Semantic Analysis of Russian Prepositional Constructions

Kirill Boyarsky, Anastasia Kozlova
ITMO University

Victor Zakharov, Anastasia Golovina
Saint Petersburg University

Russian prepositional constructions: the project

Prepositional meaning: complex

Our interpretation: **relation between content words** in a given prepositional construction

G.A. Zolotova, Syntactic Dictionary: **syntaxeme** = preposition + case government + contextually dependent relation conveyed

Goal of the project: 3-level hierarchical **ontology of Russian prepositions**

A **node** in prepositional ontology: a **syntaxeme**

  e.g. *в* 'in' + Accusative case (Directive meaning)

Через 'through, across, in, after'

Zolotova: 5 meanings (relation types)

Corpus study: Russian National Corpus, expert evaluation of contexts with через

Result: 3 meanings present in real texts

Automated prepositional construction extraction and classification: possible, preferable

Distribution of через meanings in RNC samples

RNC subcorpora
- Main
- Newspaper
- Spoken
- Poetic
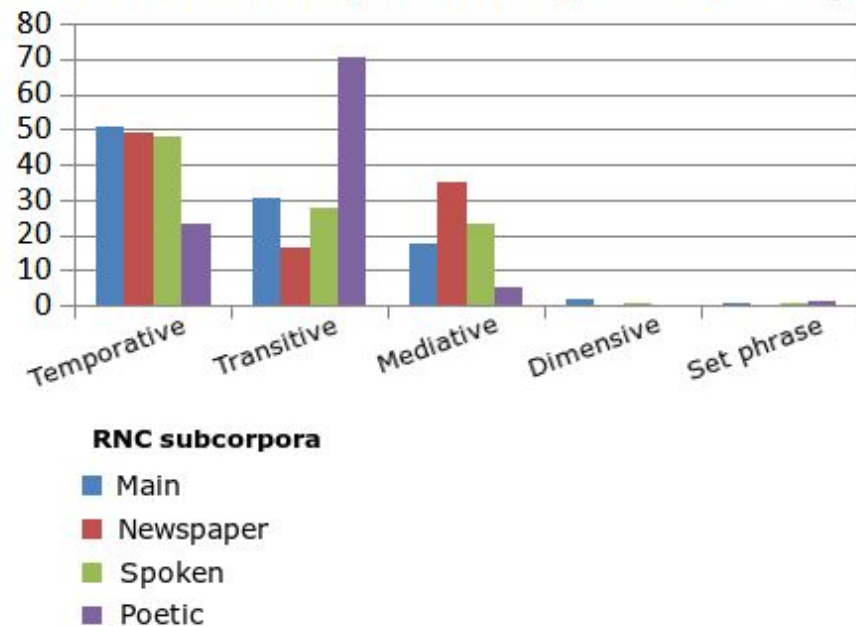
SemSin: semantic and syntactic parser

Builds **dependency tree** for each sentence + detects **types of relations** between nodes

Components:

1) morphological analyzer: tokenization + morphology

2) semantic dictionary (V.A. Tuzov): semantic classification of content words

3) lexical analyzer + production rules: conversion of token sequences into dependency trees

Prepositional construction analysis: table of **prepositional relations**

SemSin analysis of a sentence containing *через*
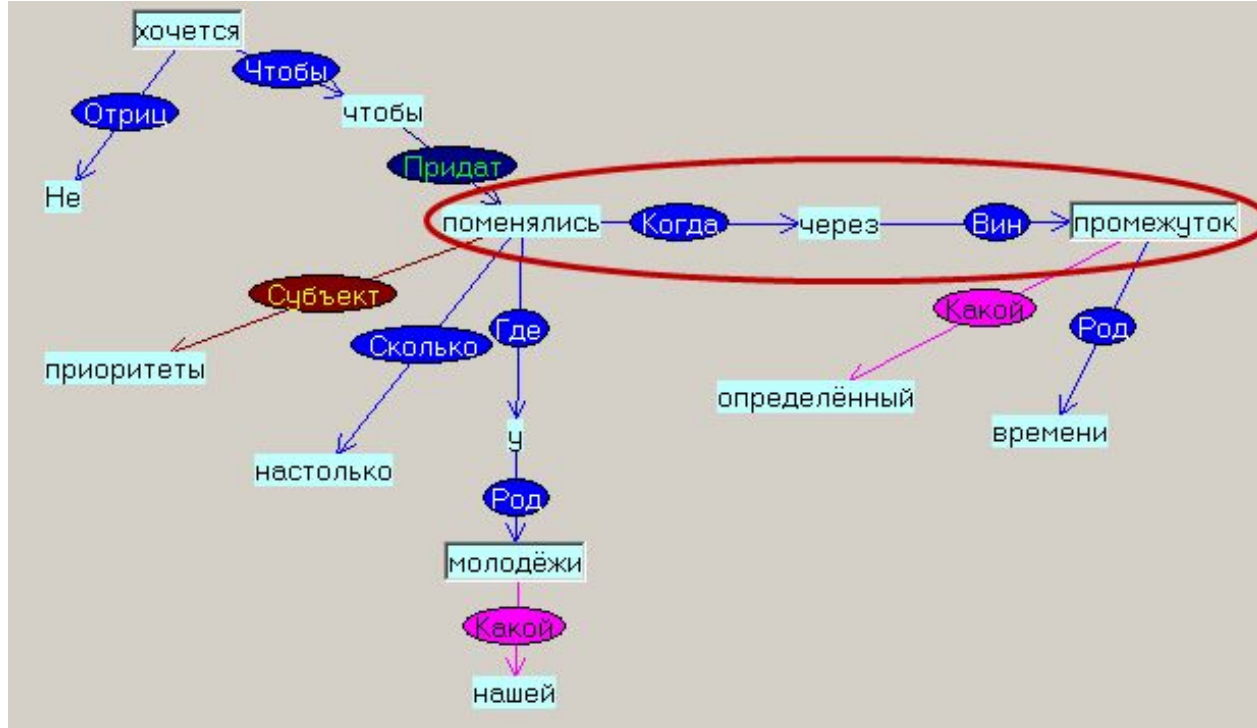
Circled:

поменялись через промежуток 'changed in a span [of time]'

Через relation types: Zolotova vs. SemSin

Zolotova	SemSin
Transitive: trajectory or obstacle *перевозки нефти через Атлантику 'oil transits across the Atlantic'*	**Через_Сквозь 'Across_Through'**: through + midway stop + area coverage *тянуться через всё село 'stretch across the whole village'*
	Через_Транзитив 'Through_Transitive': obstacle + overcoming *работать через боль 'work through the pain'*
Dimensive: distance *живут через дорогу 'live across the street'*	**Через_Мера 'Across_Measure'**: distance + time span + alternation *через один километр 'in one kilometer'*
Temporative: time span *прибудет через минуту 'arrives in a minute'*	
Mediative: way/means of doing sth *выражается через слово 'expresses through words'*	**Через_Медиатив 'Through_Mediative'**: through + means + medium (mediator) *пить через трубочку 'drink through a straw'*

SemSin and через: results

Text type	Governor/governee accuracy	Relation type accuracy
Oral transcripts	87%	87%
Newspaper articles	87%	85%
Scientific texts	92%	73%
Fiction	92%	87%
Legislative texts	65%	78%

Overall: satisfactory for the task

Legislative texts: atypically long and complex syntax structures

To improve: parsing rules, preposition-noun semantic class pair detalization

Conclusion

SemSin: automatic prepositional construction...

    <u>extraction</u>: satisfactory

    <u>meaning classification</u>: to be improved

Future work:

1) expansion of preposition inventory
2) prepositional semantics identification: study and description of preposition-noun compatibility

References

Azarova, I., Zakharov, V., Khokhlova, M., Petkevič, V.: Ontological description of Russian prepositions. In: CEUR Workshop Proceedings, 2552, pp. 245-257 (2020). PRLEAL-2019 (R. Piotrowski's Readings in Language Engineering and Applied Linguistics). Proceedings of the III International Conference on Language Engineering and Applied Linguistics (Saint Petersburg, Russia, November 27, 2019). Pp. 245-257 URL : http://ceur-ws.org/Vol-2552/.

Boyarsky, K., Kanevsky, Ye. Semantic-syntactic parser SemSin [Semantiko-sintaksicheskiy parser SemSin]. In: Scientific and technical bulletin of information technologies, mechanics and optics [Nauchno-tekhnicheskiy vestnik informacionnykh tekhnologiy, mekhaniki i optiki]. Vol.15. No. 5, pp. 869-876 (2015).

Tuzov, V.A. Computational semantics of the Russian language. Saint Petersburg: SpbU publishing (2004).

Acknowledgements

Thank you for your attention!

Any questions?