

Government Models for Cross-Lingual Transformations

Elena B. Kozerenko
kozerenko@mail.ru

Institute for Informatics Problems
Russian Academy of Sciences,
Moscow, Russia

Main objectives of the research

- Establishment of cross-language matches and *inter-structural* synonymy
- Creating presentations for computational models of *cross-lingual correspondencies* which take into account the transformations of language structures indispensable for the adequate translation

Valence: the head-dependent relation

- *Subcategorization* is understood as the enumeration of the expected categorial features of the language objects co-occurring with the head element
- *Government* specifies both categorial and morphosyntactic features of the dependent language objects

The two presentation mechanisms

- The first mechanism is based on dependency grammar, and it is applied for the design of the multilingual knowledge extraction systems and for the *interlingua*-based machine translation
- The second mechanism takes into account both constituency and dependency relations, and it is used for parallel texts alignment and for the *transfer*-based machine translation system development



The extended semantic networks ESN)

- ESN represent highly embedded structures of natural language.
- The basic structural element of the ESN is the named N -ary predicate, called "fragment".
- The ESN basis is the set of vertices (V), from which the following elementary fragments are comprised:
 $V_0(V_1, V_2, \dots, V_k/V_{k+1})$, where
 $V_0, V_1, V_2, \dots, V_k, V_{k+1} \quad V, k > 0$



The transformations result in the shift of the government models

Predicates are "support" elements:
shooting the ducks from the rifle –
strel'ba po utkam iz ruzh'ia /
strelyaiuschii po utkam iz ruzh'ia
shooting (N, process) at ducks from
rifle / shooting (Part) at ducks
from rifle
/ strelyaiia po utkam iz ruzh'ia
shooting (AdvPart) at ducks from
rifle



Nominalization

The Russian language is about 35% more "nominative" than English:

- *In vacuum molecules have large space in which to move* (V).
- *V vakuume molekuly imeiut bol'shoe prostranstvo dlia dvizhenia.* (Rus.-translit)
- In vacuum molecules have large space for movement (N)



The most productive types of verbal-nominal transformations

- Correlate with the following functional values
- Adverbial modifier of purpose and of consequence, expressed by infinitive (58%)
 - Composite predicate with the infinitive (be + infinitive) (51%)
- This statistics is taken into account in the multivariant cognitive transfer rules

Statistical portraits of translatable language objects

