



Онтология на ситуациите за
състояние – лингвистично моделиране.
Съпоставително изследване за
български и руски



Работа по събиране на примери за предикативни конструкции от Българския национален корпус (БНК) и паралелни данни от Руския национален корпус

Цветана Димитрова, Мария Тодорова, Валентина Стефанова
Институт за български език
Българска академия на науките



Онтология на ситуациите за
състояние – лингвистично моделиране.
Съпоставително изследване за
български и руски

Колекция от примери, извлечени от БНК

- Създадена е чрез заявки за търсене.
- Събрани са предикативни конструкции за сегашно, бъдеще и минало време, в които се съдържат винителни и дателни клитики.
- Данните са подбрани ръчно от получените резултати.
- Материалът е представен в табличен вид.
- Към момента:
 - са използвани 34 заявки
 - извлечени са над 10 000 изречения.
- В процес на обработка и допълване.



Онтология на ситуациите за
състояние – лингвистично моделиране.
Съпоставително изследване за
български и руски

Резултати от БНК

заявка БНК	бр. изр.	време / клитика	бележка
<е *{POS=ADV}>	18	Сегашно време	блестящо, болезнено, възможно, лошо, необходимо, несъвместимо, подходящо, препоръчително, различно, разумно, рано, реално, срамно, често
<{*{POS=ADV} е>	7	Сегашно време	вероятно, вярно, обикновено, очевидно, ужасно, често
<ме е *{POS=N}>	3441	Сегашно вр. + винителна	бъз гнус гняв грижа гъдел еня жал мързел смах срам страх шубе яд
<{*{POS=N} ме е>	373	Сегашно вр. + винителна	грижа гъдел срам страх шубе яд
<бъз гнус гняв грижа гъдел еня жал мързел смах срам страх шубе ядте го я ни ви ги е>	413	Сегашно вр.+ винителна	грижа, еня, жал, яд, срам, страх, шубе



Онтология на ситуациите за
състояние – лингвистично моделиране.
Съпоставително изследване за
български и руски

Резултати от БНК

заявка БНК	бр. изр.	време / клитика	бележка
<ме беше *{POS=N}>	383	Минало вр. + винителна	грижа срам страх шубе яд
<{*POS=N} ме беше>	130	Минало вр. + винителна	гнус срам страх шубе яд
<бъз гнус гняв грижа гъдел еня жал мързел смях срам страх шубе яд те го я ни ви ги беше>	181	Минало вр. + винителна	гнус гняв жал смях срам страх яд
<беше те го я ни ви ги бъз гнус гняв грижа гъдел еня жал мързел смях срам страх шубе яд >	138	Минало вр. + винителна	грижа еня жал срам страх яд
<беше ми *{POS=A}>	127	Минало вр. + винителна	гадно, драго, забавно, забранено, изгодно, интересно, комфортно, криво...
<беше ми *{POS=ADV}>	595	Минало вр. + винителна	безразлично, весело, все едно, горещо, добре, досадно, достатъчно...
<{*POS=A} ми беше>	218	Минало вр. + винителна	болно, възможно, досадно, драго, интересно, лесно, необходимо...
<{*POS=ADV} ми беше>	376	Минало вр. + винителна	безразлично, весело, гадно, добре, досадно, достатъчно, жално, забавно...
<{*POS=ADV} ми беше *{POS=A} *{POS=ADV} *{POS=N}>	272	Минало вр. + винителна	безразлично, болно, вкусно, възможно
<беше ти му й ни ви им *{POS=A} *{POS=ADV}>	1199	Минало вр. + винителна	безразлично, забранено, приятно, весело, безинтересно, ясно, гадно...



Онтология на ситуациите за
състояние – лингвистично моделиране.
Съпоставително изследване за
български и руски

Резултати от БНК

заявка БНК	бр. изр.	време	бележка
<*{POS=ADV} ще е>	425	Бъдеще вр.	безнадеждно, безопасно, безполезно, безсмислено, безспорно, важно, великолепно, вероятно, вярно, гадно, далечко, действително ...
<*{POS=A} ще е>	118	Бъдеще вр.	весело, вълнуващо, глупаво, жестоко, интересно, изгодно, лесно, мръсно, необходимо...
<*{POS=A} ще бъде>	133	Бъдеще вр.	
<*{POS=ADV} ще бъде>	472	Бъдеще вр.	
<*{POS=ADV} ще е бъде *{POS=ADV} *{POS=A}>	250	Бъдеще вр.	



Онтология на ситуациите за състояние – лингвистично моделиране.
Съпоставително изследване за български и руски

Резултати от БНК

заявка БНК	бр. изр.	време / клитика	бележка
<бъз гнуc гняв грижа гъдел еня жал мързел смях срам страх шубе яд ще *{POS=P RT=L} е>	1	Бъдеще вр. + винителна	Срам ще ме е
<ще те го я ни ви ги бъде бъз гнуc гняв грижа гъдел еня жал мързел смях срам страх шубе яд>	1	Бъдеще вр. + винителна	Ще ни бъде срам
<ще ме те го я ни ви ги е бъде *{POS=N}>	181	Бъдеще вр. + винителна	
<бъз гнуc гняв грижа гъдел еня жал мързел смях срам страх шубе яд ще те го я ни ви ги бъде>	1	Бъдеще вр. + винителна	срам ще те бъде



Онтология на ситуациите за
състояние – лингвистично моделиране.
Съпоставително изследване за
български и руски

Резултати от БНК

заявка БНК	бр. изр.	време / клитика	бележка
<ще ми е бъде *{POS=A}>	366	Бъдеще вр. + дателна	възможно, мъчно, досадно, достатъчно...
<ще ми е бъде *{POS=ADV}>	502	Бъдеще вр. + дателна	трудно, интересно, неприятно...
<ще ти е бъде *{POS=ADV} *{POS=A}>	483	Бъдеще вр. + дателна	
<ще му е бъде *{POS=ADV} *{POS=A}>	434	Бъдеще вр. + дателна	
<ще й е бъде *{POS=ADV} *{POS=A}>	220	Бъдеще вр. + дателна	
<ще ни е бъде *{POS=ADV} *{POS=A}>	296	Бъдеще вр. + дателна	
<ще ви е бъде *{POS=ADV} *{POS=A}>	449	Бъдеще вр. + дателна	
<ще им е бъде *{POS=ADV} *{POS=A}>	155	Бъдеще вр. + дателна	
<{*{POS=ADV} *{POS=A} ще ти му й ни ви им е бъде>	83	Бъдеще вр. + дателна	весело, полезно, скучно, срамно, тежко...



Онтология на ситуациите за
състояние – лингвистично моделиране.
Съпоставително изследване за
български и руски

Паралелни резултати от РНК



- Колекция от примери с предикативни конструкции за състояние в български и в руски език;
- Типовете са дефинирани в: Тодорова, Димитрова, Стефанова (под печат).
- Въз основа на данните може да настъпят промени.



Паралелни резултати от РНК бг-ру



Онтология на ситуациите за
състояние – лингвистично моделиране.
Съпоставително изследване за
български и руски



- Етап по създаване на база от данни с примери за употреба на предикативи за състояние с прикопулна позиция от:
 - Успореден корпус на руски и български текстове (Гочев 2011) - художествена литература в превод (руско-български, българо-руски): руски словоформи: 2 855 265; български словоформи: 3 230 364; част от Руския национален корпус.



Онтология на ситуациите за
състояние – лингвистично моделиране.
Съпоставително изследване за
български и руски

Паралелни резултати от РНК бг-ру

- Данните са извлечени въз основа на класификация на предикативните конструкции в български, само в текстовете на български (независимо дали оригинални, или преводни)
 - Екскерпирани чрез заявки: 997 примера (двойки) за разновидности на предикативни конструкции (ru-bg: 703; bg-ru: 294):
 - с дателна клитика: -о (*ясно ми е*), N (*жсал ми е*): 230, PP (*за пръв път ми е*): 13; други с дателна клитика (*познат ми е...*): 93;
 - с винителна клитика: N е (*страх ме е*): 127; причастие (*нямало ме е*): 3; -о е (*ясно е*): 464; *добре че*: 95;
 - със съществително N е: *срамота/грях/грехота... е*: 25; *време е*: 8.



Онтология на ситуациите за
състояние – лингвистично моделиране.
Съпоставително изследване за
български и руски

Благодарности



Работата, представена в презентацията, е част от проекта „Онтология на ситуациите за състояние – лингвистично моделиране. Съпоставително изследване за български и руски“ с подкрепата на РФФИ № 20-512-18005 и Фонд „Научни изследвания“ № КП-06-П РУСИЯ-78, 2020 г.



ФОНД
НАУЧНИ
ИЗСЛЕДВАНИЯ

МИНИСТЕРСТВО НА ОБРАЗОВАНИЕТО И НАУКАТА