WELCOME TO THE CLUB: DESIGNING THE INVENTORY OF SEMANTIC ROLES FOR ADJECTIVES

Olga Lyashevskaya (HSE / IRL RAS) olesar@yandex.ru
Egor Kashkin (IRL RAS) egorkashkin@rambler.ru
SEMANTIC ROLES AND THEIR INVENTORIES

- A key concept in both theoretical and computational tasks (semantic-syntax interface → natural language understanding).
- However: a huge diversity in the inventories of semantic roles.
SEMANTIC ROLES AND GRAMMATICAL CLASSES

- The existing inventories mostly deal with verbs.
- Arguments of nouns or adjectives?

  *dyra v polu* ‘a hole in the floor’, *kofe s molokom* ‘coffee with milk’, *dalekij ot Moskvy* ‘distant from Moscow’, 
  *izvestnyj svoimi publikacijami* ‘famous for his / her publications’

- The same role inventory for verbal and non-verbal roles or different inventories?
OBJECTIVE

- Building a role inventory for Russian adjectives.
- This has been beyond the existing research so far (cf. Arkhangelskiy et al. 2010; Kustova 2007, 2009; Rakhilina et al. 2010 on the semantics and morpho-syntax of adjectives; Shelmanov, Smirnov 2014; Kuznetsov 2016 on SRL).
- What is the empirical evidence to rely on?
- What is the “gold standard” to avoid arbitrary suggestions and to make the results verifiable for computational tasks?
OUR APPROACH

- Collecting expert judgements as the benchmark for assessing our decisions.
- Instead of building a theory from scratch, we propose to form the opinion of the community and to see the limits of possible divergences.
- A common practice in computational linguistics: checking to what extent the output of a machine could approximate human judgements (cf. sentiment analysis, machine translation, semantic similarity).

Some deviations in our approach:

- The need to assess not the output of a system but rather the training data.
- Judgements about semantic roles require linguistic background of the assessors.
Verbal Roles: Russian FrameBank

- The inventory of verbal roles as the starting point.
- Its source: Russian FrameBank (www.framebank.ru)

+ The task to create a role inventory good for all purposes is rather vague, so we need to adopt a particular framework.

FrameBank and its role inventory (Lyashevskaya, Kashkin 2015):

✔ Russian lexical constructions mostly for verbs (ca. 4000) & corpus examples (ca. 50000), 91 semantic roles.

✔ The roles correlate with the semantic classification of the lexicon & prototype and periphery (e.g. different types of Agent for destruction vs. speech vs. motion).

✔ Hierarchy of roles, flexible search options.
DESIGN OF THE SURVEY

- 20 native Russian speakers, either professional linguists or students of linguistic departments.
- Score-assignment test.
- The task: to rate the similarity between the target pair ADJECTIVE – ITS ARGUMENT and the control pair VERB – ITS ARGUMENT, the scale is from 1 to 7.
  + A free answer option to suggest a better variant (however it was not popular).
IV. Близкий

IV-b. Автор статьи высказывает близкие НАМ идеи.

ИВАН ПЕТРОВИЧ болеет уже несколько недель. *

ТЕТКА думает, что ему придется уехать. *

ПРОДАВЕЦ режет сыр на тонкие куски. *

Лиза открыла СВЕТЕ дверь. *
CONTENTS OF THE SURVEY


- 10 predicative and 6 attributive constructions.

- 3-4 control pairs VERB – ITS ARGUMENT after each target sentence.

- All the questions for each participant, no time limit.

- NB no universal correspondences between cognate verbs and adjectives, cf. Ja gotov pomoč’ tebe ‘I am ready to help you’ vs. *On gotovit men’a pomoč’ tebe, expected meaning ‘He is making me ready to help you’; Ja vinovat pered Vami ‘I am guilty towards you’ vs. *On obvinil men’a pered Vami, lit. ‘He accused me towards you’
RESULTS: AN OVERVIEW
VERBAL ROLES → ADJECTIVAL ROLES

The roles of adjectival arguments can be adequately tagged using the inventory of roles describing verbal arguments.

6 of the 16 experimental blocks with an example evaluated from 5 to 7 on average, 7 blocks with an average mark from 4 to 5.

Natal'ja Jur'jevna byla očen' blizka s otcom ‘Natalya Jurievna was very close to her father’ – Kol'a družit s Natašej ‘Kolya is friends with Natasha’ (COUNTER-AGENT OF SOCIAL RELATION; 6.4).

Pet’a sil’en v matematike ‘Petya is good (lit.: strong) at mathematics’ – On vseh obošel v učeb ‘He left everyone behind in his studies’ (SPHERE; 6).

Samymi ščedrymi na novogodnie podarki v etom godu stanut rukovoditeli rossijskih kompanij i gosslužaščie ‘This year the most generous in giving New Year’s presents will be the managers of Russian companies and civil servants’ – Deduška dal rebenku konfetu ‘Grandfather gave a sweet to a child’ (PATIENT; 4.8).
**PREDICATIVE VS. ATTRIBUTIVE USES**

- An issue for further discussion.
- Many predicative uses with high scores, but:
  - Some attributive uses are also there.
  - There are plenty of predicative uses with low scores, and the factors causing difference between the examples anyway have to be accounted for.
ROLE INVENTORY FOR ADJECTIVES

- A draft, not all the possible adjectival constructions and semantic classes in the survey.
- 14 roles (overall rating 4 or higher).
- Verbal roles can be in principle transferred to adjectival constructions.
- The inventory can be enlarged using the principles highlighted by the experiment.
# A Subset of the Role Inventory for Adjectives

<table>
<thead>
<tr>
<th>Semantic role</th>
<th>Adjective and morpho-syntactic pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>Counter-Agent of social relation</td>
<td><em>blizkij s + Sins</em> ‘close to smb (e.g., a friend)’, <em>vinovatyj pered + Sins</em> ‘guilty towards smb’</td>
</tr>
<tr>
<td>Patient</td>
<td><em>svobodnyj ot + Sgen</em> ‘free from sth’, <em>schedryj na + Sacc</em> ‘generous in sth (e.g., gifts)’</td>
</tr>
<tr>
<td>Content of thought</td>
<td><em>gotovyj Vinf</em> ‘ready to do sth.’</td>
</tr>
<tr>
<td>Content of action</td>
<td><em>gotovyj Vinf</em> ‘ready to do sth.’ (a competing role in the results of the survey)</td>
</tr>
<tr>
<td>Content of utterance</td>
<td><em>schedryj na + Sacc</em> ‘lavish with sth (speech etc.)’</td>
</tr>
<tr>
<td>Addressee</td>
<td><em>vinovatyj pered + Sins</em> ‘guilty towards smb’ (a competing role in the results of the survey)</td>
</tr>
<tr>
<td>Beneficiary</td>
<td><em>svobodnyj dl’a + Sgen</em> ‘free for smb/sth’</td>
</tr>
<tr>
<td>Role</td>
<td>Example</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------------------------------------------------------</td>
</tr>
<tr>
<td>Location</td>
<td><em>blizkij k + Sdat / ot + Sdat</em> ‘close to sth.’</td>
</tr>
<tr>
<td>Point of destination</td>
<td><em>blizkij k + Sdat</em> ‘close to sth.’ (a competing role in the results of the survey)</td>
</tr>
<tr>
<td>Sphere</td>
<td><em>sil’nyj v + Sloc</em> ‘strong in sth.’</td>
</tr>
<tr>
<td>Social environment</td>
<td><em>izvestnyj v + Sloc</em> ‘famous among smb’</td>
</tr>
<tr>
<td>Goal</td>
<td><em>svobodnyj dl’a + Sgen</em> ‘free for smb/sth’ (a competing role in the results of the survey)</td>
</tr>
<tr>
<td>Feature</td>
<td><em>sil’nyj Sins</em> ‘strong with sth.’</td>
</tr>
<tr>
<td>Property of Reason for mental state</td>
<td><em>izvestnyj Sins</em> ‘famous with sth.’</td>
</tr>
</tbody>
</table>
SEMANTICS VS. MORPHO-SYNTAX

- Higher scores for closer semantic classes with the same morpho-syntactic marking.

  *Pet'a sil'en v matematike* ‘Petya is good (lit.: strong) at mathematics’ – *On vseh obošel v učebe* ‘He left everyone behind in his studies’ (SPHERE; 6) vs. *Ja živu v Moskve* ‘I live in Moscow’ (PLACE; 2.4).

  *Etot žurnalist izvesten svoimi razoblačitel'nymi publikacijami* ‘This journalist is famous for his unmasking publications’ – *Pavel porazil vseh dlinnymi volosami* ‘Pavel amazed everyone with his long hair’ (PROPERTY OF THE REASON FOR EMOTIONAL STATE; 5.2) vs. *Ivan razbil okno palkoj* ‘Ivan broke the window with a stick’ (INSTRUMENT; 3.2).
SEMANTICS VS. MORPHO-SYNTAX

- For several semantically adequate candidates: similarity in the morpho-syntactic pattern.

  Avar’ijnye vyhody i prohody dolžny byt’ svobodny ot ručnoj kladi ‘Emergency exits and passages must be free of hand luggage’ — My očistili čerdak ot hlama ‘We cleared the attic of junk’ (PATIENT; 6) vs. My ubrali al'bomy s polok ‘We removed the albums from the shelves’ (PATIENT; 3.1)

  Policija zaderžala narkotorgovca, izvestnogo v opredelennykh krugakh pod klíčkoi “Korotyška” ‘The police arrested a drug pushed known as “Shorty” in criminal circles’ — Soobščenie posejalo paniku v r'adah vraga ‘The message spread panic among the ranks of the enemy’ (SOCIAL CIRCLE; 4.9) vs. Ivan znaet, čem končilos’ delo ‘Ivan knows how the things have finished’ (SUBJECT OF MENTAL STATE; 3.2).
**Statative vs. Dynamic Semantics: BLIZKIJ ‘close’**

- Point of destination - like argument:
  
  *Ekspedicija obsledovala blizkie k Saransku sela* ‘The expedition explored the villages close to the city of Saransk’

- Initial point - like argument:
  
  *V tu že subbotu, rannim večerom, uspel Aleksandrov sbegat' s kon'kami na nebol'šoj, no ujutnyj i blizkij ot doma katok Patriarših prudov* ‘On the same Saturday, early in the evening, Aleksandrov had time to run with his skates to the small but cosy skating-rink of Patriarshie ponds, which was close to (lit.: from) his home’

- Verbal candidates for assessment:
  
  - Point of destination (*podojti k domu* ‘approach the house’)
  - Initial point (*otojti ot dveri* ‘move away from the door’)
  - Location (*hodit' u reki* ‘go along the river’).
## Stative vs. Dynamic Semantics: *blizkij* ‘close’

<table>
<thead>
<tr>
<th></th>
<th>Point of dest.</th>
<th>Init. point</th>
<th>Loc.</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>blizkij + k</em> (Point of dest. - like)</td>
<td>4.3</td>
<td>2.6</td>
<td>4.7</td>
</tr>
<tr>
<td><em>blizkij + ot</em> (Init. point - like)</td>
<td>3.6</td>
<td>3.3</td>
<td>4.1</td>
</tr>
</tbody>
</table>

- Stative nature of adjectives => disposition to a stative role of Location.
- Morpho-syntactic factor in the choice between two dynamic roles.
BENEFITS FOR VERBAL ROLES

- Hierarchy of verbal roles in FrameBank.
- Mainly on semantic grounds.
- Certain verification in this experiment.
- High average scores for two verbal examples => the roles are semantically similar.
- Significantly different scores for two verbal examples => the roles are semantically different.
- Low scores for two verbal examples => no conclusion.
BENEFITS FOR VERBAL ROLES
BENEFITS FOR VERBAL ROLES

Natal’ja Jur’jevna byla očen’ blizka s otcom ‘Natalya Jurievna was very close to her father’

1). On vospityvaet trjoh synovej ‘He brings up three sons’ (SUBJECT OF SOCIAL RELATION; 2)

2). Kol’a družit s Natašej ‘Kolya is friends with Natasha’ (COUNTER-AGENT OF SOCIAL RELATION; 6.4)

3). Mit’a podrals’a s Lešej ‘Mitya fought with Lyosha’ (COUNTER-AGENT; 3.4)

4). Krest’janin rubit drova ‘The peasant is chopping firewood’ (AGENT; 1.2)

- Counter-Agent of social relation can possibly be a subtype of Counter-Agent, which supports the hierarchy in FrameBank.
- However, Counter-Agents differ much from Agents, contrary to the hierarchy.
CONCLUSIONS

- Adjectives and verbs can in principle share the same role inventory, the inventory of adjectival roles is at least a subset of the inventory of verbal roles.
- Some principles which could govern assigning semantic roles to adjectives.
- Semantic similarity + morpho-syntactic similarity for semantically adequate candidates + stative semantics.
- Much work to do in elaborating the role inventory for adjectives and in studying their valency patterns in comparison to verbs.