**METAPHOR SHIFTS IN CONSTRUCTIONS:**
THE RUSSIAN METAPHOR CORPUS

Yulia Badryzlova, Olga Lyashevskaya

AAAI Spring Symposium Series
Palo Alto, California

MARCH 28, 2017

1. This presentation is available for download at: http://web-corpora.net/~badryzlova/AAAI_2017/Presentation

National Research University Higher School of Economics
School of Linguistics
Moscow, Russia
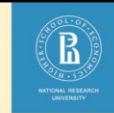
YULIA BADRYZLOVA
PhD Student

OLGA LYASHEVSKAYA
Professor

2.

CREMe: THE EXPERIMENTAL RUSSIAN METAPHOR CORPUS
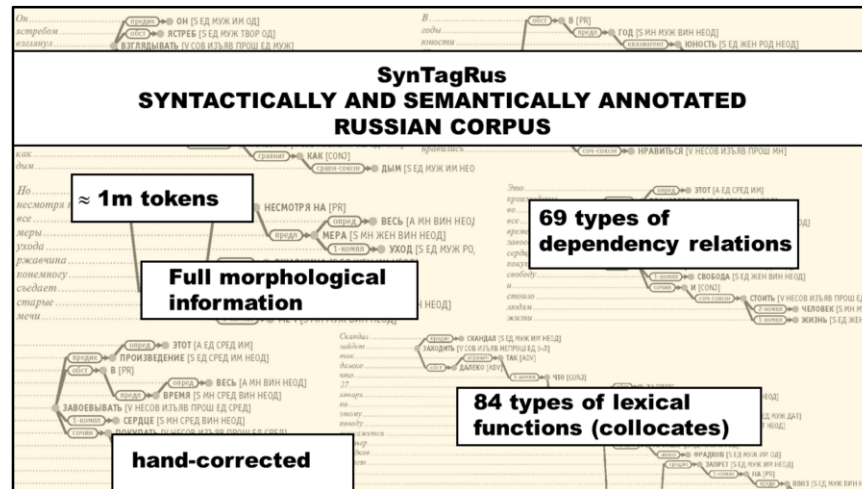
3. The title of our project is "CREMe", which is a frivolous acronym of "The Experimental Russian Metaphor Corpus".
This is an ongoing project, so we will not report any quantitative results; rather, we will present the scheme the we offer for the annotation of metaphor, as metaphor annotation is by far not a trivial task.

SynTagRus
SYNTACTICALLY AND SEMANTICALLY ANNOTATED RUSSIAN CORPUS

≈ 1m tokens

69 types of dependency relations

Full morphological information

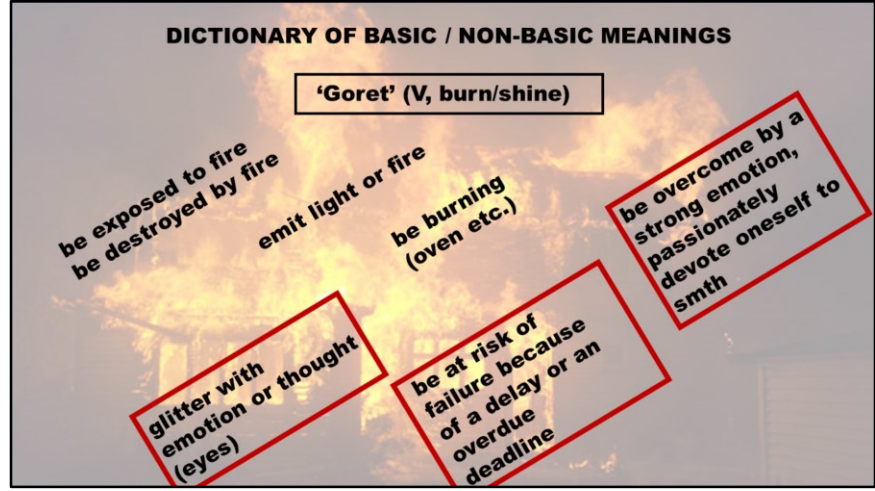84 types of lexical functions (collocates)

hand-corrected

4. What we do is: we annotate metaphor on a corpus that contains rich and high-quality morphological, syntactic and semantic markup, which allows us to observe the behaviour of constructions in metaphoric occurrences. SynTagRus comprises approximately 1m tokes with full morphological information, including (non-)animacy of nominal forms (which is their inherent morphological feature in Russian).

The syntax in SynTagRus is annotated with 69 types of dependency relations: predicative, agentive, quasi-agentive, first complement, modifier, and prepositional, and others.

Besides, SynTagRus contains annotation of 84 types of collocates (or Lexical Functions). For example, the lexical function OPER 1 in 'osuschestvlyat' (v, exercise) + 'kontrol' (n, control) denotes a semantically empty verb such that the first actant of a certain situation functions as the subject of this verb and the name of the situation itself is the verb's first object.

SynTagRus is recognized as a golden standard of Russian syntactic annotation because its initial automatic parsing was manually checked and corrected by linguists.

See [12, 13]

5. Our aim in annotating metaphor is to capture a variety of types of metaphor and the factors that contribute to metaphoricity, and their dynamic interactions in dicsourse.
We annotate the three major types of metaphor: Indirect Metaphor, Personification, and Direct Metpaphor.
The additional factors of metaphoricity annotated in the corpus are: the shift of the semantic class of the argument between Concrete and Abstract, metonymic shifts, newly attested meanings, morphologically motivated forms and semantic shifts.
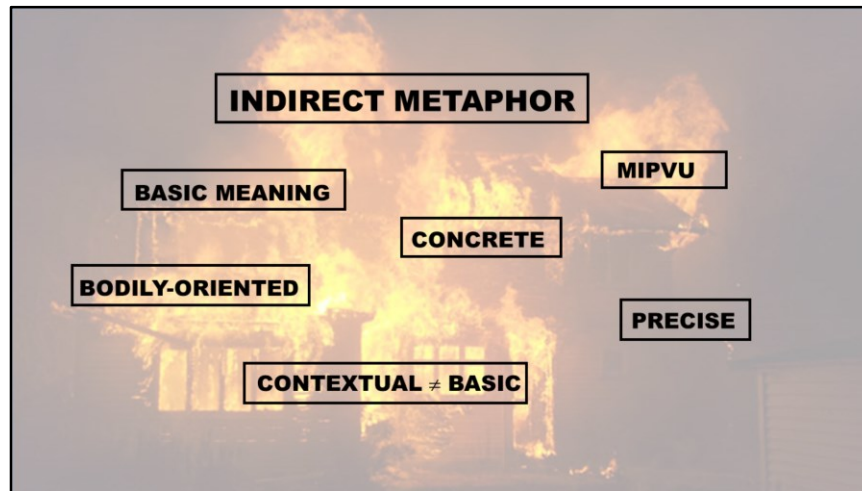
6. The first type of metaphor we annotate is Indirect Metaphor (this is a term borrowed from MIPVU, the Vrije University Metaphor Identification Procedure, proposed by G. Steen et al [5].)

Basically, Indirect Metaphor is another term for conventional lexicalized metaphor.

For each word in a sentence, we establish its Basic Meaning(s) (BM) and Contextual Meaning.

The BM is the most concrete, bodily-oriented and precise meaning of the word. The remaining meanings that are less concrete, bodily-related and precise are Non-Basic Meanings (NBM).

Indirect Metaphor is attested when the contextual meaning of a word in a given context is not BM.

DICTIONARY OF BASIC / NON-BASIC MEANINGS

'Goret' (V, burn/shine)

be exposed to fire
be destroyed by fire

emit light or fire

be burning (oven etc.)

be overcome by a strong emotion, passionately devote oneself to smth

glitter with emotion or thought (eyes)

be at risk of failure because of a delay or an overdue deadline

7. For example, among the six meanings of the Russian verb 'goret' (to burn/to shine), two meanings will be Non-Basic – the ones that relate to emotions and deadlines. Therefore, whenever 'goret' occurs in a sentence in one of these meaings, it will be annotated as IndirMet.

The meaning 'glitter with emotion or thought' is tagged as WIDLII (an acronym borrowed from MIPVU for borderline cases).

We compile a dictionary of word meanings where each meaning is tagged as either BM, NBM or WIDLII.

Our inventory of word meanings comes from an electronic version of The Dictionary of the Russian Language [15] as the most representative and authoritative dictionary of contemporary Russian.

NO POTENTIAL FOR INDIRECT METAPHOR

'Khokhotat' (V) – Scream with laughter
- BM Laugh loudly
- BM Produce sounds resembling laughter

'Uvazhat' (V) – Respect
- NBM Feel respect or reverence towards smb or smth
- NBM Take smth into account or consideration, treat smth with respect
- NBM Love and value smth for its qualities

8. Lexemes that have only Basic or only Non-Basic meanings cannot produce Indirect Metaphor.

PHRASEMIC UNITS, IDIOMS
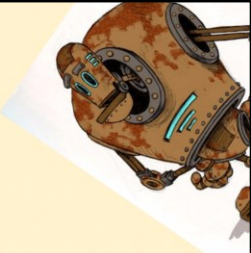
'zayti (V)   daleko (ADV)'
to have walked   far

	predic	advmod		delim	1-compl	
Skandal	zaydet		tak		daleko,	chto...
		phrasemic unit				
A,sg,m, nom inan	V,perf,ind,nonpast, sg,3pers		Adv	Adv		Conj
scandal	go		so	far		that

Met type	LexPers_Target	IndirMet LexPers_Source		IndirMet	
Basic Meaning		'To have walked beyond a certain point'		'A long distance away'	
Contextual Meaning		NBM To have progressed beyond a certain limit'		NBM To have progressed beyond a certain limit'	

I blame it on the GPS-System myself: Hardly anybody gets lost around here anymore...

9. Phrasemic units and idioms are annotated as a subtype of Indirect Metaphor.
In this example, the Non-Basic Meaning of the phrasemic unit 'zayti (V) daleko (ADV) – to have progressed beyond a certain limit – is opposed to the Basic Meaning of the free word combination made up of the same constituents – to have walked far.

LEXICALIZED PERSONIFICATION

'Syedat' (V) – to eat

Rust gradually eats (=corrodes) old swords.

	Rzhavchina	ponemnogu	syedayet	staryye	mechi.
	S,sg,f,Nom, inan	Adv	V,impf,ind, nonpast,3pers	A,pl,Acc inan	S,pl,m, Acc,inan
	rust	gradually	eat	old	sword
Met type	LexPers_Target		LexPers_Sourse		
Basic Meaning			'consume food, destroy smth by eating'		
Contextual Meaning			NBM corrode smth (about acrid and toxic substances, rust etc.)'		

10. The second type of metaphor we annotate is Personification – the shift of the argument class from Animate to Inanimate.
Lexicalized Personification is observed when the central BM of a word requires an animate argument, while a dictionary-attested peripheral BM or an NBM requires an inanimate argument.
The word that triggers the shift of the class is called Source of Personification; the inanimate argument that this word takes is called Target of Personification.

	Eto	proizvedenie	zavoevyvalo	serdtsa,	pokupalo	svobodu.
	A,sg,n, Nom	S,sg,n, Nom inan	V,impf,ind, past,sg,n	S,pl,n.Acc,i nan	V,impf,ind, past,sg,n	S,sg,f,Acc, inan
	this	work of art	conquer	heart	buy	freedom
Met type		MetPers_ LexPers_ Target	MetPers_ LexPers_ Source		MetPers_ Source	AbstrArg
Basic Meaning			'subdue by means of armed force, conquer'		'To purchase'	
Contextual Meaning			NBM 'manage to attract one's friendly disposition		'To purchase'	

This work of art ... conquered hearts, bought freedom.

METAPHORIC PERSONIFICATION

11. Metaphoric Personification is observed when the contextual meaning (either BM or NBM) normally requires an animate argument, but takes an inanimate argument instead, as the verb 'pokupat' – to conquer – in this example. (Beside MetPers, this sentence also contains an instance of Lexicalized Personification (triggered by the verb 'zavoyevyvat' – to concquer and an instance of the shift from Concrete to Abstract, see Slide 15.
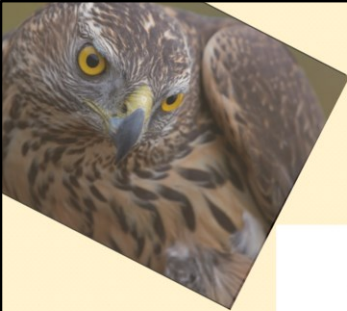
DIRECT METAPHOR

LEXICAL CUE

The old man disappeared like smoke.

	predic	comp	compconj	
	Starik	**ischez**	**kak**	**dym.**
	S,sg,m, Nom,anim	V,perf,indic, past,sg,m	Conj	S,sg,m, Nom,inan
	old man	**disappear**	**like**	**smoke**
Met type	DirMet_Target		LexCue	DirMet_Source
Basic Meaning				'Small hard particles and gases released in the process of burning'
Contextual Meaning				'Small hard particles and gases released in the process of burning'

12. The third type of metaphor we annotate is Direct Metaphor (this term is also borrowed from MIPVU). Contextual meaning (either BM or NBM) refers to a concept from a different domain than the regular domain of the contextual meaning. In terms of Conceptual Metaphor Theory, there is a mapping between a conceptual Source and a Target domains.
Direct Metaphors can be signaled in the sentence by different cues: graphic, lexical or morphological.
In this example, the Direct Metaphor is signaled by the lexical cue 'kak' – like.
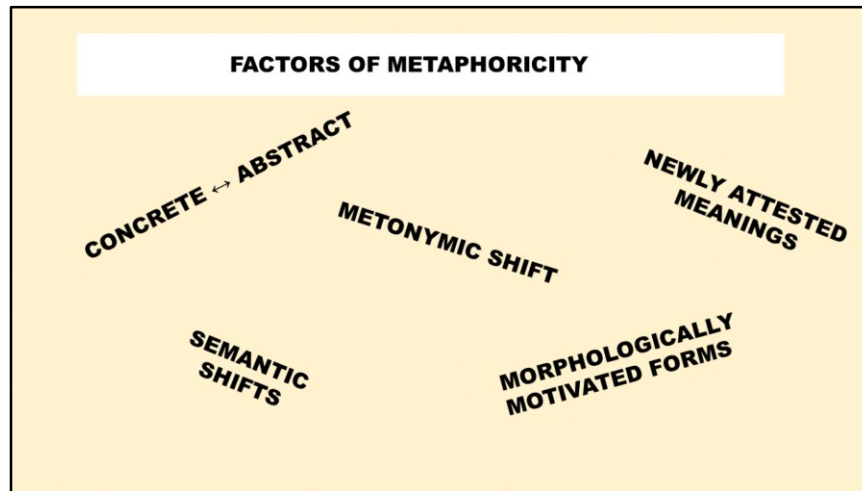
DIRECT METAPHOR

MORPHOLOGICAL CUE

He hawkishly (=fiercely) glanced at the old man.

	On	yastrebom	vzglyanul	na	starika.
	S,sg,m, nom,anim	S,sg,m, Instr anim	V,perf,ind, past,sg,m	Prep	S,sg,m, acc,anim
	He	hawk	glance	at	old man
Met type	DirMet_Target	DirMet_Source MorphCue	DirMet_Target		
Basic Meaning		'Bird of prey'			
Contextual Meaning		'Bird of prey'			

13. Morphological cues are the morphological features that express affinity.
For example, the instrumental construction licensed by the Instrumental Case of a noun is a regular means of expressing resemblance in Russian, as in this sentence.

FACTORS OF METAPHORICITY

CONCRETE ↔ ABSTRACT

METONYMIC SHIFT

NEWLY ATTESTED MEANINGS

SEMANTIC SHIFTS

MORPHOLOGICALLY MOTIVATED FORMS

14. Beside Indirect and Direct Metaphor and Personification, we annotate the additional factors of metaphoricity: shift of semantic class of the argument between Concrete and Abstract, metonymic shifts, newly attested meanings, morphologically motivated forms and semantic shifts.

SHIFT OF SEMANTIC CLASS: CONCRETE ↔ ABSTRACT

This work of art ... conquered hearts, bought freedom.

		attr.	predic	1-compl	coord	1-compl	
	Eto	proizvedenie	zavoevyvalo	serdtsa,	pokupalo	svobodu.	
	A,sg,n, Nom	S,sg,n, Nom inan	V,impf,ind, past,sg,n	S,pl,n.Acc,inan	V,impf,ind, past,sg,n	S,sg,f,Acc inan	
	this	work of art	conquer	heart	buy	freedom	
Met type		MetPers_ LexPers_ Target	MetPers_ LexPers_ Source		MetPers_ Source	AbstrArg	
Basic Meaning			'subdue by means of armed force, conquer'		'To purchase'		
Contextual Meaning			NBM 'manage to attract one's friendly disposition		'To purchase'		

15. Shift of the semantic class of the argument Concrete <-> Abstract: the selectional preference of the contextual meaning (BM or NBM) is restricted to concrete arguments, but takes an abstract argument instead (or vice versa). E.g. 'pokupat (V) svobodu (N) – to buy freedom

	Lyzhnitsa	upustila	zolotuyu	medal.
	S,sg,f, Nom,anim	V,perf,ind, past,sg,f	A,sg, f,Acc	S,sg,f, Acc,inan
	skier	let escape	gold	medal
Met type		LexPers_Source IndirMet →		Metonymy LexPers_Target
Basic Meaning	1)'loose hold on smth, let smth fall' 2)'fail to control smb, let smb leave, run away, swim away etc'			'round metal badge with a relief image awarded for civil or military achievement'
Contextual Meaning	→ NBM 'miss an opportunity'			'rank in a competition marked by a round metal badge'

The skier let the _gold medal_ _escape_.'

METONYMIC SHIFT

"I always thought 'bunny slope' was just an expression."

16. Metonymic shifts: irregular metonymic shifts of any type not attested in the dictionary.
E.g. 'zolotaya (ADJ) medal (N)' – gold medal – which stands for a rank in competition.

MORPHOLOGICALLY MOTIVATED FORMS

'ISCHERPAT' (V)

NBM 'deplete, exhaust smth'

NBM 'use smth up'

NBM 'complete smth, bring smth to an end'

NBM 'come to an end, be used up'

'CHERPAT' (V)

'TO SCOOP, EXTRACT LIQUID OR LOOSE SUBSTANCES (USU. FROM BELOW OR DEEP)'

17. Morphologically motivated forms

When stem(s) of one word is etymologically connected with the stem of another word which possesses stronger basic meaning(s) than the first word, and this connection is felt by the contemporary speakers the language.

E.g. 'ischerpat' (v) has the following dictionary-attested meanings none of which is basic.

NBM 'expend, exhaust smth'

NBM 'use smth up'

NBM 'complete smth, bring smth to an end'

NBM 'come to an end, be used up'

However, the stem of 'ischerpat' is related to the verb of the same stem, 'cherpat', which possesses a strongly basic meaning: 'to scoop, extract liquid or loose substances (usu. from below or deep)'.

	prep	quaslag	quaslag	advmod	predic	advmod	prep	attr
	V	gody	ego	yunosti	devushki	begali	v	domotkanykh sarafanakh.
PR	S,pl,m, Acc,inan	S,sg,m,G en,anim	S,sg,f, Gen,inan	S,pl,f, Nom,anim	V,imperf, ind,past,pl	PR	A,pl,Prp	S,pl,m, Prp,inan
	in	year	his	youth	young woman	run	in	homespun	dress
Met type					SemShift			
Basic Meaning					'rapidly move in a certain direction by alternately putting one foot in front of the other'			
Contextual Meaning					+ 'wear smth, decorate oneself according to the current fashion'			

SEMANTIC SHIFT

'In the years of his youth young women used to run (=wear) in homespun dresses.

18. Semantic shifts.
Non-conventionalized semantic components of a word meaning that are realized in certain constructions.
However, the original semantic content of the source meaning is preserved (unlike in newly attested meanings).
For example, the verb 'begat' – run – in this sentence has a tangible additional semantics of 'wearing' which is realized by the construction N_{ANIM} $V_{'begat'}$ $PRP_{'v'}$ $N_{TYPE\ OF\ CLOTHING}$

NEWLY ATTESTED MEANINGS

BM 'make smth fall down from somewhere'

'Sbrasyvat' (V) – drop smth

NBM 'send via electronic means of communications; record smth to an electronic data storage'

		prep		3-compl		predic		attr		prep	
	Na		sayt		sbrasyvayutsya		poslaniya		na		arabskom.
	Pr		S,sg,m, Acc,inan		V,imperf,pass,ind, nonpast,pl,3pers		S,pl,n, Nom,inan		Pr		A,sg, m,Prp
	to		websites		are dropped		message		in		Arabic
Met type					IndirMet_ NewMeaning						
Basic Meaning					'make smth fall down from somewhere'						
Contextual Meaning					NBM send via electronic means of communications; record smth to an electronic data storage'						

19. Newly attested meanings

Secondary meanings that are not attested in the dictionary (including domain-specific terminology, slang etc.). The original semantic content of the source meaning is replaced by the new semantics (to a varying extent).

REFERENCES

1. Sullivan K. Frames and constructions in metaphoric language. John Benjamins Publishing, 2013. Vol. 14.
2. Sullivan K. Grammar in metaphor: A construction grammar account of metaphoric language. 2007.
3. Cameron L. Metaphor in educational discourse. A&C Black, 2003.
4. Shutova E., Teufel S. Metaphor Corpus Annotated for Source-Target Domain Mappings. // LREC. 2010. Vol. 2. P. 2–2.
5. Steen G. A Method for Linguistic Metaphor Identification: From MIP to MIPVU. John Benjamins Publishing, 2010. 251 p.
6. Shutova E., Sun L., Korhonen A. Metaphor identification using verb and noun clustering // Proceedings of the 23rd International Conference on Computational Linguistics. Association for Computational Linguistics, 2010. P. 1002–1010.
7. Strzalkowski T. et al. Robust extraction of metaphors from novel data // Proceedings of the First Workshop on Metaphor in NLP. 2013. P. 67–76.
8. Li L., Sporleder C. Using gaussian mixture models to detect figurative language in context // Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics. Association for Computational Linguistics, 2010. P. 297–300.
9. Krishnakumaran S., Zhu X. Hunting elusive metaphors using lexical resources // Proceedings of the Workshop on Computational approaches to Figurative Language. Association for Computational Linguistics, 2007. P. 13–20.
10. Neuman Y. et al. Metaphor Identification in Large Texts Corpora // PLOS ONE. 2013. Vol. 8, № 4. P. e62343.
11. Gandy L. et al. Automatic Identification of Conceptual Metaphors With Limited Knowledge. // AAAI. 2013.
12. SynTagRus: the syntactically annotated corpus of Russian [Electronic resource]. URL: http://www.ruscorpora.ru/instruction-syntax.html.
13. Boguslavsky I. et al. Development of a Russian Tagged Corpus with Lexical and Functional Annotation // Proceedings of MONDILEX Third Open Workshop. Bratislava, Slovakia, 2009. P. 194.
14. Dictionary of the Russian Language. 2nd ed. / ed. Yevgenyeva A. Moscow: Academy of Sciences of the USSR; Russian Language Institute, 1981.

THANK YOU

22. This presentation is available for download at: http://web-corpora.net/~badryzlova/AAAI_2017/Presentation