Functional Generative Description (FGD)

Markéta Lopatková

Institute of Formal and Applied Linguistics, MFF UK

lopatkova@ufal.mff.cuni.cz

Basic characteristics of FGD

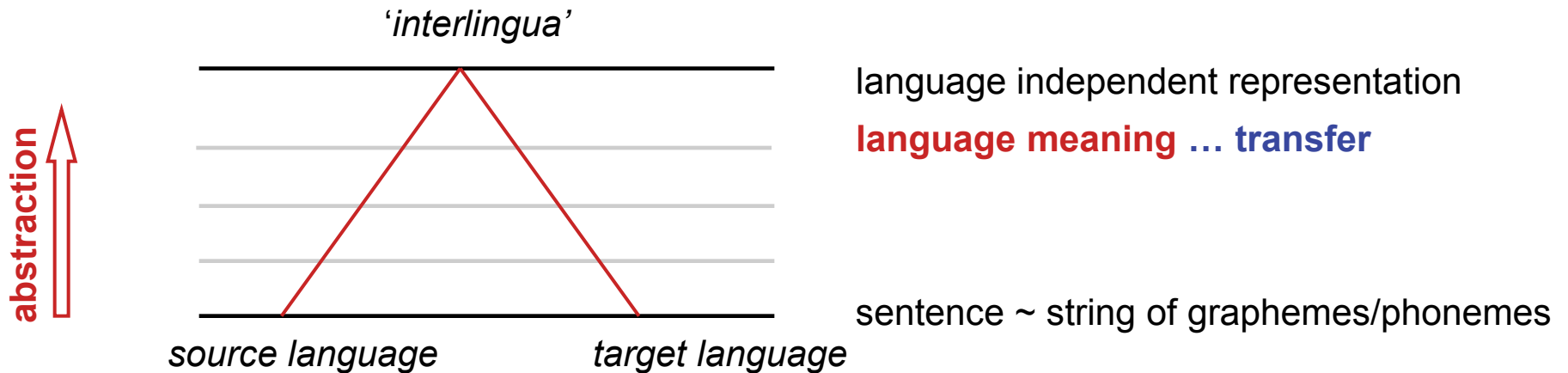- motivation: machine translation

'interlingua'

language independent representation

language meaning … **transfer**

abstraction →

sentence ~ string of graphemes/phonemes
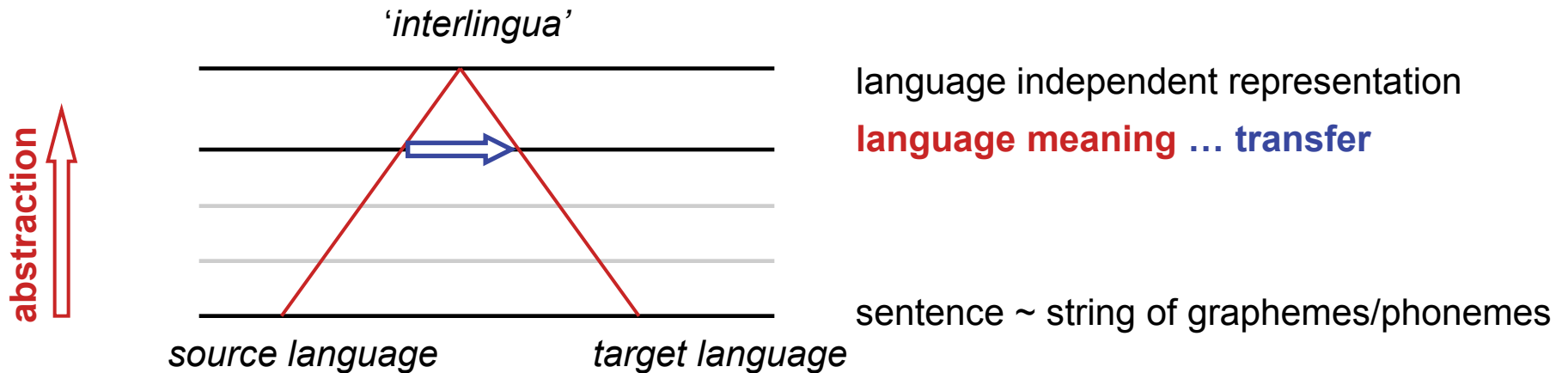
source language   *target language*

- Petr Sgall (1967) *Generativní popis jazyka a česká deklinace.* Academia, Praha
- since 1970s … together with Eva Hajičová, Jarmila Panevová
- Petr Sgall, Eva Hajičová, Jarmila Panevová (1986) *The Meaning of the Sentence in Its Semantic and Pragmatic Aspects.* Reidel, Dordrecht

Basic characteristics of FGD

- motivation: machine translation

'interlingua'

language independent representation

language meaning … transfer

sentence ~ string of graphemes/phonemes

source language   *target language*

abstraction

- Petr Sgall (1967) *Generativní popis jazyka a česká deklinace*. Academia, Praha
- since 1970s … together with Eva Hajičová, Jarmila Panevová
- Petr Sgall, Eva Hajičová, Jarmila Panevová (1986) *The Meaning of the Sentence in Its Semantic and Pragmatic Aspects.* Reidel, Dordrecht
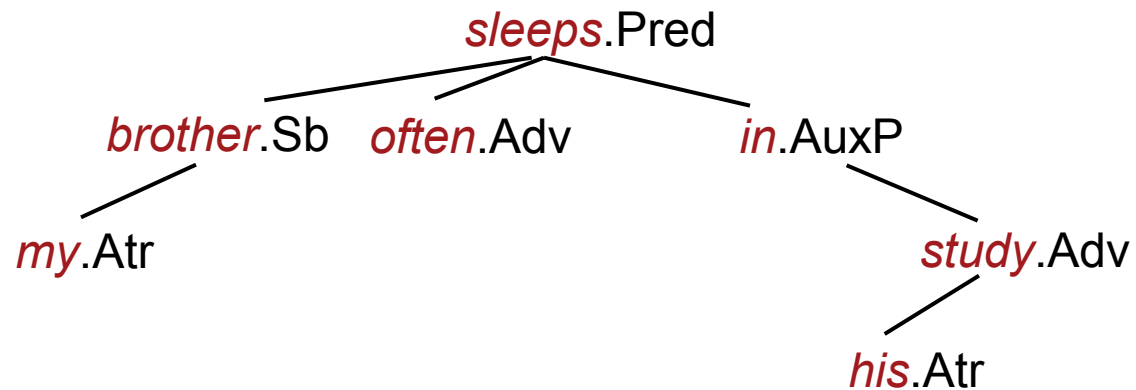
Basic characteristics of FGD (cont.)

'classical' version of FGD:

- dependency framework
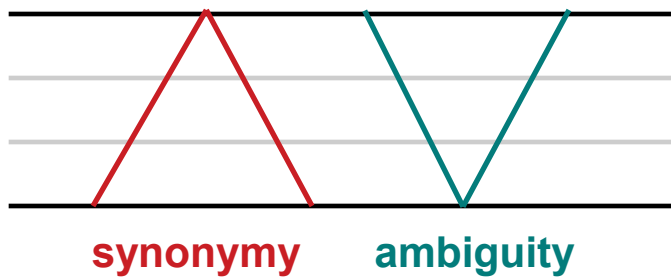 - formal description
  - suitable mathematical formalism

sleeps.Pred

brother.Sb *often*.Adv *in*.AuxP

my.Atr *study*.Adv

his.Atr

Basic characteristics of FGD (cont.)

'classical' version of FGD:

• dependency framework

• stratificational approach

synonymy   ambiguity

language meaning **~ function**

string of graphemes/phonemes **~ form**

Basic characteristics of FGD (cont.)

'classical' version of FGD:

• dependency framework

• stratificational approach

• relation between a form and its function
         / a function and its form $\Rightarrow$ *functional*

Basic characteristics of FGD (cont.)

'classical' version of FGD:

- dependency framework

- stratificational approach

- relation between a form and its function
                      / a function and its form   ⟹   *functional*

- language meaning (not cognitive content)

- *generative* vs. analytical

Basic characteristics of FGD (cont.)

- tradition of Prague Linguistic Circle
  - structural school, since 1926
  - Mathesius, Trnka, Havránek, Mukařovský, Jakobson, Trubeckoj, Karcevskij, …
- language as a system ~ langue
  vs. individual utterances ~ parol
- stress on testable criteria for distinguishing lang. phenomena
- higher layers of language description (syntax)
- topic focus articulation as a part of language meaning

Two components of FGD

- generative component
  - ~ to define all formally correct meaning representations
    (of possible sentences of a given language)
    - formalism: 1) phrase rules, phrase structure trees + functors
      2) dependency trees
    - push-down automaton

- translation component
  - ~ translating meaning representations to lower layers
    - sequence of push-down transducers plus finite-state automaton

Main pillars of FGD

- system of layers
- valency theory
- topic focus articulation
- anaphora / coreference

System of layers in FGD

meaning

deep / underlying syntax
tectogrammar

surface syntax

morphematics

morphonology

expression

phonology/phonetics

System of layers in FGD (cont.)

sentence … full representation on each layer of description

each layer ~ set of descriptions for all possible sentences
- finite set of elementary units
- finite set of operations and relations → set of complex units
- finite set of relations between sentence representations
  on a particular layer and its representations on adjacent layers

System of layers in FGD (cont.)

sentence … full representation on each layer of description

each layer ~ set of descriptions for all possible sentences
- finite set of elementary units
- finite set of operations and relations → set of complex units
- finite set of relations between sentence representations
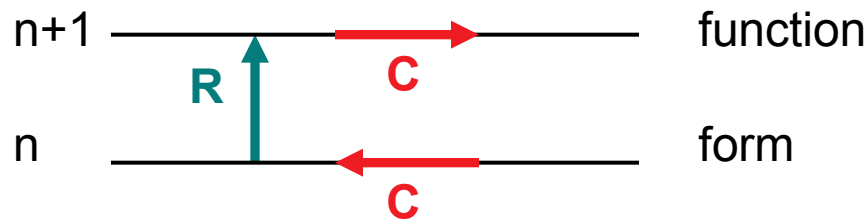 on a particular layer and its representations on adjacent layers
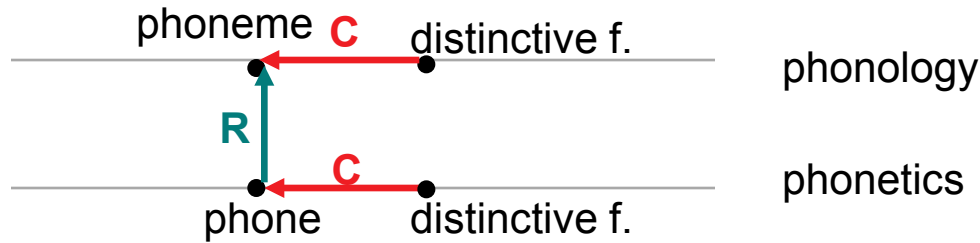

type C relations (composition): elementary units constitute complex units
     i.e., relations between units of the same layer

type R relations (representation): form-function relation
     i.e., relation between adjacent layers

System of layers in FGD (cont.)

phoneme **C** distinctive f.
                                phonology

R

        C
                                phonetics
phone   distinctive f.

layer of phonetics
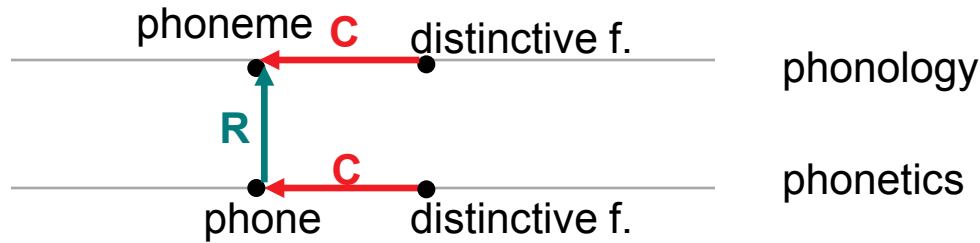   <u>distinctive features</u> … elementary units
    <u>phones</u> (~ a speech sound) … complex units

    suprasegmental units … prosody, intonation

System of layers in FGD (cont.)

phoneme **C** distinctive f.

R

C

phone distinctive f.

phonology

phonetics

layer of phonetics

<u>distinctive features</u> … elementary units

<u>phones</u> (~ a speech sound) … complex units

suprasegmental units … prosody, intonation

layer of phonology
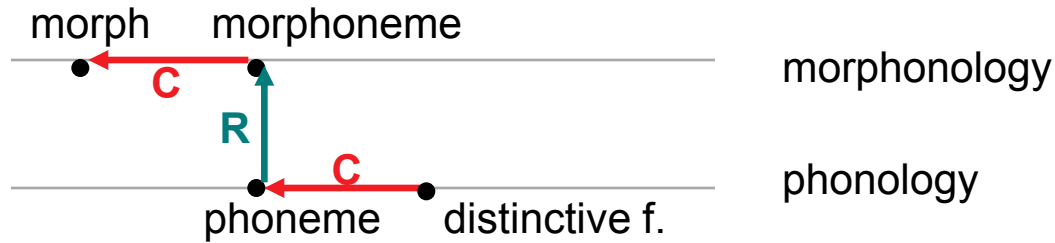
<u>distinctive features</u> … elementary units

<u>phonemes</u> (~ 'smallest' units that distinguish meaning) … complex units

asymmetry … allophones ~ variants of a single phoneme

language dependent   (sing vs. sin)
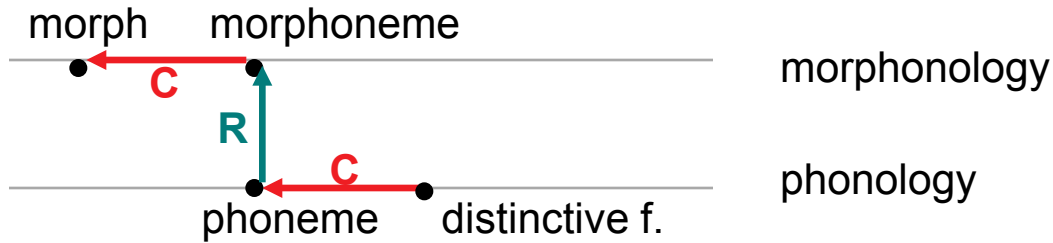
System of layers in FGD (cont.)

morph   morphoneme

morphonology

R

C

phonology

phoneme   distinctive f.

layer of morphonology

morphoneme ~ set of phoneme variants   e.g. *k|c|č|.k* in *"matka"*

oo|ee   in *"foot"*

System of layers in FGD (cont.)

morph   morphoneme

morphonology

phoneme   distinctive f.

phonology

layer of morphonology

morphoneme ~ set of phoneme variants   e.g. *k|c|č|.k* in *"matka"*

morph ~ string of morphonemes

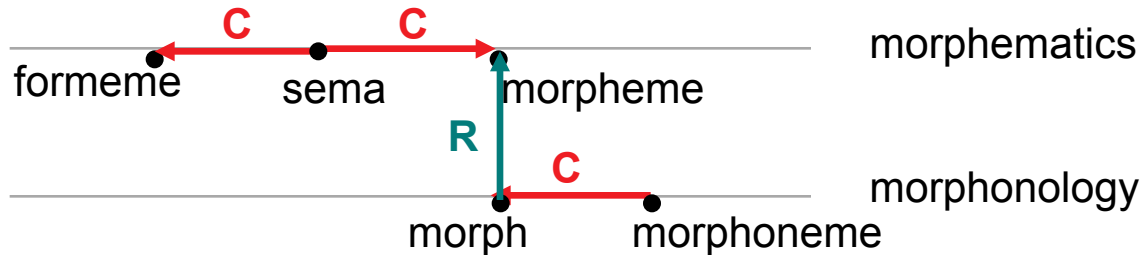lexical variants *(matk, matc, matč, mat.k)* ...   4 allomorphs

mat(k|c|č|.k)   1 morph

lexical variants (foot, feet) ...   2 allomorphs

f(oo|ee)t   1 morph

System of layers in FGD (cont.)

layer of morphematics

    morpheme ~ the smallest component that has semantic meaning

- <u>lexical morpheme</u>
  - roots
        e.g. lex. morpheme for *matka* consists of 4 allomorphs *(matk, matc, matč, mat.k)* ;
           for *to write (writ, wrot);* for *leaf (leaf, leav)*
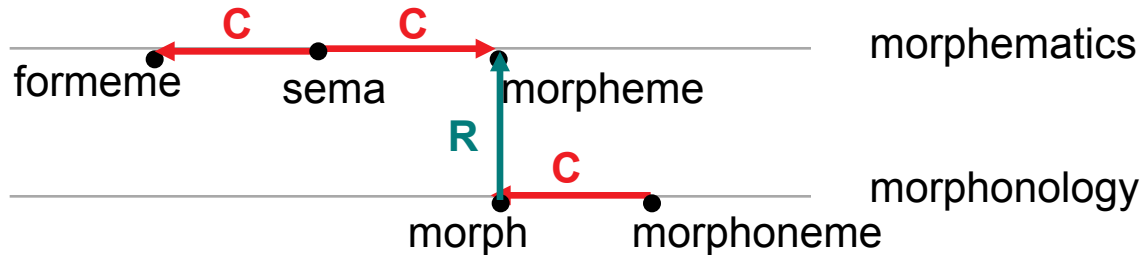 - derivational morphemes (affixes: prefixes, infixes, suffixes, …)
        il- (as in *illegal*), *non-* (as in *nonproblematic*)
        -ly (as in *legally*), *-ess* (as in *actress*)

System of layers in FGD (cont.)

layer of morphematics

morpheme ~ the smallest component that has semantic meaning

- lexical morpheme

- grammatical morpheme

    - inflectional affixes   e.g. Cz: suffixes   nouns: case, gender, number, …
                                        verbs: gender, number, tense, voice,

                            Eng: suffixes   nouns: plural *-s*
                                        verbs: past tense *-ed* , continuous *–ing*

System of layers in FGD (cont.)

layer of morphematics
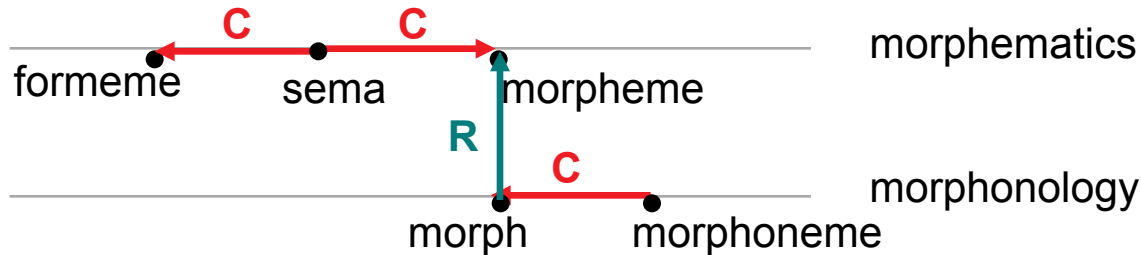
morpheme ~ the smallest component that has semantic meaning

- <u>lexical morpheme</u>
- <u>grammatical morpheme</u>
    - inflectional affixes   e.g. Cz: suffixes   nouns: case, gender, number, …
      verbs: gender, number, tense, voice,
      Eng: suffixes nouns: plural *-s*
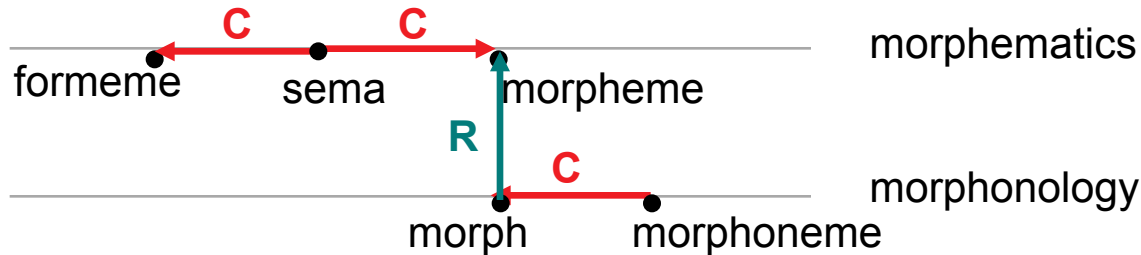   verbs: past tense *-ed* , continuous *–ing*

e.g.   *matk + a*   *boy + s*
       koup + il   *play + ed*

- sema … a combination of grammatical morphemes that characterize
  a lexical morpeme (or strings of lexical morphemes)

System of layers in FGD (cont.)

layer of morphematics

    morpheme ~ the smallest component that has semantic meaning
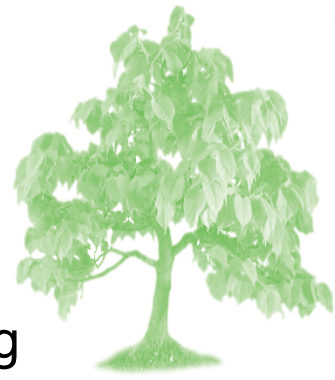
- <u>lexical morpheme</u>
- <u>grammatical morpheme</u>
- <u>formeme:</u>

      sequence of morphs realizing a single tagmeme / sentence member

      lexical f., case f. (i.e., prep+case), conjunction formemes (i.e., conj+verb mood)

    Cz: *vysoká škola;   lamp+a;   na+lavic+i;   chod+í;   bude+chodit*
    Eng: *white-collar;   lamp;   on+ table;   walk+s;   will+be+walk+ing*

System of layers in FGD (cont.)

morpheme ~ the smallest component that has semantic meaning

Czech:

nejneobhospodařovatelnější
nej - ne - obhospodař - ova - teln - ějš - í

most - non - cultivate - [iter] - [adj] - [super] - [sg+nom+fem | sg+acc+neutr | … pl+voc+masc]

(a single root)   23 combinations

Basque:
dakartzat   [I bring them]
da – kar – tza – t
present -   - plural - Sb
  root =verb *ekarri* [bring]

Hungarian:
fiaiéi
the root "fi-" means "son", the subsequent four vowels are all separate suffixes, and the whole word means "[properties] of his/her sons"

Two layers of syntax in FGD

tree-based dependency structure
- nodes for tagmemes / sememes (complex symbols)
- edges labeled with a type of a respective syntactic relation

The layer of surface syntax

layer of surface syntax
 ~ structure of sentence members / *tagmemes*
  incl. lexical and morphological information

3 types of elementary units:
 • lexical: units from a dictionary
 • morphological: set of morphological features ~ *tags*
  (a pair of) trousers … sema - plural
 • syntactic: *sentence members* / tagmemes
  subject, object, attribute, adverbial, complement,…

 +

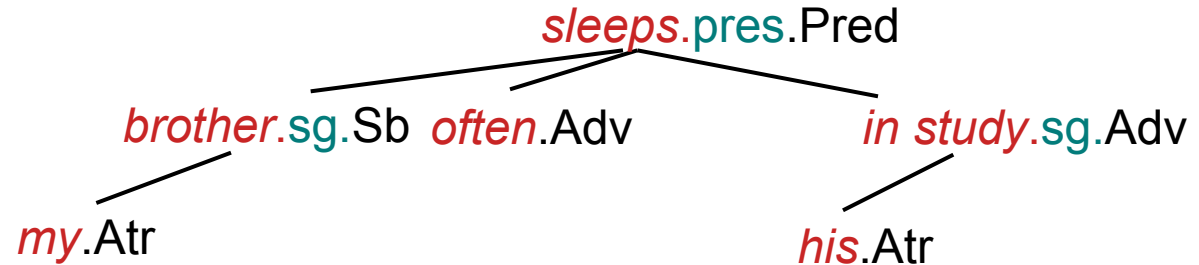surface word order … linear ordering of tree nodes

The layer of surface syntax

My brother often sleeps in his study.

```
                              sleeps.pres.Pred
                    _____|_____
                   |               |               |
           brother.sg.Sb       often.Adv      in study.sg.Adv
               |                                    |
            my.Atr                               his.Atr
```
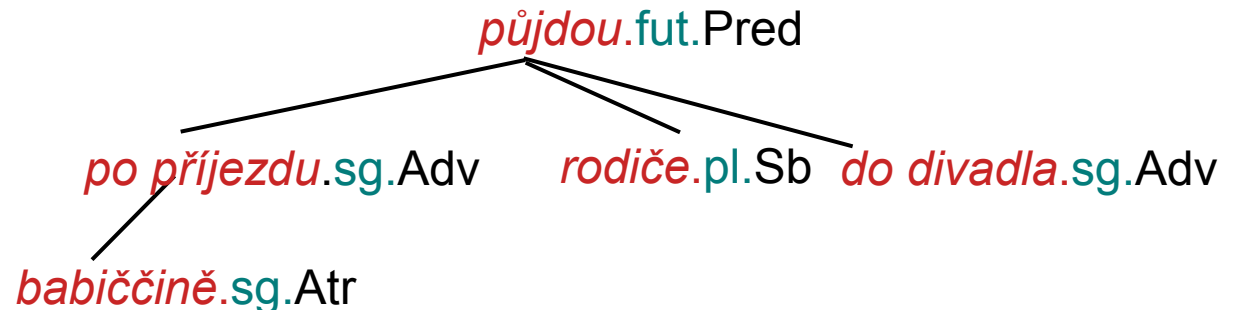
Po babiččině příjezdu půjdou rodiče do divadla.
[After grandma's arrival the parents will go to the theatre.]

```
                              půjdou.fut.Pred
                    _____|_____
                   |           |               |
          po příjezdu.sg.Adv  rodiče.pl.Sb  do divadla.sg.Adv
               |
        babiččině.sg.Atr
```

The layer of deep syntax

~ meaning of a sentence:
   semantemes: lexical (autosemantic) words, their lexical and
   morphological features and mutual relations

terminology: deep / underlying / tectogrammatical representation (TR)

3 basic types of elementary units:
 • lexical: units from a (tectogrammatical) dictionary
 • morphological: *grammatemes*

   meaning of individual morphological categories
   (a pair of) trousers … singular
   denominating (*pojmenovávací*)

                          vs. co*rrelating* (*usouvztažňující*) categories
 • syntactic: types of relation, *functors* and *subfunctors*
      Actor, Patient, Addressee, … local, temporal modifications …
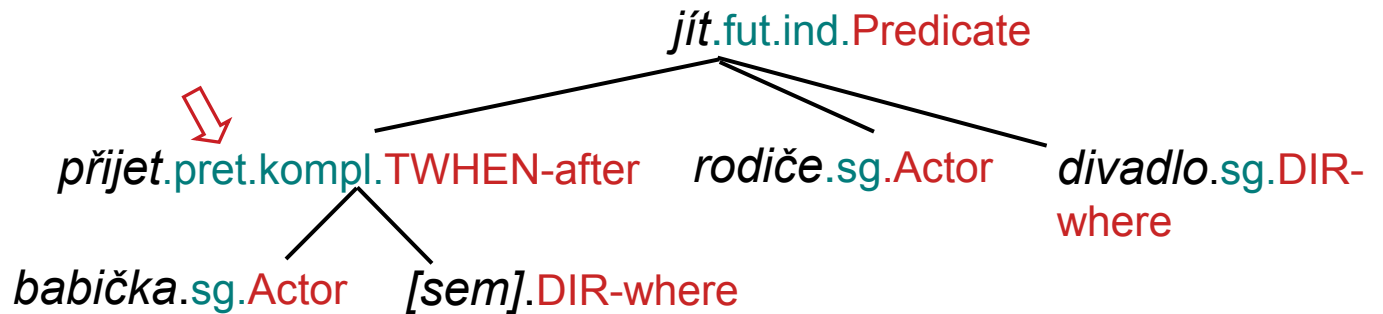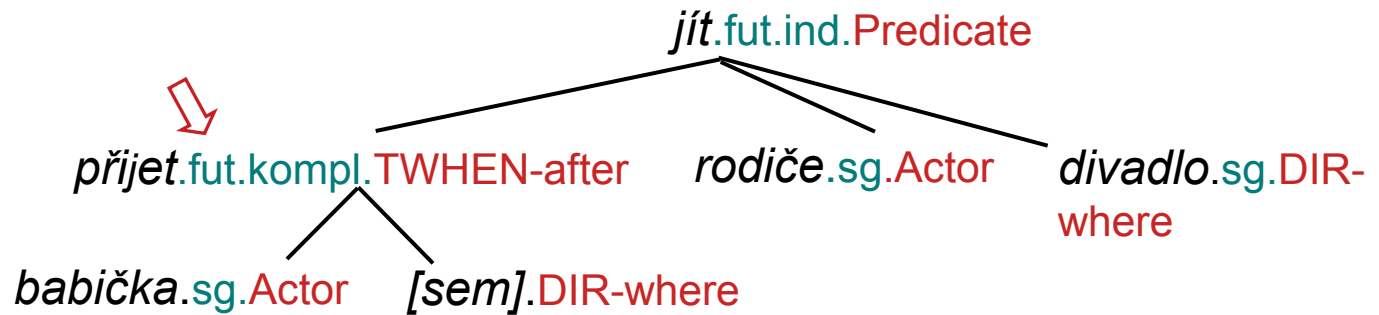
The layer of deep syntax

- *deep word order*
    - increasing communicative dynamism:
      word order reflects "relative degree of importance in comparison with other expressions in the sentence […]"
    - topic focus articulation
- condition of projectivity !!!

The layer of deep syntax

Po babiččině příjezdu půjdou rodiče do divadla.
[After grandma's arrival the parents will go to the theatre.]

jít.fut.ind.Predicate

⇦ *přijet*.fut.kompl.TWHEN-after   *rodiče*.sg.Actor   *divadlo*.sg.DIR-where

babička.sg.Actor   *[sem]*.DIR-where

jít.fut.ind.Predicate

⇦ *přijet*.pret.kompl.TWHEN-after   *rodiče*.sg.Actor   *divadlo*.sg.DIR-where

babička.sg.Actor   *[sem]*.DIR-where

The layers of deep and surface syntax

- different sets of elementary units
  - 'morphological' vs. tectogrammatical lemma
  - morphological categories vs. grammatemes
  - surface sentence members vs. functors

⟹

- sentence members / *tagmemes* vs. *semantemes*
  - only autosemantic / lexical words at TR
    – modal verbs
      Peter <u>wants to</u> attend the concert. [to attend + volitive]
      Charles <u>has to</u> pass the exam. [to pass + debitive]
    – nominalization
      After grandma's <u>arrival</u> … → [to arrive]
    – active / passive verbs → [active form]
      Tato krásná kniha <u>byla vydána</u> nakladatelstvím Albatros.
      [This beautiful book was published by the Albatros publishing house.]

The layers of deep and surface syntax

- different sets of elementary units
  - 'morphological' vs. tectogrammatical lemma
  - morphological categories vs. grammatemes
  - surface sentence members vs. functors

⇒

- sentence members / *tagmemes* vs. *semantemes*
  - only autosemantic / lexical words at TR
  - completeness of the representation – (surface) ellipses are restored
    omitted surface subject (Czech: pro-drop language):
    - Czech: *Vidíš bratra? Vidím. Přichází.*
      - → [Ty] vidíš bratra? [Já] vidím [ho]. [On] přichází [sem].
    - Russian: *Ты видел брата? Вижу [его]. Идёт.*
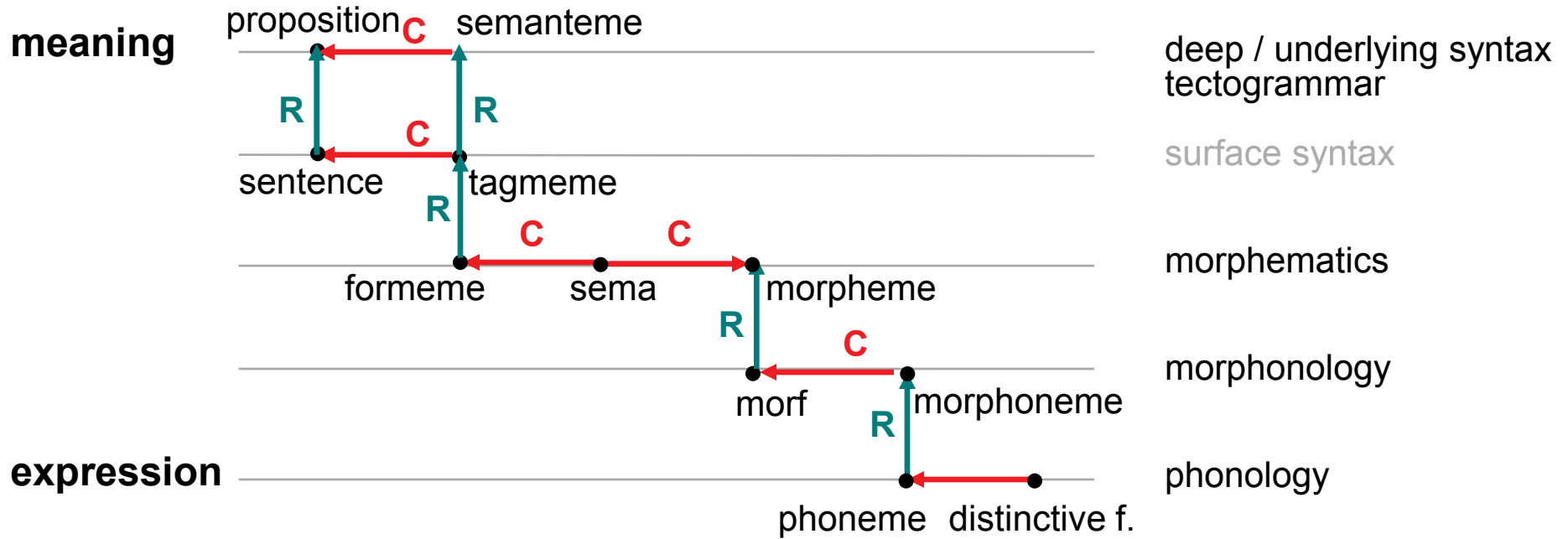   - Spanish: *¿Ves este tronco?* [(Do) **you** see this log?]

The layers of deep and surface syntax

- different sets of elementary units
    - 'morphological' vs. tectogrammatical lemma
    - morphological categories vs. grammatemes
    - surface sentence members vs. functors

⇒

- sentence members / *tagmemes* vs. *semantemes*
    - only autosemantic / lexical words at TR
    - completeness of the representation – (surface) ellipses are restored
- surface vs. deep word order:
    - TR: projective trees
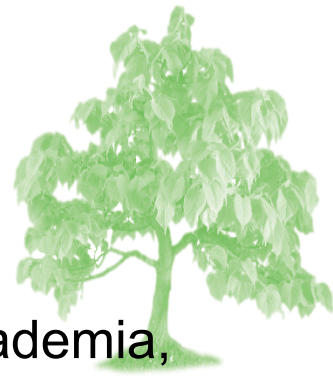   - increasing communicative dynamism

System of layers in FGD

meaning

proposition **C** semanteme

deep / underlying syntax
tectogrammar

R   **C**   **R**

surface syntax

sentence   tagmeme

R

C   **C**

formeme   sema   morpheme

morphematics

R

C

morf   morphoneme

morphonology

R

expression

phonology

phoneme   distinctive f.

References

- Sgall, P. (1967) *Generativní popis jazyka a česká deklinace.* Academia, Praha
- Sgall, P., Hajičová, E., Panevová, J. (1986) *The Meaning of the Sentence in Its Semantic and Pragmatic Aspects.* Reidel, Dordrecht.
- Sgall, P. a kol. (1986) *Úvod do syntaxe a sémantiky.* Academia, Praha
- Hajičová, E., Panevová, J., Sgall, P. (2002) *Úvod do teoretické a počítačové lingvistiky*, sv. I. Karolinum, Praha.
- Petkevič, V. (1995) A New Formal Specification of Underlying Structure. *Theoretical Linguistics* Vol.21, No.1.