

XVII международный съезд славистов
Париж, Сорбонна
25–30 августа 2025 г.

Еще раз о лицензировании славянских неопределенных местоимений: корпусное исследование



Universität Regensburg

Roman Fisun
Institut für Slavistik
Fakultät Sprach-, Kultur- und Literaturwissenschaften

DFG Deutsche
Forschungsgemeinschaft

План доклада

1. Лицензирование неопределенных местоимений
2. Цели исследования
3. Данные и методология
4. Результаты: корпусные наблюдения и кластерный анализ
5. Основные выводы

DFG-проект (рук. Б. Ханзен, Университет Регенсбурга):
«Составные местоимения в славянских языках.
Семантическая карта неопределенности второго
поколения» (HA 2659/10-1)

1. Лицензирование неопределенных местоимений

Неопределенные местоимения (НМ)

- референтные: употребляются свободно, но типичны для эпизодических контекстов

(1) UA *Щось (*що-небудь) різко зашурхотіло праворуч*

[...] (ukTenTen20)

- Нереферентные неопределенные местоимения: требуют особого контекста = **лицензирования**

(2) UA *Якщо ви по **яких-небудь** причинах вибрали другий*

варіант, то вам варто читати далі. (ukTenTen20)

- В (2) лицензирующий контекст – условное придаточное. Среди других – разные типы отрицания (\sim NPIs), модальные контексты (\sim FCIs, NSIs) и некоторые другие

Традиционное представление – качественное лицензирование: «элементы класса X допустимы в контекстах A, B, C, но исключены в других»

Основные подходы к анализу лицензирующих контекстов:

- Неверидиктальность (Giannakidou 1998)
- Контексты снятой утвердительности (КСУ) у Падучевой (2015, 2018)

Контексты снятой утвердительности (КСУ)

Определение:

- Контексты, где говорящий не берет на себя ответственность за достоверность высказывания

Основные характеристики:

- 25 типов КСУ: отрицания, вопросы, модальные конструкции, стандарт сравнения и др.
- На русском материале Падучева (2015, 2018) выделяет 4 класса НМ:
 - референтные: *кое-кто, кто-то*
 - нереферентные неопределенные (NSIs): *кто-нибудь*
 - отрицательной поляризации (NPIs): *кто-либо*
 - свободного выбора (FCIs): *кто угодно, любой*

Контексты снятой утвердительности (Падучева 2015, 2018)

Universität Regensburg

	Референтные	NSI	NPI	FCI
0.1. Утвердительный контекст	+	-	-	
1.1. Прямое отрицание		-		+
1.2.1. Отрицание вышестоящего инфинитива		-		
1.2.2. Отрицание вышестоящего предиката с союзом		-		
1.5. Предлог с имплицитным отрицанием		-	+	
4.2. Деепричастный оборот				+
9.1. Желание		+	-	+
9.2. Просьба, предложение, императив		+	-	+
10. Неуверенность: предположение, нереальность, мнение		+	-	+
11. Стандарт сравнения				

Таблица 1. Лицензирование различных типов неопределенных местоимений по данным Падучевой



Universität Regensburg

2. Цель исследования

Проблема:

- Существующие исследования по лицензированию славянских НМ не опираются на корпусные данные.
- Корпусный материал чаще используется только для иллюстрации теоретических результатов.

Исследовательские вопросы:

1. Можно ли выявить лицензирование славянских НМ в количественных корпусных данных?
2. Если да, какой характер лицензирования наблюдается?
3. Позволяет ли корпусный анализ определять типы лицензирования и, следовательно, типы НМ в славянских языках?

3. Данные и методология

Данные

- 5 славянских языков: **русский, украинский, польский**; хорватский (данные и анализ Б. Ханзена), чешский (данные и анализ М. Рыбовой)
- основано на корпусных данных больших славянских веб-корпусов, доступных через SketchEngine:

Язык	Корпус	Размер, токенов
русский	ruTentTen11	18.280.486.876
украинский	ukTentTen20	3.282.586.754
польский	plTentTen12	9.387.142.186
чешский	csTentTen17	12.586.415.546
хорватский	hrWac 2.2	1.405.794.913

Неопределенные местоименные серии

- 51 серия кодифицированных (выделены полужирным) и некодифицированных (=«новых составных») неопределенных местоимений

	русск.	украинск.	польск.	чешск.	хорватск.
префиксальные	абы Бог весть черт-те мало кое- не пойми	багато казна- невідомо мало Бог зна(є)	bądź (nie)byle cholera wie lada mało	būchví: jedno kdoví: lhostejno málo: nevím: sotva:	bilo Bog zna gdje- i- koje- malo mnogo
постфиксальные	бы то ни было -то -либо -нибудь угодно попало	завгодно попало б (то) не було -сь -небудь	bądź popadło -ś -kolwiek śkolwiek	chceš/chcete -koli(v) je libo	-god bilo

Обработка данных и анализ

Формализация 26 контекстов (CQL-запросы)

Извлечение частот 51 серии местоимений

ручная очистка данных

нормализация (mod. PMI, шкала 0–10)

иерархический кластерный анализ

4. Результаты исследования

4.1. Корпусные наблюдения

Universität Regensburg

Наблюдается лишь количественное, но не качественное лицензирование:

specific known, модальный контекст

(1) Мне нужно кое-что принести. (ruTenTen11)

existential, эпизодический контекст

(2) В современном мире значение признанных авторитетов в какой-либо сфере деятельности многократно возросла.

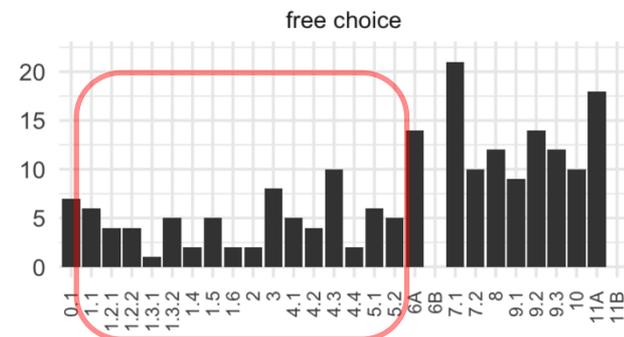
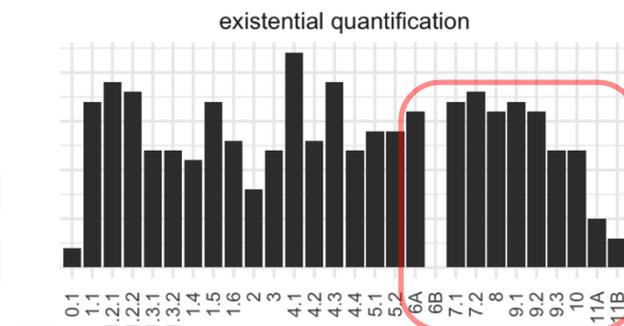
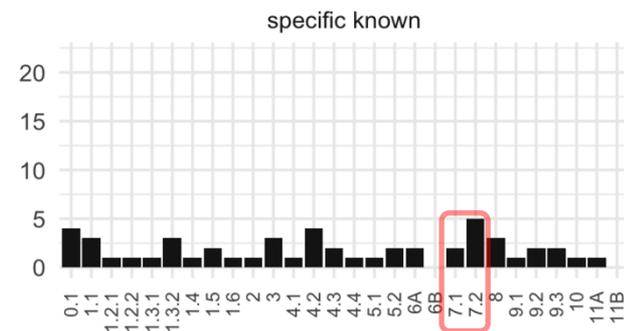
(ruTenTen11)

existential, модальный контекст

(3) Доп.плюсом будет знание какой-либо системы бронирования (сейчас у вас есть время пройти обучение). (ruTenTen11)

free choice, отрицание вышестоящего предиката

(4) Почти в любой "аналитической" передаче ТВ они не стесняются смачно плюнуть во что угодно русское - историю ли, национальный русский характер ли. (ruTenTen11)



Количество зафиксированных контекстов для местоимений разных семантических типов

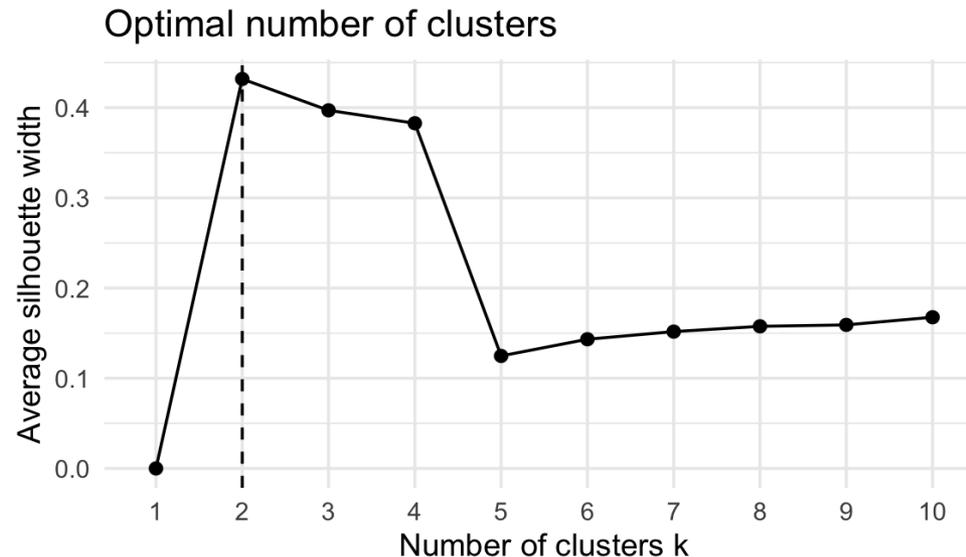
4.2. Кластерный анализ

Иерархическая кластеризация

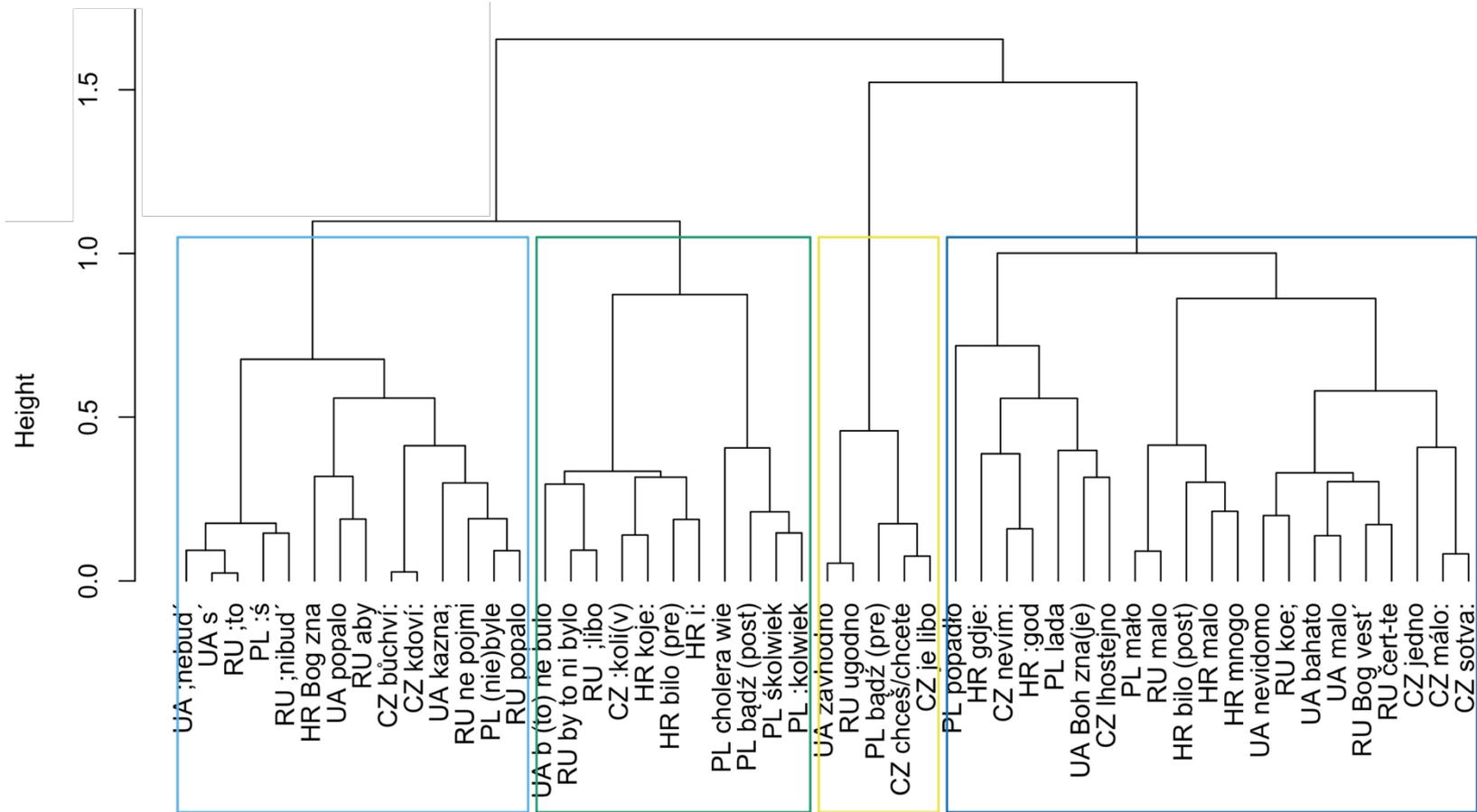
- Данные: относительные частоты в контекстах (mod. PMI)
- Метрика: корреляция Пирсона (сравнение профилей, а не значений)
- Метод: кластеризация по методу Уорда (компактные кластеры, меньше внутрикластерная дисперсия)
- Результат: группировка серий по сходству распределений

Силуэтный анализ кластеров

- Оптимальное разбиение: 2 кластера \Rightarrow слабая структурированность данных
- Для дальнейшего анализа выбрано 4 кластера (порог = 0,25)



Дендрограмма местоименных кластеров по контекстам употребления



distances_P
hclust (*, "ward.D2")

Universität Regensburg

- Кластер A: разнородна, делится на NSIs и местоимения случайного выбора; деление статистически неустойчиво, но отражает теорию
- FCIs (кластер C) образуют чёткий кластер (исключение: *CZ koli(v)*)
- Другие кластеры: содержат смешанные типы

Итог:

- ✓ только FCIs показывают статистически надёжную дистрибуцию
- ✗ для других классов лицензирование – скорее тенденция, чем правило

A	B	C	D
UA nebud´	UA b (to) ne	UA zavhodno	PL popadło
UA -s´	bulo	RU ugodno	HR gdje
RU to	RU by to ni bylo	PL bądź (pre)	CZ nevím
PL -ś	RU libo	CZ	HR god
RU nibud	CZ koli(v)	chceš/chcete	PL lada
	HR koje	CZ je libo	UA Boh zna(je)
	HR bilo (pre)		CZ lhostejno
	HR i		PL mało
	PL cholera wie		RU malo
	PL bądź (post)		HR bilo (post)
HR Bog zna	PL školwiek		HR malo
UA popalo	PL kolwiek		HR mnogo
RU aby			UA nevidomo
CZ bůchví			RU koe
CZ kdoví			UA bahato
UA kazna			UA malo
RU ne			RU Bog vest´
pojmi			RU čert-te
PL (nie)byle			CZ jedno
RU popalo			CZ málo
			CZ sotva
14	11	5	21



Universität Regensburg

5. Выводы

1. Ограничение дистрибуционного анализа НМ

Определение семантического типа славянских НМ только по их контекстуальной дистрибуции невозможно (исключение – FCIs)

2. Исключение – FCIs, они формируют формируют статистически надёжный кластер

3. Лицензирование количественное, а не качественное

Различия между контекстами проявляются в частотах, но не в абсолютной запретности или допустимости

4. Лицензирование – тенденция, а не правило

Данные крупных корпусов показывают скорее склонность к определённым контекстам, чем жёсткое правило

5. Необходимость пересмотр теории

Существующие модели лицензирования требуют уточнения с учётом количественных корпусных данных

Литература

- Падучева, Елена Викторовна (2018) Русские местоимения свободного выбора. *Russian Linguistics* 3(42), 291–319
- Падучева, Елена Викторовна (2015) Снятая утвердительность и неверидиктальность (на примере русских местоимений отрицательной полярности). *Russian Linguistics* 2(39), 129–162
- Giannakidou, Anastasia (1998) *Polarity sensitivity as (non)veridical dependency*. Amsterdam
- van der Wouden, Ton (1997) *Negative Contexts Negative Contexts. Collocation, Polarity and Multiple Negation*. New York



Universität Regensburg

DFG Deutsche
Forschungsgemeinschaft

Спасибо за внимание!