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

\[\text{`interlingua'}\]

language independent representation

language meaning \ldots transfer

\text{sentence} \sim \text{string of graphemes/phonemes}

• Petr Sgall (1967) Generativní popis jazyka a česká deklinace. Academia, Praha
• since 1970s \ldots together with Eva Hajičová, Jarmila Panevová
Basic characteristics of FGD

- motivation: machine translation

```

<table>
<thead>
<tr>
<th>'interlingua'</th>
<th>language independent representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>source language</td>
<td>target language</td>
</tr>
<tr>
<td>sentence ~ string of graphemes/phonemes</td>
<td></td>
</tr>
</tbody>
</table>
```

- Petr Sgall (1967) *Generativní popis jazyka a česká deklinace*. Academia, Praha
- since 1970s ... together with Eva Hajičová, Jarmila Panevová
Basic characteristics of FGD (cont.)

'classical' version of FGD:

• dependency framework
  • formal description
  • suitable mathematical formalism
Basic characteristics of FGD (cont.)

'classical' version of FGD:

- dependency framework
- stratificational approach

language meaning $\sim$ function

string of graphemes/phonemes $\sim$ form

synonymy ambiguity
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
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
• 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

<table>
<thead>
<tr>
<th>meaning</th>
<th>expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>deep / underlying syntax</td>
<td>phonology/phonetics</td>
</tr>
<tr>
<td>surface syntax</td>
<td></td>
</tr>
<tr>
<td>morphematics</td>
<td></td>
</tr>
<tr>
<td>morphonology</td>
<td></td>
</tr>
</tbody>
</table>

**PDT – FGD**

Lopatková
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 \( \rightarrow \) 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

\[
\begin{align*}
 & n+1 \quad \text{function} \\
 & n \quad \text{form}
\end{align*}
\]

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.)

layer of phonetics

**distinctive features** … elementary units
**phones** (~ a speech sound) … complex units

suprasegmental units … prosody, intonation

---

layer of phonetics
System of layers in FGD (cont.)

layer of phonetics
- distinctive features ... elementary units
- phones (~ a speech sound) ... complex units
- suprasegmental units ... prosody, intonation

layer of phonology
- distinctive features ... elementary units
- phonemes (~ ‘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.)

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.)

layer of morphonology

- morphoneme ~ set of phoneme variants  e.g. k|č|č.k in "matka"
- morph ~ string of morphonemes
  - lexical variants (*matk*, *matc*, *matč*, *mat.k*) ... 4 allomorphs
  - *mat(k|č|č.k)* ... 1 morph
- *lexical variants* (*foot*, *feet*) ... 2 allomorphs
  - *f(oo|ee)t* ... 1 morph

PDT – FGD

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

- **lexical morpheme**
  - 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, …
  - Eng: suffixes verbs: gender, number, tense, voice, nouns: plural -s
      verbs: past tense -ed , continuous –ing

morphematics
morphophonology
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  Eng: suffixes
    nouns: case, gender, number, …  verbs: gender, number, tense, voice,
    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

- lexical morpheme
- grammatical morpheme
- formeme:
  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+lavici; chodí; bude+chodit
Eng: white-collar; lamp; on+ table; walk+s; will+be+walking
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:

fiaiēi
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.

Po babiččině příjezdu půjdou rodiče do divadla.
[After grandma's arrival the parents will go to the theatre.]
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. *correlating* *(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.]
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 *wants to attend the concert.* [to attend + volitive]
      - Charles *has to pass the exam.* [to pass + debitive]
    - nominalization
      - After grandma's *arrival* … → [to arrive]
    - active / passive verbs → [active form]
      - *Tato krásná kniha byla vydána 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):


- 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

semanteme

c


tectogrammar

deep / underlying syntax

sentence

c

tagmeme

c

surface syntax

formeme

c

morpheme

c

morphematics

sema

c

morphonology

expression

morphoneme

morf

c

phonology

phoneme

distinctive f.
References

- Sgall, P. a kol. (1986) *Úvod do syntaxe a sémantiky*. Academia, Praha