

Russian NLP Evaluation 2012

Syntax

Svetlana Toldova (toldova@yandex.ru)^{1,2}, **Elena Sokolova** (minegot@rambler.ru)²,

Olga Lyashevskaya (olesar@gmail.com)³,

Maxim Ionov¹, **Anastasia Gareyshina**¹,

Anastasia Bonch-Osmolovskaya^{1,3}

Irina Astaf'eva¹, **Anna Koroleva**¹, **Dmitry Privoznov**¹, **Eugenia Sidorova**¹,

Lyubov' Tupikina¹, **Julia Grishina**¹, **Anna Lityagina**¹, **Natalia Menshikova**¹,

Alexandra Semenovskaya¹

¹Lomonosov Moscow State University, Moscow, Russia

²Russian State University for the Humanities, Moscow, Russia,

³National Research University Higher School of Economics, Moscow, Russia

Background

Evaluation of NLP tools

- ❖ CLEF
- ❖ Morpho Challenge
- ❖ AMALGAM
- ❖ GRACE
- ❖ EVALITA
- ❖ PASSAGE
- ❖ SEMEVAL
- ❖ РОМИП
- ❖ ...

ROOT неакт-компл @Дат|@Род нум-аппоз @Куда|@кДат @Имс @Узо|@От пролетн @между|@в компл-аппозизъясн @КакВ|@Тв @Вин|@Лат|@Род @Дат|@Род @ЛееКак|@вПред 1-компл @Где|@аналитн @оКакой|@и|@зо @КакВ|@подВинном-аппоз @Дат|@Род|@Тв @Для|@КакВ @заТв @ОКакой|@вПредрелят @ДееКак|@поДат аппрокс-колич

Challenges for Russian



What about standard corpora?

- Annotated
- Freely distributed
- Generally accepted as a Gold Standard



What about systems?

- Existing systems for Russian language
- Their quality



Ru-Eval initiative

- ❖ Evaluation of NLP in Russian
- ❖ 2010 — Morphology
- ❖ 2011-2012 — Syntax
- ❖ Objectives:
 - ❖ To track the state-of-art systems and technologies
 - ❖ To provide a benchmark
 - ❖ To disseminate the information to the community



Participants

11 participants $\xrightarrow{\text{got data}}$ 7 participants $\xrightarrow{\text{submitted results}}$ 6 participants

- ❖ SynAutom
- ❖ DictaScope Syntax
- ❖ SemSin
- ❖ ЭТАП–3
- ❖ SemanticAnalyzer Group
- ❖ AotSoft
- ❖ ABBYY Syntactic and Semantic Parser (ASSP)
- ❖ Link Grammar parser
- ❖ Russian Malt



Principles & design

- ❖ 2 tracks:
 - ❖ General — style mixture
 - ❖ News
- ❖ Tasks:
 - ❖ Dependency parsing
 - ❖ Extracting:
 - ❖ Token head (its ID)
 - ❖ Syntactic role (*optional*)
 - ❖ Morphological tag (*optional*)

⇒ Evaluation metrics is unlabelled attachment score

ROOT неакт-компл @Дат|@Род нум-аппоз @Куда|@кДат @Вин|@Лат|@Род @ЛееКак|@вПред 1-компл @Где|@КакВ|@подВинном-аппоз @Дат|@Род|@Тв @Для|@КакВ @заТв @0Какой|@вПред релят @ДееКак|@поДат аппрокс-колич

Principles & design

- ❖ Gold Standard:
 - ❖ > 700 sentences
 - ❖ Manually tagged
 - ❖ Tagging instructions by E.G. Sokolova
 - ❖ Main purpose of the instruction
 - ❖ To provide robust procedure
 - ❖ **Not** to build theoretically right trees

неакт-компл @Дат|@Род
нум-аппоз @Куда|@кДат
суб-копр @Вин|@Дат|@Род
компл-аппозизъясн @КакКто|@Тв
примыкат @КакВ|@Тв
аппрокс-колич @Дат|@Род|@Тв @Для|@КакВ @заТв @0Какой|@вПредрелят

Preparation

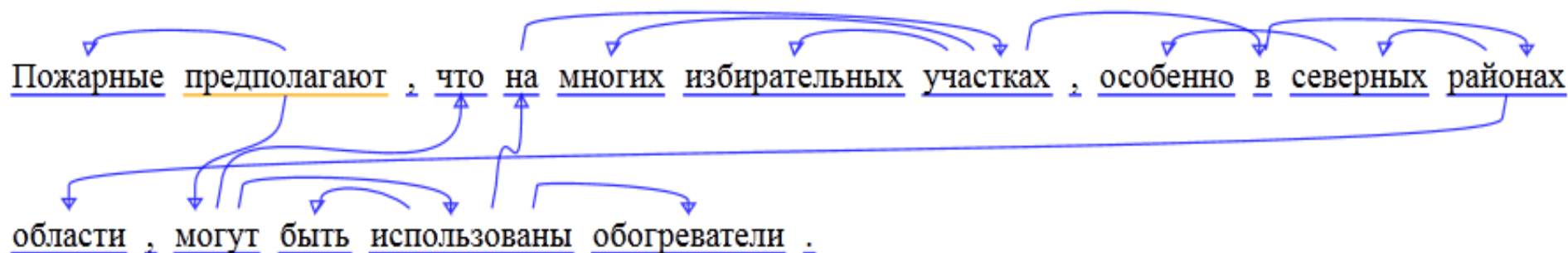
- ❖ Preliminary stage:
 - ❖ 100 random sentences from corpus
 - ❖ Parsed by participants
 - ❖ Manually reviewed and compared
 - ❖ Results:
 - ❖ Variation in tags
 - ❖ Variation in dependency directions
- ⇒ No tag assignment evaluation is possible for now



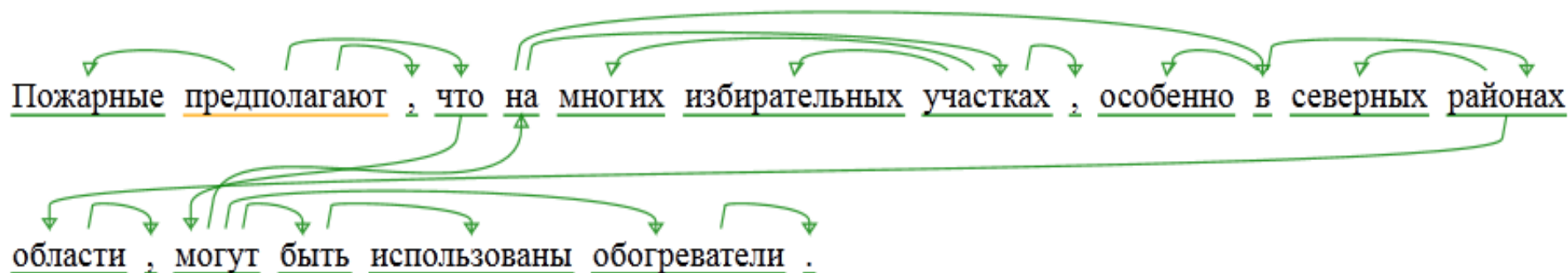
неакт-компл
@Дат|@Род
нум-аппоз
@Куда|@кДат
@Имс|@Изо|@От
пролен|@между|@в
компл-аппозизъясн
@КакКто|@Тв
суб-копр
@КакВ|@Сравн|@Вин|@КакЧто
@Вин|@Дат|@Род|@ДееКак|@вПред
1-компл
@КакВ|@Тв
примыкат
@Где|@КакВ|@подВинном-аппоз
@Дат|@Род|@Тв
@Для|@КакВ
@заТв
@0Какой|@вПред
релят
@ДееКак|@поДат
аппрокс-колич

Diversity of solutions

Trieste



Manchester



ROOT неакт-компл @Дат|@Род нум-аппоз @Куда|@кДат @Вин|@Лат|@Род @ЛееКак|@вПред 1-компл @Где @КакВ|@подВинном-аппоз @Дат|@Род|@Тв @Для|@КакВ @заТв @0Какой|@вПред релят @ДееКак|@поДат аппрокс-колич

Diversity of solutions

Пожарные предполагают, что на многих избирательных участках, особенно в северных районах области, могут быть использованы обогреватели.

- differences from Gold Standard:
- 6 (times) – preposition phrase adjoining: *на многих избирательных участках* and *в северных районах области*.
- 5 – head for *особенно*.
- 4 – head for *обогреватели*, 4 – head for *могут*. (depends on (a) whether conjunction ‘что’ (“that”) is a head of subordinate sentence or not and (b) how to parse a string like *могут быть использованы*).

So far: PP adjoining, emphasizing adverbs/particles, passive construction, modal verbs problems

ROOT неакт-компл @Дат|@Род нум-аппоз @Куда|@кДат @Вин|@Лат|@Род @ЛееКак|@вПред 1-компл @Дат|@Род|@Тв @Для|@КакВ @КакВ|@подВинном-аппоз @Дат|@Род|@Тв @Для|@КакВ @заТв @0Какой|@вПред релят @ДееКак|@поДат аппрокс-колич

Preposition phrase coordination

~~sentence 63838:~~

Наш регион приближается к эпидемии по времени, но не по числу заболевших.

Barcelona –

по времени, но не по числу заболевших.

Napoli –

по времени, но не по числу заболевших.

Toulon –

по времени, но не по числу заболевших.

Nice –

по времени но не по числу заболевших

Brana —

по времени , но не по числу заболевших .

Trieste –

по времени, но не по числу заболевших.

Marceille –

по времени, но не по числу заболевших.

Manchester –

по времени, но не по числу заболевших

неакт-компл @Дат|@Род
@Имс нум-аппоз @Куда|@кДат
@Узо|@От
проект @между|в
компл-аппозизъясн @Вин @Вог
@КакКто|@Тв @Вин @Вог
примыкат @КакВ|@Тв @Где|
@КакВ|@подВинном-аппоз
@Дат|@Род|@Тв @Для|@КакВ
@заТв @ОКакой|@вПредрелят
@ДееКак|@подат
аппрокс-колич

Conversion “effects”

ответственности в соответствии с федеральным законом .

На имя Главного федерального инспектора по Кировской области Валентина Пугача поступило письмо от службы судебных приставов в ответ на открытое письмо

ROOT неакт-компл @Дат|@Род нум-аппоз @Куда|@кДат @Вин|@Лат|@Род @ЛееКак|@вПред 1-компл @ДееКак|@поДат аппрокс-колич
пролетт @между|в @КакВ|@Тв @КакВ|@подВинном-аппоз @Дат|@Род|@Тв @Для|@КакВ @заТв @0Какой|@вПред релят
компл-аппоз изъясн @КакКто|@Тв @Вин @Вин примыкат @Где @Дат|@Род|@Тв @Для|@КакВ @заТв @0Какой|@вПред релят
суб-копр @КакВ|@Сравн @Вин|@КакЧто @Дат|@Род|@Тв @Для|@КакВ @заТв @0Какой|@вПред релят

Flexibility of evaluation

1. Computation of feasible solutions
2. Recompile of associated relations
3. Separate computation of feasible relations within each system

13в	circ	7	14	1	3
14соответствии	lexmod	13	17	1	3
15с	lexmod	14	14	0	0
16федеральным	mod	17	17	0	0
17законом	рcomp	15	12	1	3

неакт-компл
@Дат|@Род
нум-аппоз
@Куда|@кДат

пролен @между | в
компл-аппозизъясн @Вин @Вог
@КакКто|@Тв @Вин @Вог

примыкат
@Где|
@КакВ|@подВинном-аппоз
@Дат|@Род|@Тв @Для|@КакВ
@заТв @ОКакой|@вПредрелят

суб-копр
@КакВ|@Сравн@Вин|@КакЧто
@Вин|@Лат|@Род@ДееКак|@вПред 1-компл @ДееКак|@поДат

аппрокс-колич

(unlabelled attachment score)

неакт-компл @Дат|@Род
нум-аппоз @Куда|@кДат
суб-копр @Вин|@Дат|@Род
компл-аппозизъясн @КакКто|@Тв
примыкат @КакВ|@Тв
аппрокс-колич @Дат|@Род|@Тв @Для|@КакВ @заТв @0Какой|@вПредрелят

Results. News

	P	
Trieste	0,957	Compreno
Marceille	0,900	ЭТАП–3
Barcelona	0,879	SyntAutom
Brega	0,809	Dictum
Nice	0,807	Semantic analyzer group
Toulon	0,780	SemSyn
Napoli	0,732	AotSoft

ROOT неакт-компл @Дат|@Род нум-аппоз @Куда|@кДат @Вин|@Лат|@Род @ЛееКак|@вПред 1-компл @Где| @КакВ|@подВинном-аппоз @Дат|@Род|@Тв @Для|@КакВ @заТв @0Какой|@вПредрелят @ДееКак|@поДат аппрокс-колич

Conclusions

Three basic approaches to parsing :

- systems manually enriched with expert linguistic knowledge
- automata-based systems
- machine-learning

Positive results:

1. Advantages of the systems manually enriched with expert linguistic knowledge
 - ✓ drawbacks – resource- and time-consuming procedure
2. The existence of Treebank that enables reliable machine learning for parsing
3. Low-time-consuming systems are also reliable



Outcome. A treebank with parallel annotation

1 mln. tokens annotated by three different systems - participants of RU-EVAL 2012

new perspectives for applied and theoretical studies

<http://testsynt.soiza.com>

неакт-компл
@Дат|@Род
нум-аппоз
@Куда|@кДат
@Имс|@От
@Вин|@Дат|@Род|@ЛееКак|@вПред
1-компл
@Дат|@Род|@ЛееКак|@вПред
@КакВ|@Тв
@КакВ|@подВинном-аппоз
@Дат|@Род|@Тв
@Для|@КакВ
@заТв
@0Какой|@вПред
релят
аппрокс-колич

RU-EVAL 2012: positive results

optimizing organization routines

forum for participants

a new option to compare results with gold standard for all participants after the competition -> feedback

conference papers of participants

educational outcome:

students involved in evaluation

a course on syntactic parsers in MSU by S.Toldova with participants as invited guests



неакт-компл @Дат|@Род
нум-аппоз @Куда|@кДат
суб-копр @КакВ|@Сравн@Вин|@КакЧто
компл-аппозизъясн @КакКто|@Тв @Вин @Вог
примыкат @КакВ|@Тв @Где|@КакВ|@подВинном-аппоз
@Дат|@Род|@Тв @Для|@КакВ
@заТв @0Какой|@вПредрелят
аппрокс-колич @ДееКак|@поДат