Dialogue-2014
Bekasovo, 04-08.06.2014

Evaluation of frame-semantic role labeling in a case-marking language

Olga Lyashevskaya[1,2] (olesar@gmail.com)
Egor Kashkin[2] (egorkashkin@rambler.ru)
[1]Higher School of Economics, Moscow
[2]Vinogradov Institute of Russian Language RAS, Moscow

SEMANTIC ROLE LABELING (SRL): BACKGROUND

История

не

знает

привилегий

'History does not take into account privileges'

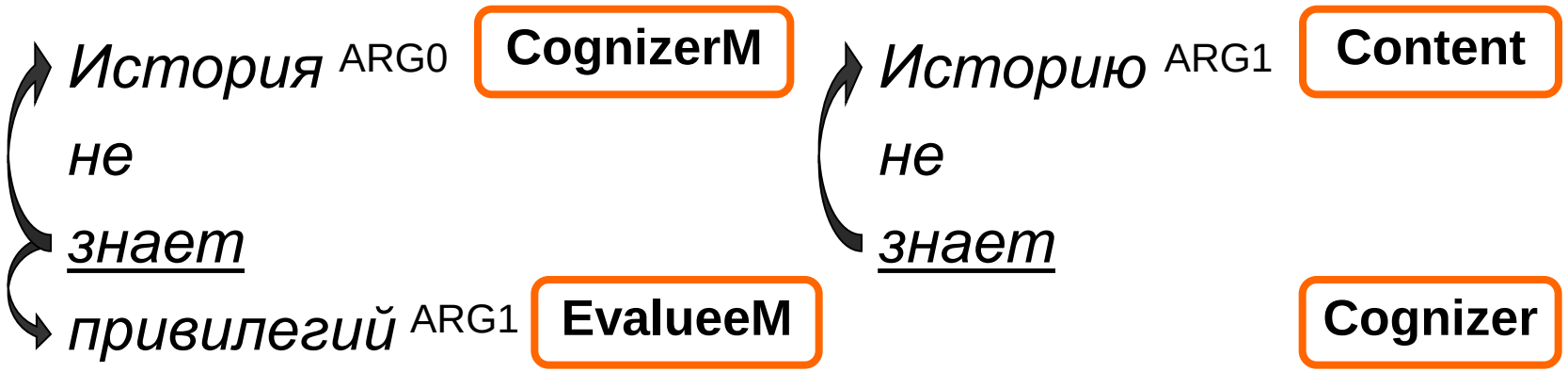
Историю

не

знает

'[She/He] does not know history'

SEMANTIC ROLE LABELING (SRL): BACKGROUND



'History does not take into account privileges'

'[She/He] does not know history'

CoNLL 2008 shared task (Surdeanu et al. 2008): PropBank-like SRL

SemEval 2007 (Baker et al. 2007): FrameNet-like SRL

SRL FOR RUSSIAN

- Anisimovich et al. 2012, Petrova 2013
  Azarova 2008
  Shelmanov & Smirnov 2014
  Ermakov & Pleshko 2009
  Kuznetsov 2012
  ... yet project in progress

- Many possible classifications of roles and frames

- Russian FrameBank as a benchmark for SRL evaluation + class mapping

OUTLINE OF EVALUATION EXPERIMENT

- a prototype of SRL module (rule-based)
- training and test data sets from FrameBank
- how it works (cues and challenges of Russian PP SRL)
- induced roles VS gold standard roles:
  - metrics
  - the goodness of fit for non-matching pairs

Data source: Russian FrameBank

- Berkeley FrameNet : extralinguistic situations (frames) → a set of participants → lexemes.
- Russian FrameBank (www.framebank.ru): lexemes (about 2200, primarily verbs) → constructions ([Apresjan, Pall 1982] & added by annotators) → examples from RNC (about 100 for each lexeme, manually tagged)
- Theoretical framework: FrameNet and Construction Grammar + Moscow Semantic School

- *резать* 'to cut, to carve':
- ✓ *Продавщица* ***режет*** *сыр* 'The shop assistant is cutting cheese'*; Она* ***режет*** *хлеб на тонкие куски* 'She is slicing bread (lit.: cutting bread into thin slices)'*…*
- ✓ *Старик* ***резал*** *четки из кипариса* 'The old man carved rosaries out of cypress'*; Он* ***резал*** *деревянные ложки ножом* 'He carved wooden spoons with a knife'*…*
- ✓ *У него в желудке* ***резало*** 'He had griping pains in his stomach (lit.: It was cutting in his stomach)'

FRAMEBANK: CONSTRUCTION PATTERN

ID230. Cx name: *Pjatno vystupilo na rubaške* ['a stain appeared on the short']. Cx Pattern: Snom V na + Sloc.								
Cx Item ID	Pl	Letter	Head	Phrase	Explication	Syntactic Rank	Lexico-semantic constraints	Status [obligatory / optional]
2077	1	**X**	**Snom** [Nominative case]	**NPnom**	**substance**	**Subject**	**natural object**	**Oblig.**
2078	2	–	***vystupit'*** ['to appear; lit. to step forward']	–	**to appear**	**Predicate**	–	**Oblig.**
2079	3	**Y**	***na*** **+ Sloc** [preposition ***na*** 'on' + Locative case]	***na*** **+ NPloc**	**location: surface**	**Peripheral**	**space and place**	**Oblig.**

Lexical Index of target words

Index of Morphosyntactic Items

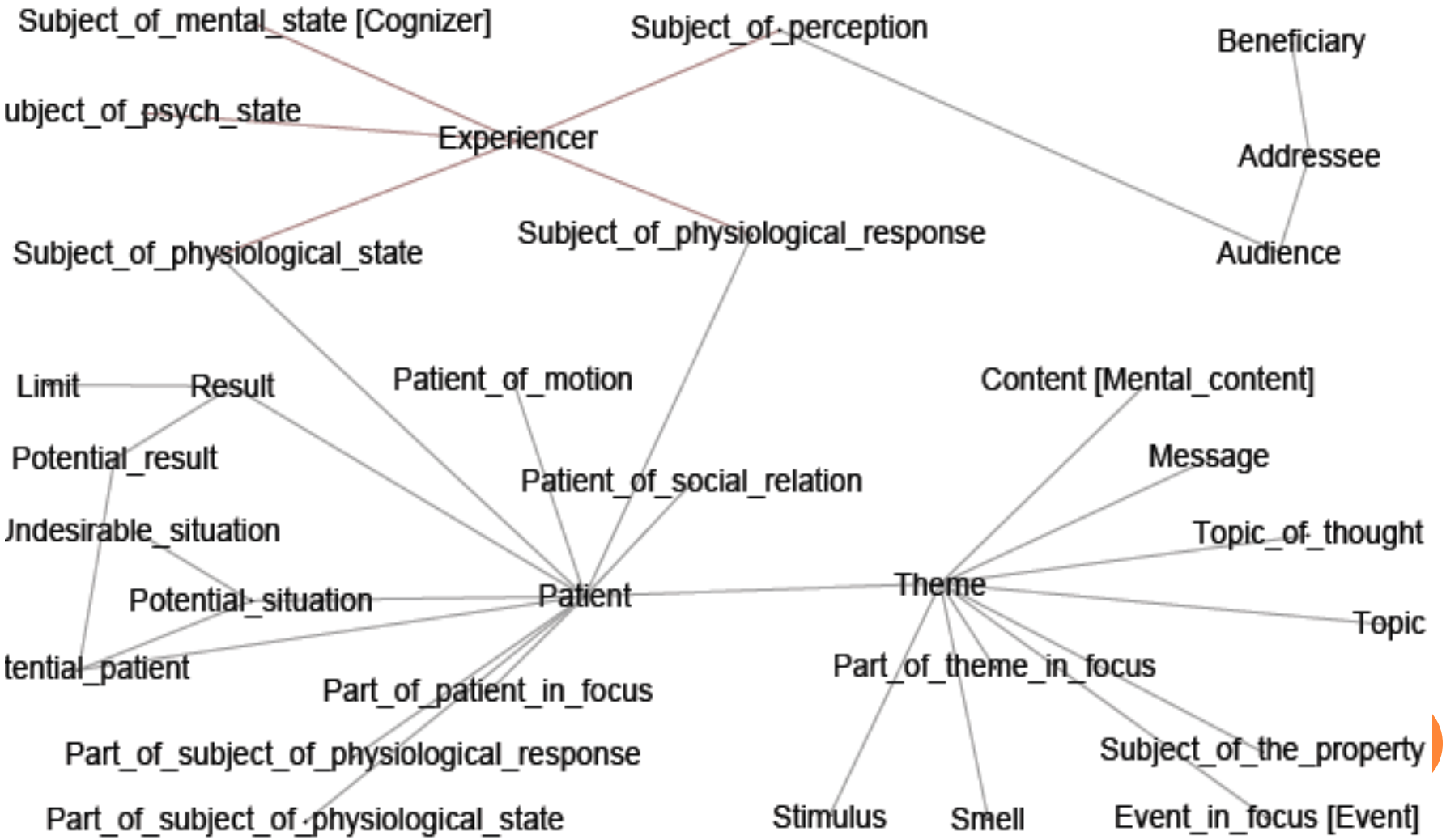
FRAMEBANK: TAGGING

ID5914. Cx name: *Soberi požinat'* ['pick up something for supper']. Cx Pattern: Snom V Vinf.								
ID	Pl	Letter	Head	Phrase	Explication	Syntactic Rank	Lexico-semantic constraints	Status / Realization
18589	1	**X**	**Snom**	**NPnom**	**Agent**	**Subject**	**human**	**Oblig.**
20089	-	X	–	–	–	No	–	Omitted. Licensed by Imperative Cx
18590	2	–	***sobrat'***	–	**to collect**	**Predicate**	–	**Oblig.**
20090	1	–	*soberi.Vimper*	–	to collect	Predicate	–	Standard
18591	3	**Y**	**Vinf**	**VPinf**	**what is collected**	**Peripheral**	**eat**	**Oblig.**
20091	3	Y	*poest'*.Vinf	*poest' v dorogu*.VPinf	what is collected	Peripheral	eat	Standard
20089	2	Z	*nam*.SPROdat	*nam*.NPdat	Beneficiary	Peripheral	human	Added. Licensed by Ditransitive Dative Cx
20089	4	W	*v dorogu*	*v dorogu*.NPdat	Goal	Adjunct	abstract	Added
+								

FrameBank: semantic roles

- Different inventories of semantic roles (cf. Ch. Fillmore, Ju.D. Apresjan, E.V. Paducheva, etc.)
- FrameBank:
  - ✓ hierarchy of semantic roles → flexible search options
  - ✓ correlation between the roles and the semantic classes of verbs.
  - ✓ semantic roles graph: 96 items → 6 domains → further smaller units.

FrameBank: semantic roles graph



OUTLINE OF EVALUATION EXPERIMENT

- a prototype of SRL module (rule-based)
- training and test data sets from FrameBank
- how it works (cues and challenges of Russian PP SRL)
- induced roles VS gold standard roles:
  - metrics
  - the goodness of fit for non-matching pairs

SRL PROTOTYPE: PREPOSITIONAL PPs

- Four PPs:

  za + NPins
  za + NPacc
  ot + NPgen
  po + NPdat

- Very frequent (e.g. ca. 900.000 hits of "*ot* + NPgen" in the RNC).
- Highly polysemous (e.g. 14 possible roles of "*za* + NPacc")

TRAINING AND TEST DATA

- Training data set: constructions from [Apresjan, Pall 1982]
- Test data set: constructions added by annotators

- NB type units (from dictionary), not tokens (hits from corpus)

PP	Training set: ‘old’ data		Test set: ‘new’ data	
	# constructions	# examples	# constructions	# examples
za + NPins	95	80	19	22
za + NPacc	228	223	37	51
ot + NPgen	266	435	70	113
po + NPdat	311	245	65	78
Total	900	983	191	264

SRL PROTOTYPE: PREPOSITIONAL PPs

- 62 heuristics

cf. [Toldova et al. 2008] on the semantic filters in the RNC

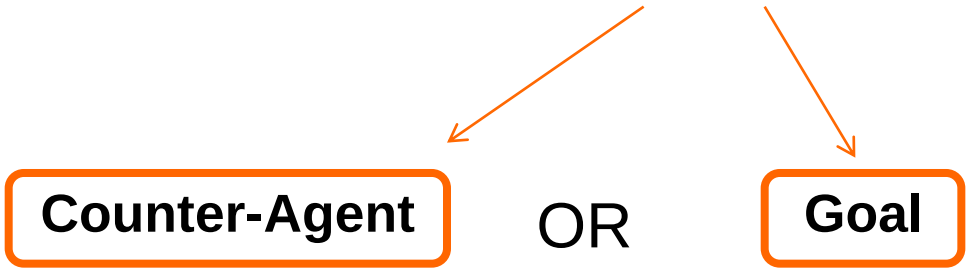
- ✓ morphosyntactic pattern
- ✓ lexical class of the PP argument
- ✓ lexical class of the predicate
- ✓ lexical classes of other arguments

- Examples: *za* + NPacc
- ✓ NPnom $V_{\text{CHANGE POSS.}}$ NPacc *za* + NPacc → **Price**
- ✓ *za* + $\text{NPacc}_{\text{TIME PERIOD}}$ → **Period**

SRL PROTOTYPE: PREPOSITIONAL PPs

- Two possible outcomes for some rules:

$$NPnom_{ANIM}\ V_{MOTION}\ za + NPins_{ANIM}$$



- Counter-Agent: *Milicioner pobezhal* ***za prestupnikom*** ‘A policeman ran **after an offender**’
- Goal: *Mal’chik pobezhal v bol’nicu* ***za vrachom*** ‘A boy ran to hospital **to call the doctor**’

SRL CUES

- Semantic class of a verb
- ✓ *Beshus'*$_{\text{EMOTION}}$ ***za doch*** *moju* 'I am in a rage **because of my daughter**' → Reason for an Emotion, cf. *bojat'sja* 'to be afraid', *bespokoit'sja* 'to worry about sth.'
- Semantic class of a participant
- ✓ ***Po radio***$_{\text{COMMUNIC. FACIL.}}$ *igrala muzyka* 'There was music broadcast (lit.: played) **by radio**' → Manner. cf. *vystupat'* ***po televizoru*** 'to speak **on TV**', *poslat' dokumenty* ***po pochte*** 'to send documents **by post**'

Challenges for SRL

- Pragmatic factors, cf. Counter-Agent vs. Goal of motion.
- No clear semantic constraints on a verb or on its arguments.
  - ✓ "*po* + NPdat$_{???}$" for Reason: *Rasskaz byl prochitan po ego pros'be* 'The story was read at his request' → ~~Information Source~~, Reason, cf. *zhenit'sja* ***po ljubvi*** 'to make a love-match (lit.: to get married **due to love**)', *uvolit'* ***po sokrascheniju shtatov*** 'to discharge sb. **on grounds of staff reduction**', *sidet' zdes'* ***po drugomu delu*** 'to stay here **on some other business**'
- Metonymies of concrete nouns
  - ✓ *Ego nevozmozhno otorvat'* ***ot knigi*** 'It is impossible to divert his attention **from the book** (lit.: to tear him=it from the book)' → ~~Patient & Location~~ Content of Action

EVALUATION METRICS

1. Precision, Recall, F1

   P = #Matches / #E_Answers

   R = #Matches / #G_Answers

2. Cluster purity and collocation, F1 (Lang & Lapata 2011)

$$\mathrm{Pu} = \frac{1}{n}\sum_{i=1}^{n_C} \max_{j=1,\ldots,n_G} |C_i \cap G_j| \qquad \mathrm{Co} = \frac{1}{n}\sum_{j=1}^{n_G} \max_{i=1,\ldots,n_C} |C_i \cap G_j|$$

3. Repulsion: how likely are the roles to distinguish frames

$$\text{repulsion} = \frac{\#\text{Verbs(RoleE\_RoleG OR RoleE!RoleG)}}{\sqrt{\#\text{Verbs(RoleE)} \times \#\text{Verbs(RoleG)}}}$$

4. Distance between roles in semantic role graph

RESULTS: PRECISION

PP	Total amount of new patterns	'Strong' matching (the role is identified correctly and unambiguously	'Weak' matching (one of the answers is correct)	P_{strong}	$P_{strong+weak}$
za + NPins	19	9	7	0.47	0.84
za + NPacc	37	22	11	0.59	0.89
ot + NPgen	70	41	24	0.59	0.93
po + NPdat	65	32	25	0.49	0.88
Total	191	104	67	0.54	0.90

THE GOODNESS OF FIT AND REPULSION

Matching Evaluation (human)	Role E	Role G	#Verbs (RoleE)	#Verbs (RoleG)	#Verbs (RoleE! RoleG)	#Ve (Ro Role
Good	Source	Reason	12	266	3	
[Source↑External_cause↑Age						
Lovit' kaif [*ot knig*] ‘To be in high from books'.						
Good	Path	Patient	105	712	46	
[Path↑Locati						
On bredit i mechetsja golovoj [*po perekladine*] ‘He raves, tossing his head over						
Good	Property	Reason	175	266	31	
Ego zabrali [*po nacional'nomu piznaku*] ‘He was arrested on ethnic grounds'.						
Average	Term	Time_point	52	42	6	
Vstrecha prodlilas' [*za polnoch*] ‘The meeting lasted past midnight'.						
		Target_				

THE GOODNESS OF FIT AND REPULSION

Role G	#Verbs (RoleE)	#Verbs (RoleG)	#Verbs (RoleE! RoleG)	#Verbs (RoleE+ RoleG)	Repulsion	Same domain
Reason	12	266	3	0	0.05	NO

[Source↑External_cause↑Agent]↑Root↓[Setting↓Reason]

high from books'.

Patient	105	712	46	3	0.18	NO

[Path↑Location↑Setting]↑Root↓[Patient]

voj [*po perekladine*] 'He raves, tossing his head over the crossbar'.

Reason	175	266	31	5	0.17	(YES)

[Property↑Setting↓Reason]

omu piznaku] 'He was arrested on ethnic grounds'.

Time_point	52	42	6	2	0.17	YES

[Term↓Time_point]

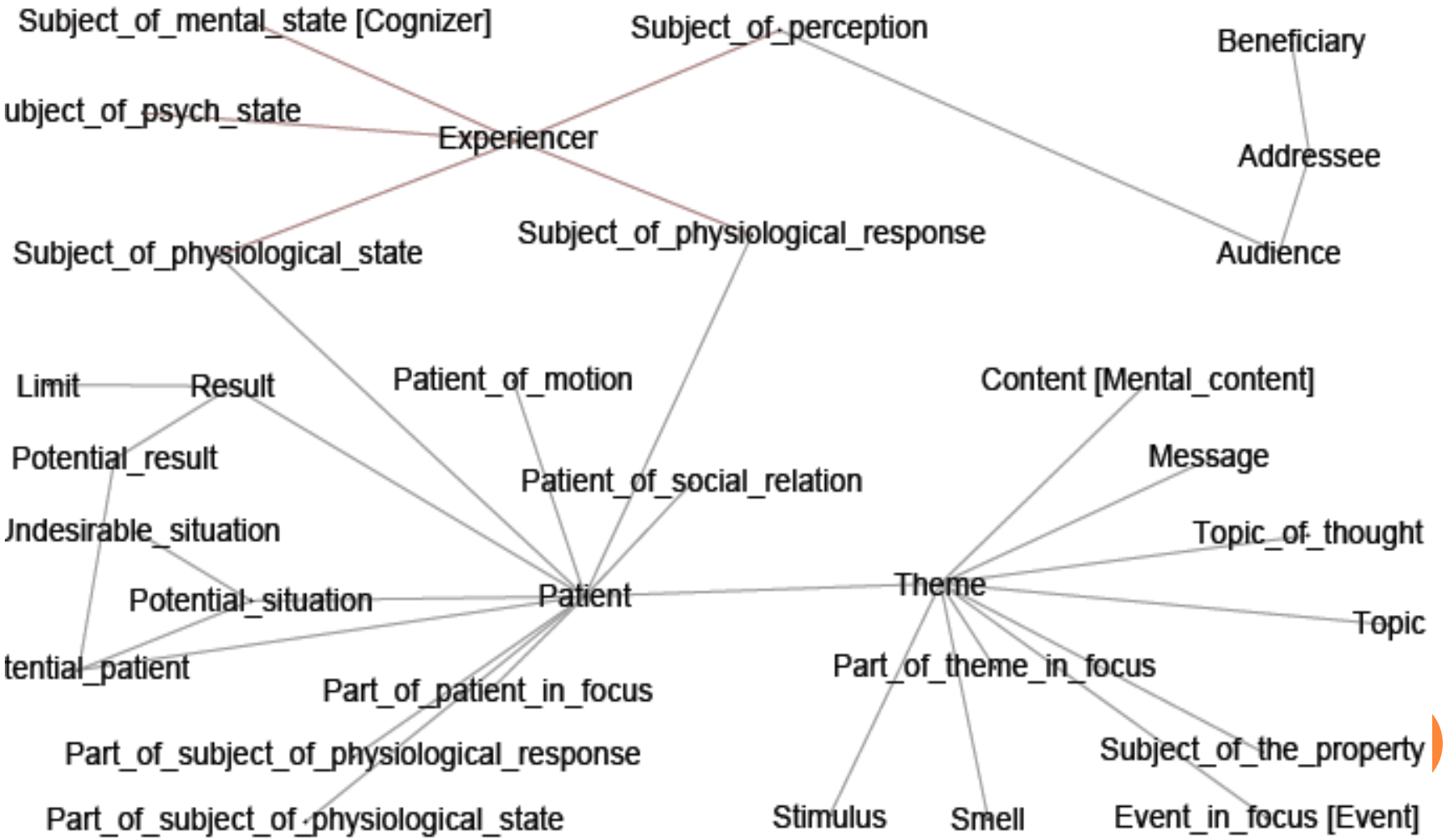
och] 'The meeting lasted past midnight'.

Target_						

CONCLUSION

- Repulsion correlate quite well with the split between Average and Bad matches (repulsion threshold .20) √
- ... but split between Good and Average matches:: FAILED χ
- Same-domain VS Good_Average_Bad matches:
  - all Bad matches do not share the same domain but this is not sufficient χ
- Graph-based distances: FAILED χ
- *Future development*
- Other SRL modules
- Token frequencies
- More (other than IS-A) types of edges in the SR graph

СПАСИБО !


Subject_of_mental_state [Cognizer]
Subject_of_perception
Beneficiary
ubject_of_psych_state
Experiencer
Addressee
Subject_of_physiological_state
Subject_of_physiological_response
Audience
Limit
Result
Patient_of_motion
Content [Mental_content]
Potential_result
Message
Patient_of_social_relation
Jndesirable_situation
Topic_of_thought
Potential_situation
Patient
Theme
Topic
tential_patient
Part_of_theme_in_focus
Part_of_patient_in_focus
Part_of_subject_of_physiological_response
Subject_of_the_property
Part_of_subject_of_physiological_state
Stimulus
Smell
Event_in_focus [Event]