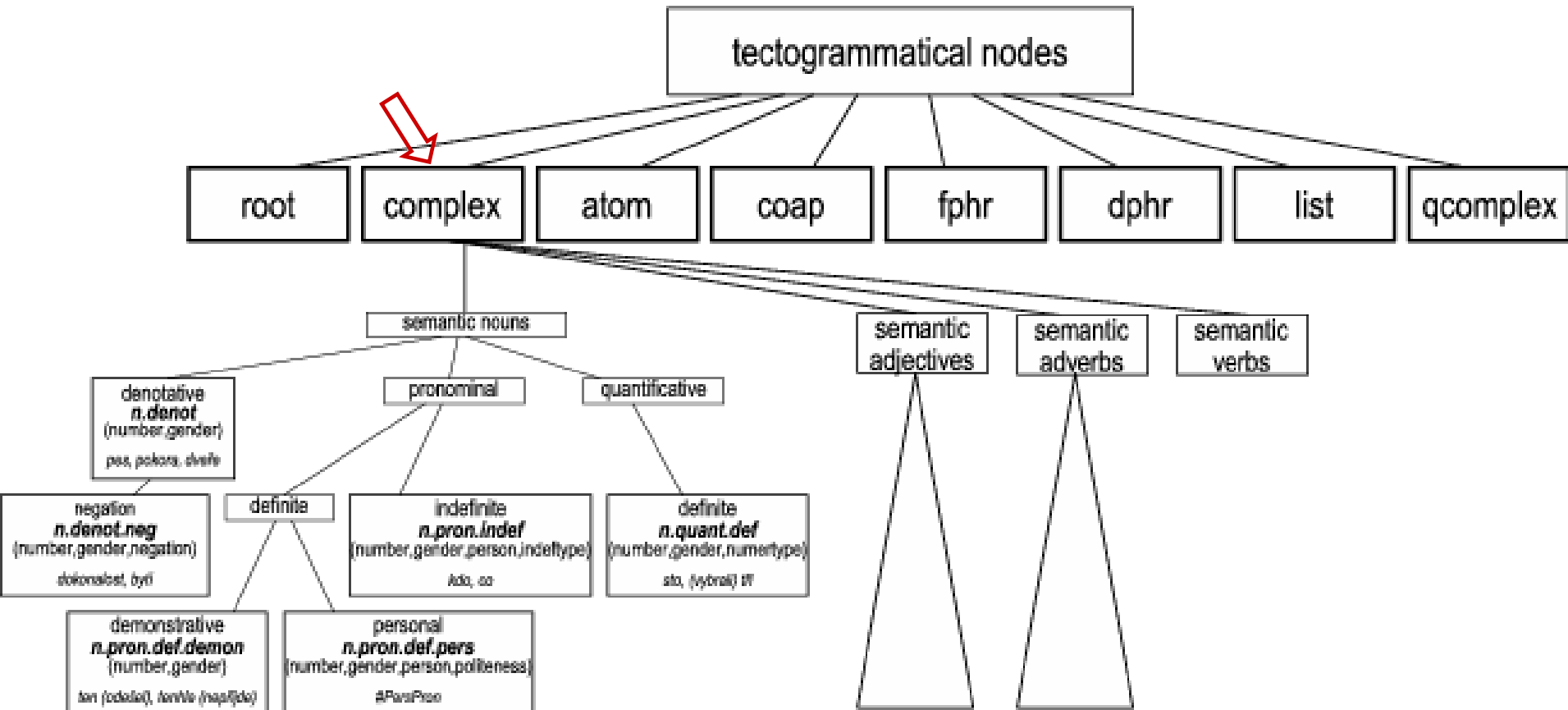
- to describe deep, semantic structure of a sentence  
~ a sentence meaning
- disambiguated
- dependency 'tree'
- information on
  - lexical items: **t-lemma** (primarily)
  - relations between lexical words ~ **deep structure:**  
**functors + subfunctors**
  - **grammatemes**
  - coreferential links
  - topic focus articulation: linear order + tfa attribute  
(cz: aktuální členění)
- PDT 3.0: multiwords expressions, discourse relations, genre specif.

documentation: <http://ufal.mff.cuni.cz/pdt3.0/documentation>

# PDT: node types



- eight node types ... attribute ***nodetype***
- defined on the basis of a *t-lemma* and/or a *functor*

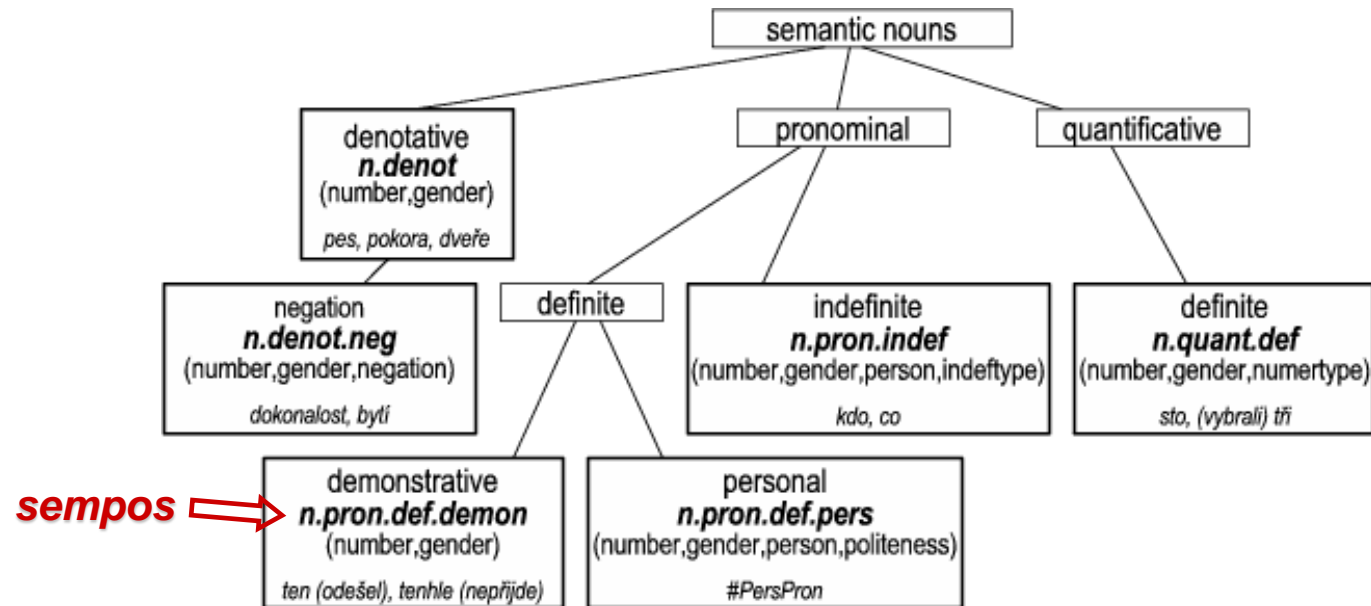


# PDT: system of grammatememes



- to capture meaning of morphological categories
- complex nodes
- attribute *gram/sempos*
  - semantic nouns
  - semantic adjectives
  - semantic adverbs
  - semantic verbs

} further classified



---

# PDT: system of grammemes (cont.)



based on a value of sempos

⇒ 15 grammemes:

- number, gender, person, politeness, **typgroup\*** (SemN)
- numertype (SemN, SemAdj)
- indeftype, negation (SemN, SemAdj, SemAdv)
- degcmp (SemAdj, SemAdv)
- ~~verbmod~~ → **factmod\***, deontmod, aspect, tense, iterativeness, **resultative**, **dispmod**, **diathgram\*** (SemV)

(attribute gram ... technical attribute gathering all grammeme attributes)

+ sentmod

\* PDT 3.0

---

# PDT: system of grammatememes (cont.)



values of grammatememes:

- basic values ... values inherent to an individual grammateme
  - two (or more) values  
e.g. *Viděl jen dva.* [gender=anim|inan] [(He) saw only two.] (not fem, neut)
- special values
  - **nr** ... all values may be assigned to  
(i.e., all basic values are possible in the given case)  
e.g. *Vypral si kalhoty.* [number=nr] [He washed his pants]
  - **nil** ... none of the basic values is suitable (limited)
  - **inher** ... the value is inherited from the coreferred node  
(for grammatical coreference)  
e.g. reflexive *se*, *svůj*, relative pronouns

---

# Grammateme *sentmod*



## sentential modality

... whether the sentence expresses an assertion, a question, a demand etc

- **enunc:** declarative *Petr tam nepřišel.* [*Petr didn't come.*]  
*Škola.* [School.]
- **excl:** exclamative *Vyhráli jsme!* [We won!]  
*Auto!* [Car!]
- **desid:** desiderative *At' se vám daří!* [*I wish you all the best*]  
*Hodně štěstí!* [Good luck!]
- **imper:** imperative *Přijďte včas!* [Come in time!]  
*Pozor!* [Warch out!]
- **inter:** interrogative *Kdy dnes přijdete?* [When do you come?]  
*Škola?* [School?]

nodes representing the main verb or noun ~~and the coordination roots~~

root nodes of subtrees representing direct speech;

root nodes of subtrees representing a parentheses (PAT)

root nodes of title subtrees (labeled with the functor *ID*),

---

---

# Grammateme *number*



- semantic nouns
- values ... **sg**, **pl**
- prototypically corresponds to the morphological number
- BUT
  - pluralia tantum  
e.g. *jedny kalhoty.sg* [one pair of trousers], *dvoje kalhoty.pl* [two pairs of trousers]
  - polite form (cz 'vykání')  
e.g. sg: *vy.sg jste přišel*



---

# Grammateme *gender*



- semantic nouns
- prototypically corresponds to the morphological gender
  - **anim**: *soused* [neighbor], *pes* [dog], (*přišli*) *tři* [three (of them came)]
  - **inan**: *hrad* [castle], *nápad* [idea], *tisíc* [thousand]
  - **fem**: *dívka* [girl], *děti* [children], *dveře* [door], *miliarda* [milliard]
  - **neut**: *děvče* [girl], *dítě* [child], *kolo* [wheel], *ono* [it], *sto* [hundred], *co* [what]

---

# Grammateme *person*



- definite personal and indefinite pronom. semantic nouns, i.e. pronouns referring to an object of communication (third person), to the speaker or hearer (first and second person)
- values:
  - 1: *Já tam nepůjdu.* [I will not go there.], *náš návrh* [our suggestion];
  - 2: *Tvůj názor nesdílím.* [I don't share your view]  
*Vy už jste se přihlásili.* [You.pl have already registered]  
*Vy už jste se přihlásil.* [You.sg have already registered]
  - 3: *On sem nepřijde.* [He will not come], *jeho názor* [his opinion]
- indefinite semantic nouns
  - typically the value 3  
unless different subject-predicate agreement

e.g. *Verše, které kdekdo známe...* [Poems which everybody/whoever know.**1.pl** ]  
*Zachraň se, kdo můžeš* [Save yourself who can.**2.sg** ]

---

# Grammateme *typgroup*



- the semantic opposition of the pair/group meaning  
vs. meaning of single entities

*kalhoty* [trousers], *ruce* [hands, arms], *klíče* [keys], *boty* [shoes]

- plural forms refer rather to a pair/group than to a larger amount of single entities
- combination with numerals → grammaticalized  
*dvoje boty* [two-pairs-of shoes] set numeral vs. *dvě boty* [two shoes] cardinal

• values:

- **group:**
- **single:**

semi-automatical annotation

→ the annotation of number is affected

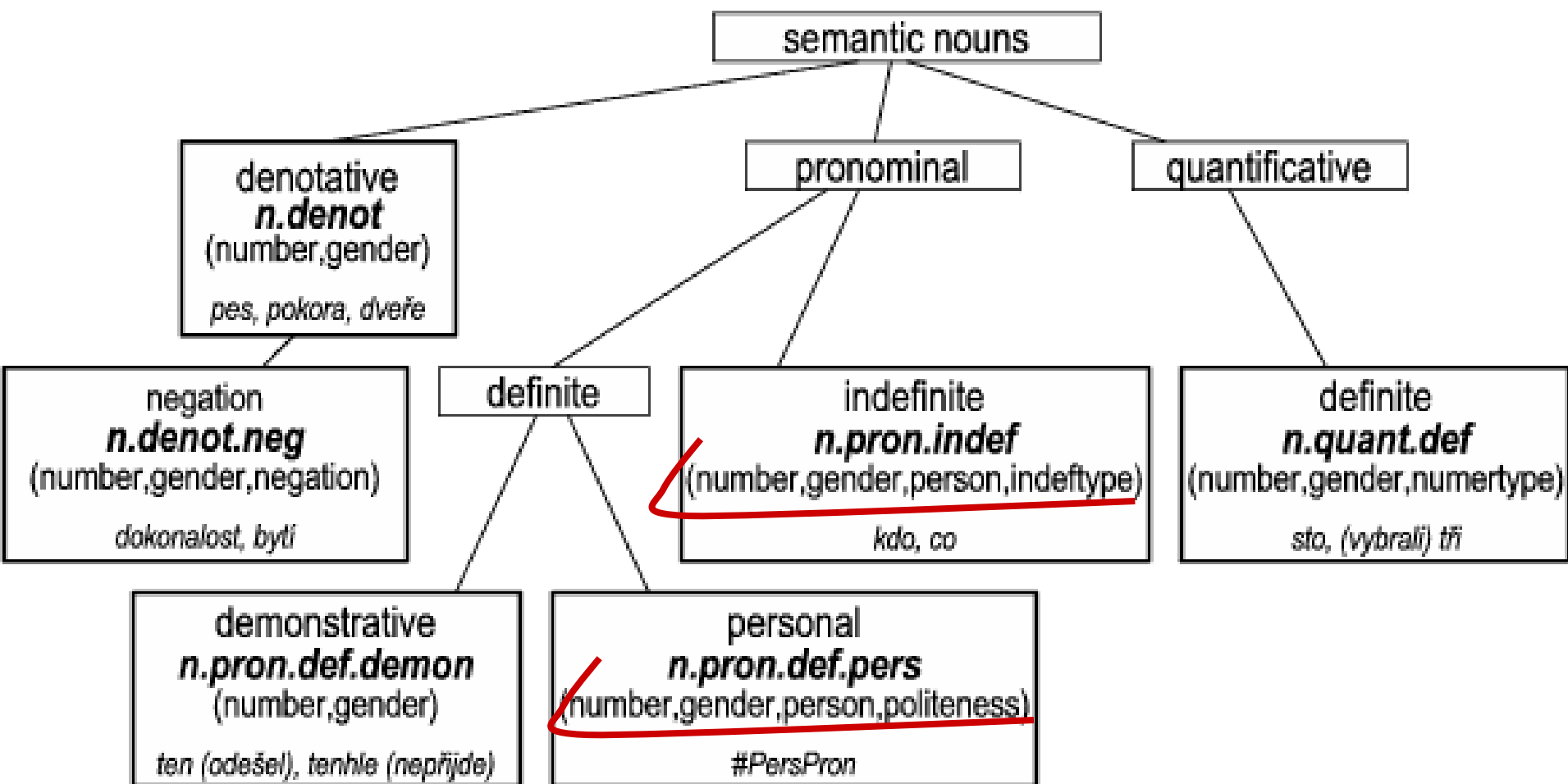
*sg.group*: one pair/group, a plural form of nouns ... (a pair of) trousers; cigarettes (a package)

*pl.group*: more pairs/groups, a plural form of nouns ... (more pairs of) trousers; cigarettes  
(more packages)

*nr.group*: one or more pairs/groups, a plural form of nouns,

*sg.single*: one entity, a singular form of nouns

*pl.single*: more entities, expressed by a plural form of nouns

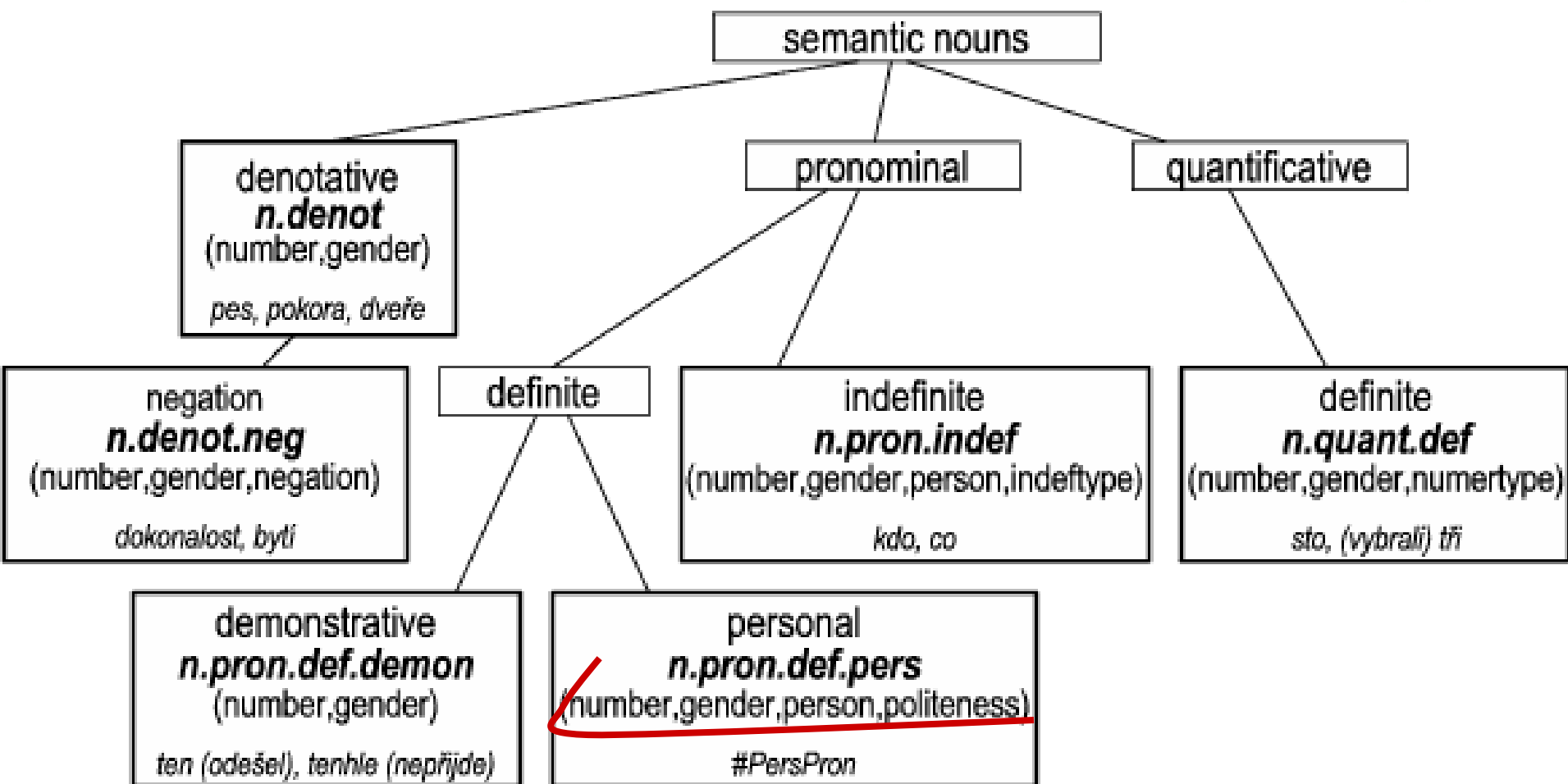


---

# Grammateme *politeness*



- polite form of addressing (cz 'vykání')
- t-lemma #PersPron
- values
  - **polite:** ('vykání')  
*Vy už jste se přihlásil.* [You.sg have already registered]
  - **basic:**  
*Já tam nepůjdu / On tam nepůjde.* [I / He will not go there],  
*Ty bys tam neměl chodit.* [You.sg should not go there]  
*Vy už jste se přihlásili.* [You.pl have already registered]

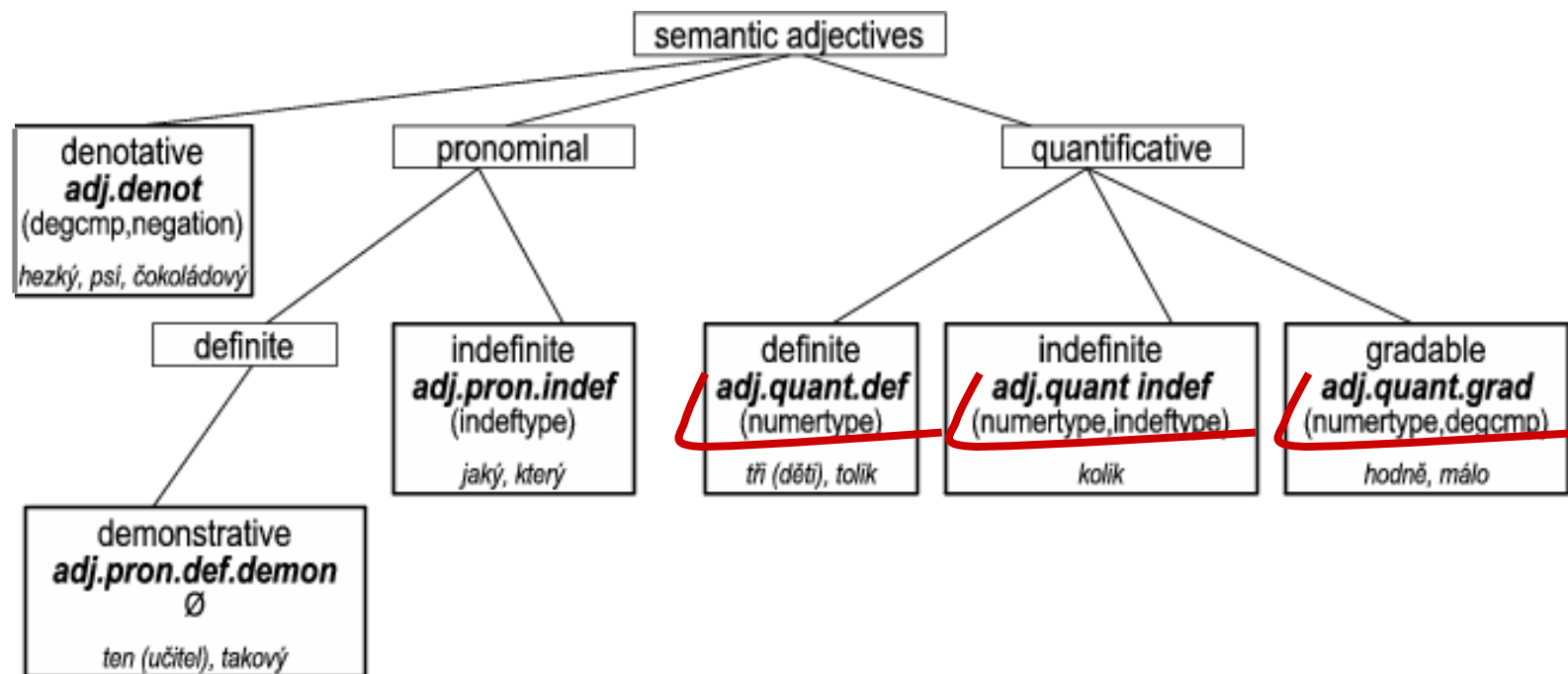
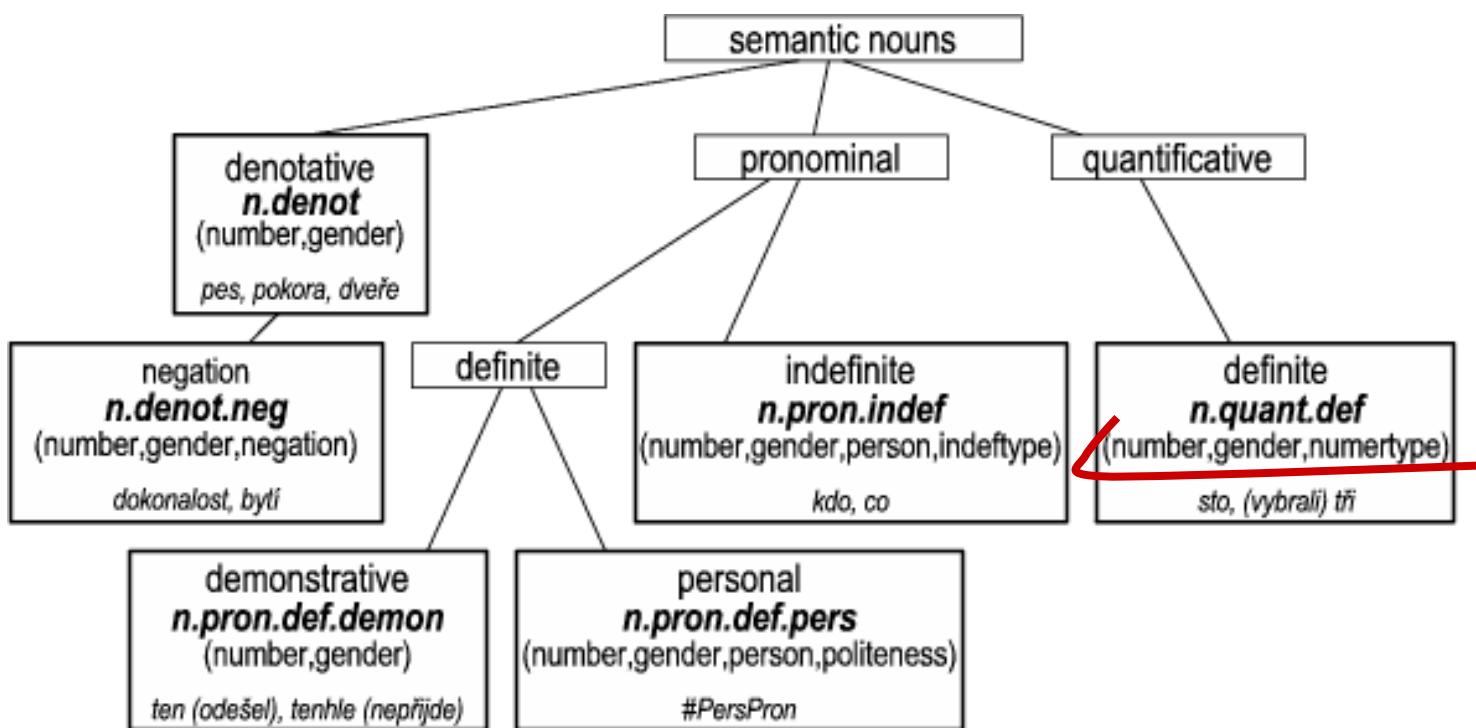


---

# Grammateme *numertype*



- lexical derivation ... ordinal, sort, set, and fraction numerals
- t-lemma + numertype  $\Rightarrow$  m-lemma
- values
  - basic: cardinal numbers  
(cz základní) ... *tři, sto, kolik* [three, hundred, how many]
  - **set**: (cz souborové) ... *troje, stery, kolikery*
  - **kind**: (cz druhové) ... *trojí, sterý, kolikerý*
  - **ord**: (cz řadové) ... e.g. *třetí, stý, kolikátý*
  - **frac**: (cz dílové) ... *třetina, setina*





---

# Grammateme *indeftype*

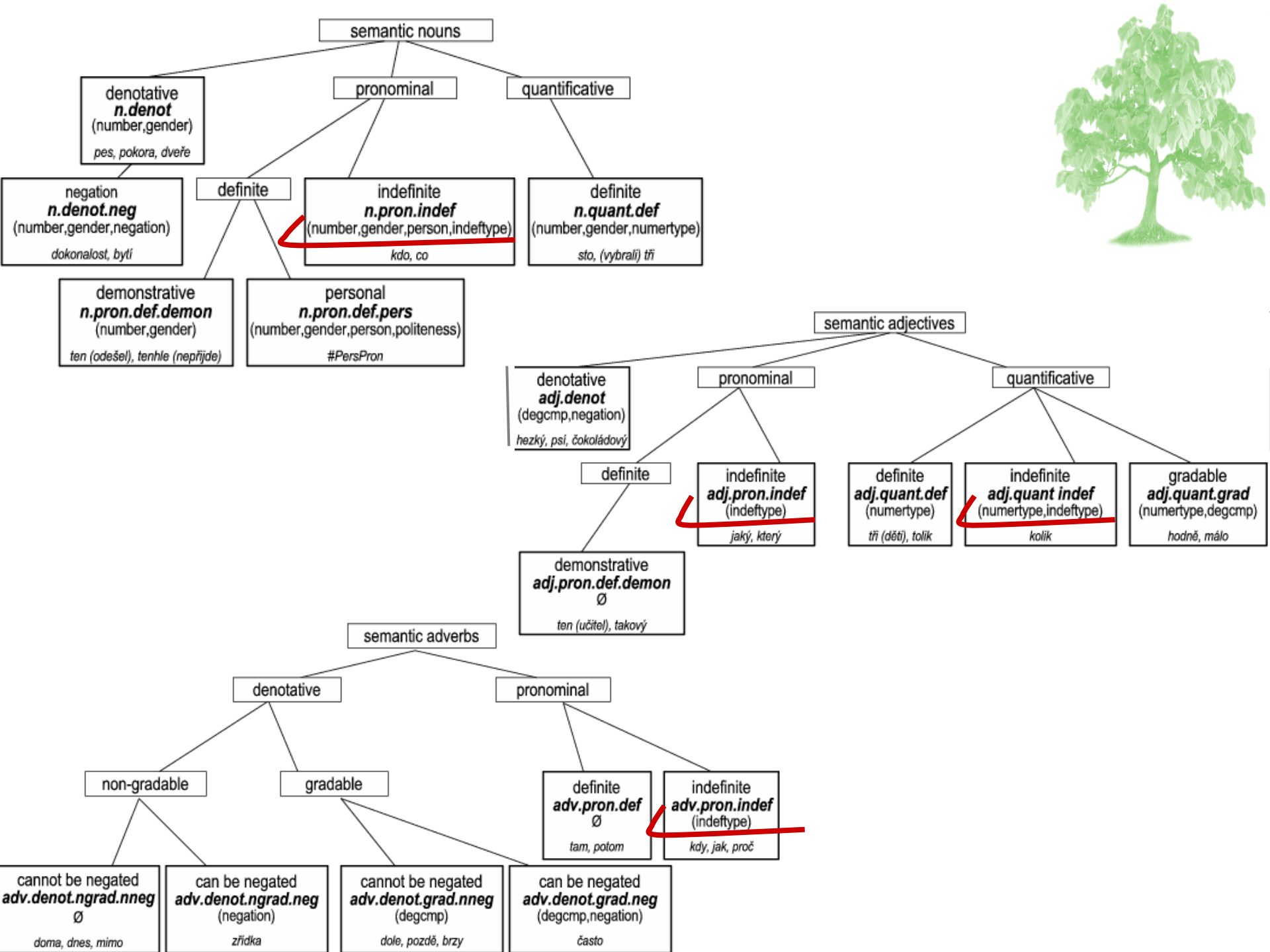


- lexical derivation:  
pronouns, pronominal numerals and pronominal adverbs  
(indefinite, interrogative, negative and totalizing) ...  
derived from the correspond. interrogative or relative pronoun /  
numeral / adverb
- t-lemma + indeftype  $\Rightarrow$  m-lemma
- values
  - **relat, indef1 ... indef6, inter, negat, total1, total2**

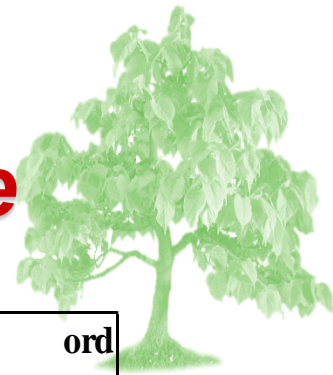
# Grammateme *indeftype*



t-lema:	<i>kdo</i>	<i>co</i>	<i>který</i>	<i>jaký</i>
hodnota gramatému <b>indeftype:</b>				
<b>relat</b>	<i>kdo</i>	<i>co</i>	<i>který, jenž</i>	<i>jaký</i>
<b>indef1</b>	<i>někdo</i>	<i>něco</i>	<i>některý</i>	<i>nějaký</i>
<b>indef2</b>	<i>kdosi, kdos</i>	<i>cosi, cos</i>	<i>kterýsi</i>	<i>jakýsi</i>
<b>indef3</b>	<i>kdokoli(v)</i>	<i>cokoli(v)...</i>	<i>kterýkoli(v)</i>	<i>jakýkoli(v)</i>
<b>indef4</b>	<i>ledakdo,</i> <i>leckdo...</i>	<i>ledaco,</i> <i>lecco...</i>	<i>leckterý,</i> <i>ledakterý</i>	<i>lecjaký,</i> <i>ledajaký</i>
<b>indef5</b>	<i>kdekdo</i>	<i>kdeco</i>	<i>kdekterý</i>	<i>kdejaký</i>
<b>indef6</b>	<i>málokdo,</i> <i>kdovíkd...</i>	<i>máloco...</i>	<i>málokterý...</i>	<i>všelijaký...</i>
<b>inter</b>	<i>kdo,</i> <i>kdopak...</i>	<i>co, copak...</i>	<i>který,</i> <i>kterýpak</i>	<i>jaký, jakýpak</i>
<b>negat</b>	<i>nikdo</i>	<i>nic</i>	<i>žádný</i>	<i>nijaký</i>
<b>total1</b>	<i>všechn</i>	<i>všechn,</i> <i>všechno, vše</i>	—	—
<b>total2</b>	—	—	<i>každý</i>	—



# Grammateme *numertype* and *indeftype*



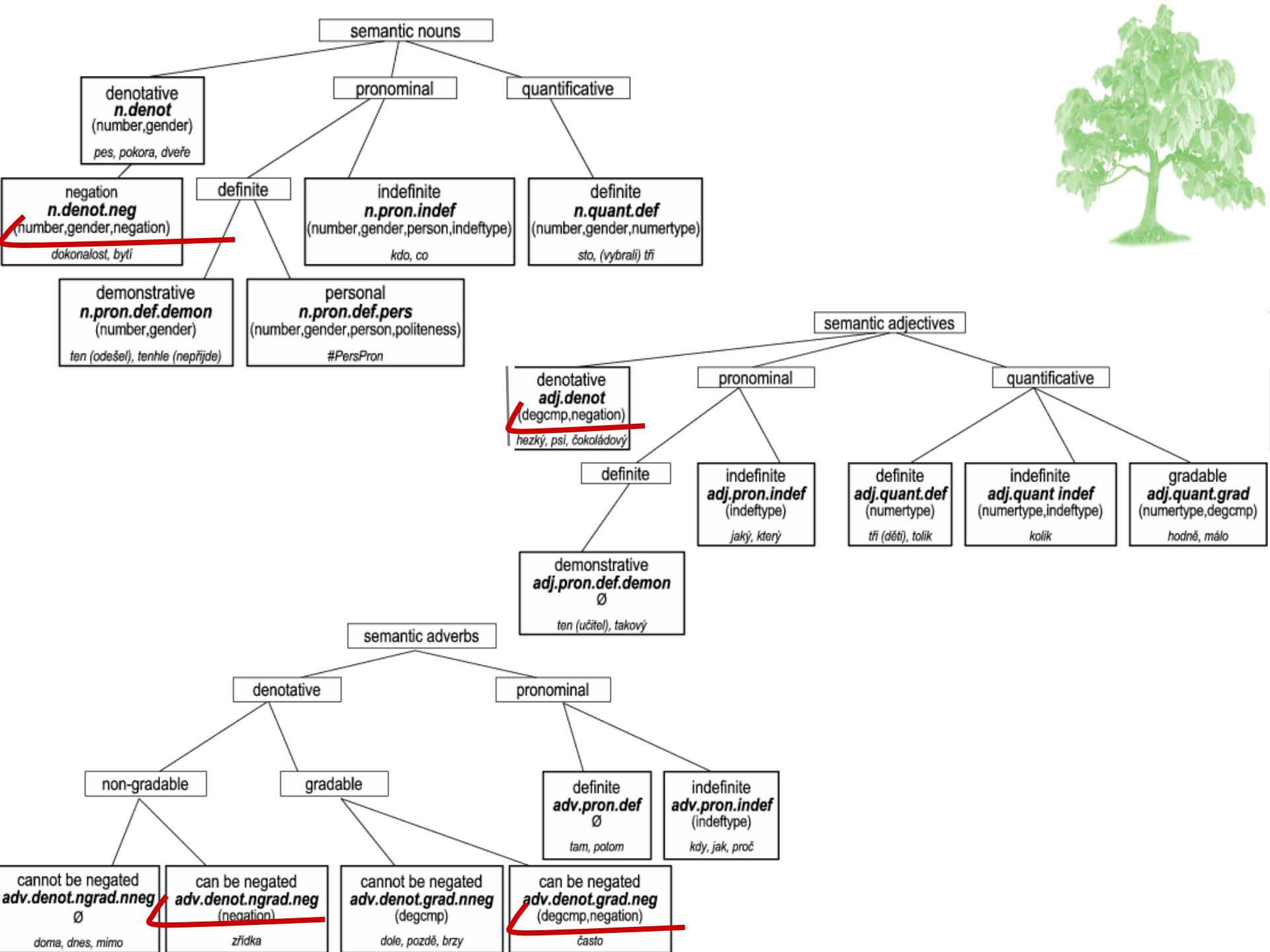
hodnota gramatému <b>numertype:</b>	<b>basic</b>	<b>set</b>	<b>kind</b>	<b>ord</b>
hodnota gramatému <b>indeftype:</b>				
<b>relat</b>	<i>kolik (klíčů), kolikery (dveře)</i>	<i>kolikery (klíče)</i>	<i>kolikerý</i>	<i>kolikátý</i>
<b>indef1</b>	<i>několik (klíčů), několikery (dveře)</i>	<i>několikery (klíče)</i>	<i>několikerý</i>	<i>několikátý</i>
<b>indef2</b>	—	—	—	—
<b>indef3</b>	—	—	—	—
<b>indef4</b>	—	—	—	—
<b>indef5</b>	—	—	—	—
<b>indef6</b>	<i>kdovíkolik (klíčů), kdovíkolikery (dveře)...</i>	<i>kdovíkolikery (klíče)...</i>	<i>kdovíkolikerý...</i>	<i>kdovíkolikátý...</i>
<b>inter</b>	<i>kolik (klíčů), kolikery (dveře)</i>	<i>kolikery (klíče)</i>	<i>kolikerý</i>	<i>kolikátý</i>
<b>negat</b>	—	—	—	—
<b>total1</b>	—	—	—	—
<b>total2</b>	—	—	—	—

---

# Grammateme *negation*



- applicable for:
  - nouns with the endings *-ní, -tí, -ost*
  - all adjectives
  - some adverbs
  - ( NOT for verbs ... node for negation, functor RHEM )
- (positive) t-lemma + negation  $\Rightarrow$  m-lemma
- values:
  - **neg1**: (negative)  
*nezralost (dítěte)* [immaturity (of the child)], *nepěkný* [not-nice], *nedodržování (pořádku)* [non-maintenance (of order)], *nepříliš* [not-very], *nedaleko* [not-far], *nebytí* [not-being]
  - **neg0**: (affirmative)  
*zralost (dítěte)* [maturity (of the child)], *pěkný* [nice], *dodržování (pořádku)* [maintenance (of order)], *příliš* [very], *daleko* [far], *bytí* [being]



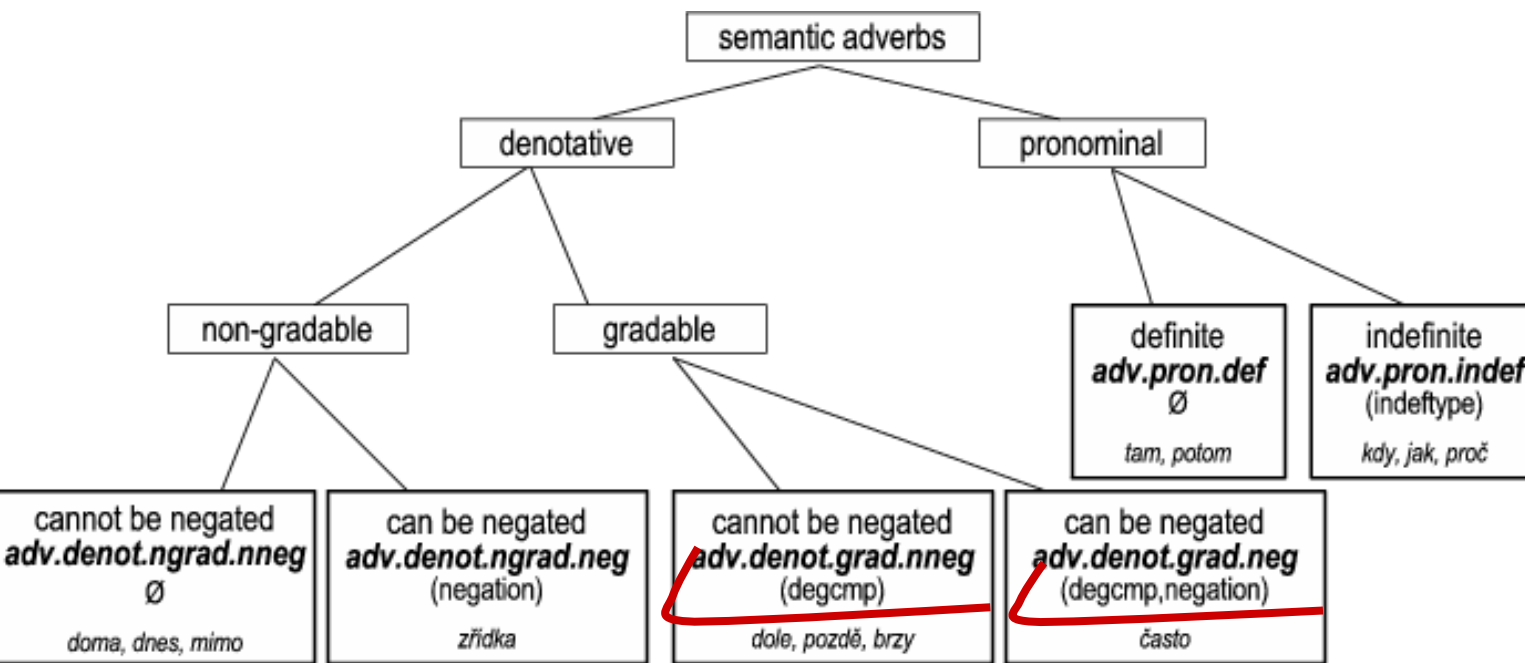
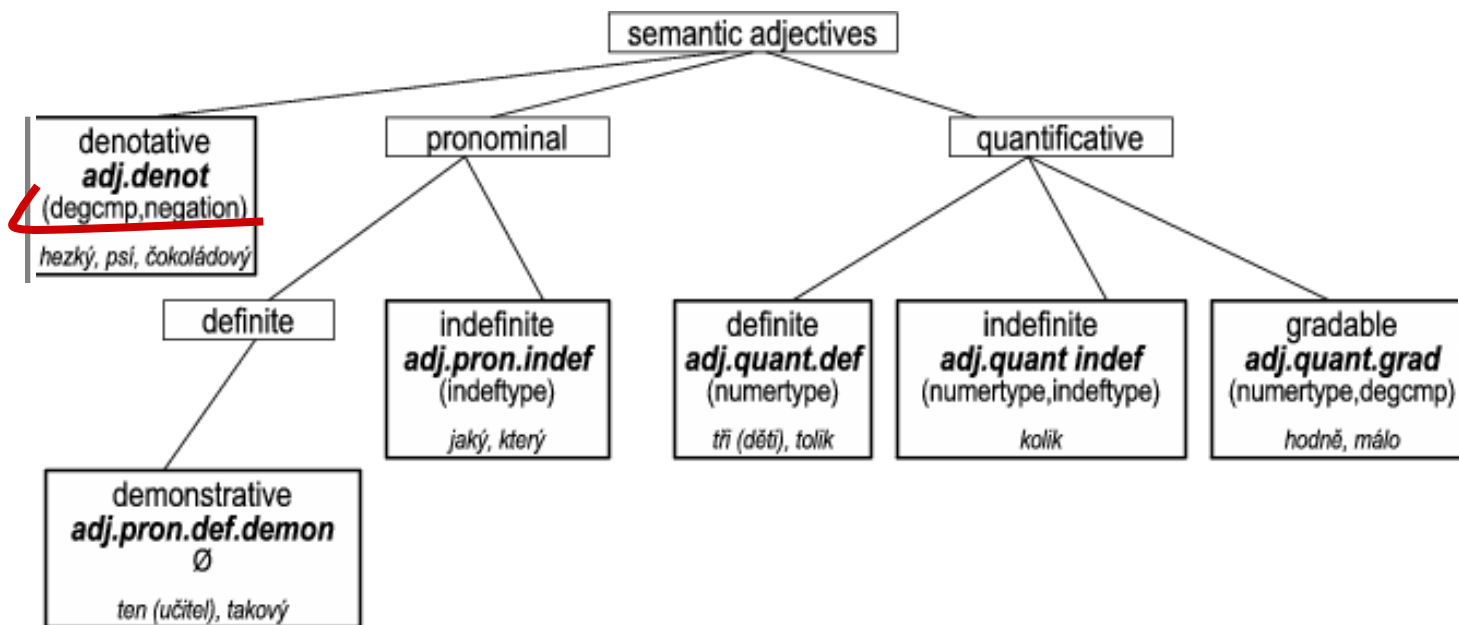
---

# Grammateme *degcmp*



degree of comparison

- applicable for
  - denotative semantic adjectives
  - gradable denotative semantic adverbs ...*blízko* [near] vs. *dnes* [today]
- correlate of the morphological category of degree
- values
  - **pos**: *příjemná hudba* [nice music]
  - **comp**: *příjemnější hudba* [nicer music], *později* [later]
  - **sup**: *nejpříjemnější hudba* [the nicest music], *nejblíže* [the nearest]
  - **acomp**: (absolute comparative)  
*muž tmavší pleti* [a man of darker skin], *starší žena* [an elder(ly) woman]







---

# Grammateme **verbmod**

- whether an event is presented by the speaker  
as **given** or **hypothetical**  
correlate of the morphological category of verbal modality (cz slovesný způsob)
- applicable for verbal predicates

## PDT 2.0

- **ind**: (indicative)

Přišli včas. [They came in time].

Kdy přišli? [When did they come?]

- **imp**: (imperative)

Přijďte tam včas! [Be there in time!]

- **cdn**: (conditional)

Přišli bychom určitě včas. kdyby ... [We would definitely come in time if ...]

- **nil**: (for infinitives, transgressives)

*Každý má povinnost poskytnout člověku první pomoc.*

[Everybody has the obligation to give first aid to the ones who need it]

*Hlasitě naříkajíc, odcházela.* [Loudly crying, she was leaving]

---

# Grammateme *verbmod* → *factmod*



- whether an event is presented by the speaker

as *given* or *hypothetical*

correlate of the morphological category of verbal modality (cz slovesný způsob)

- applicable for verbal predicates

## PDT 2.0 ... *vebmod*

- **ind**: (indicative)
- **imp**: (imperative)
- **cdn**: (conditional)
- **nil**: (for infinitives, transgressives)

## PDT 3.0 ... *factmod*

- **asserted**: (given)
- **appeal**: (required, imperative)
- **potential**: (present conditional)
- **irreal**: (perfect conditional)
- **nil**

When a mistake **occurs.asserted**, it **is.asserted** possible **to ask.nil** the new owner that he **would remit.potential** money to the right owner of securities.

To draw the game **would be.potential** a success.

**Take.appeal** the phone and **call.appeal** (us).

Of course, the majority of Bangladesh Muslim inhabitants **would not have noticed.irreal** the interview in question under common circumstances at all.

The saint drunkard Joseph Roth **would celebrate.potential** his 100th birthday today.

---

# Grammateme *deontmod*



- an event is understood as necessary, possible, permitted etc
- applicable for all verbal forms
- values:
  - **deb:** necessity, i.e. *muset* [must, have to]  
*Musíme zaplatit fakturu včas.* [We have to settle the invoice in time]
  - **vol:** volition, i.e. *chtít, hodlat* [to want, to intend]  
*Chtěl na sebe upozornit.* [He wanted to turn the attention to himself]
  - **perm:** permission, i.e. *smět, moci* [may, to be allowed]  
*Nesmíš kouřit.* [You are not allowed to smoke]
  - **hrt:** obligation, i.e. *mít* [should, to be obliged]  
*Petr ti měl ty podklady poslat.* [Petr was supposed to send you the documents]
  - **poss:** possibility, i.e. *moci, dát se* [may, can, to be possible]  
*To se dá zjistit.* [It is possible to find out]
  - **fac:** ability, i.e. *dovést, umět, moci* [can, to be able]  
*Umí se výborně přetvařovat.* [He can/is able to pretend very well]
  - **decl:** without a modal verb (or if a modal verb has its own node)

---

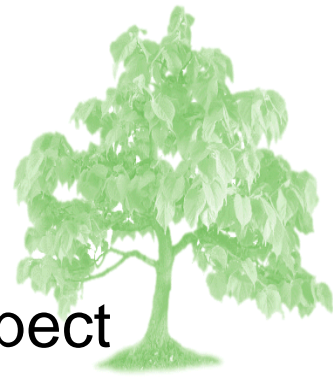
# Grammateme *tense*



- correlate of the morphological category of tense (cz. slovesný čas)
- time relation between two events (relative tense)  
btw an event and a moment of speech (absolute)
- applicable for all finite non-imperative verbal forms and transgressives
- values:
  - **sim**: (simultaneous, the same time as another event / moment of utterance)  
*Píše dopis. Psal dopis.* [He is writing a letter. He was writing a letter.]
  - **ant**: (prior to the time of another event / moment of utterance)  
*Nasal dopis.* [He wrote a letter]
  - **post**: (after the time of another event / moment of utterance)  
*Napíše dopis.* [He will be writing a letter]
  - **nil**: (imperative, infinitive)

---

# Grammateme **aspect**



- correlate of the morphological category of verbal aspect (cz slovesný vid)
- applicable for all verbal forms
- values:
  - **proc**: (progressive, imperfective aspect)  
*Nejraději kupuje / nakupuje nábytek.* [He likes to buy furniture best]
  - **cpl**: (complex, perfective aspect)  
*Koupil / nakoupil už vše potřebné.* [He already bought everything he needed]
  - biaspectual verbs
    - proc**: *Operace byla blokována až do konce 80. let.*  
[The operation was blocked until the end of the 80's ]
    - cpl**: *K 31. lednu 1995 registroval úřad 75 tisíc nezaměstnaných.*  
[By January 31st 1995, the agency registered 75 thousands of unemployed people]

---

# Grammateme *iterativeness*



- applicable for all verbal forms
- values:
  - **it1**: (multiple/iterated/repeated action)  
only verbs with suffixes *-ívát/-ávat, -ívávat/-ávávat*  
*kupovávat* → t-lemma *kupovat* + it1
  - **it0**:  
incl. lexically expressed repetition  
*Každé pondělí chodí nakupovat.* [He used to go shopping each Monday]

---

# Grammateme *diagram* ... PDT 3.0



- diathesis ~ a way how a sentence is structured  
i.e., “from perspective of which complementation of the predicate”  
John ate the cake          vs. The cake was eaten
- applicable for verbal predicates (not imperative)
- values:
  - **act**          I bought a book. I prepared a dinner for my mum.
  - **pass**        The book was bought. The dinner was prepared (by me) in advance.
  - **res1**        *Obchod je otevřen.* (past participle) [The shop is open.]
  - **res2**        *Obchod má otevřeno.* (past participle) [The shop is open.]  
*Babička má oběd uvařen.*  
*Babička má uvařeno.*
  - **recip**        *Děti dostali vynadáno od rodičů.*
  - **disp**        *Kniha se (mi) dobře čte.*  
*Spí se mu tam dobře. Tančilo se skvěle.*
  - **deagent**     *Čeká se tuhá zima. Knihy se kupují v knihkupectví.*

---

# Grammateme *dispmo* ... PDT 2.0



- dispositional modality (spec. construction in Czech)
- applicable for verbal predicates (not imperative)
- the agent's attitude towards the event is expressed

*Tato studie se studentům četla dobře.*

[It was easy for the students to read the article]

*Spalo se nám tu výborně.*

[We slept very well here]

- agent in dative (may be omitted) + refl. form of verb + manner +  
+ (compl. in accusative → compl. in nominative)
- values
  - **disp1**
  - **disp0**
  - **nil** (imperative, infinitive, transgressive)



---

# Grammateme *resultative* ... PDT 2.0



- applicable for all verbal forms
- resultative diathesis
- *být / mít* + passive participle
- values:
  - **res1**: (resulting state)  
*Oběd je uvařen. Je uvařeno. [The lunch is prepared.]*  
*Má uvařen oběd. Má uvařeno.*
  - **res0**:  
*Uvařil a uklidil. [He cooked (the dinner) and cleaned (the house)]*

---

# References



- Manual for Tectogrammatical Annotation  
<http://ufal.mff.cuni.cz/pdt3.0/documentation>