

Матюхина Екатерина Петровна

**СООТНОШЕНИЕ АРЕАЛЬНОЙ И ГЕНЕТИЧЕСКОЙ
БЛИЗОСТИ ЯЗЫКОВ В ВЫБОРКЕ
ДЛЯ ЛЕКСИКО-ТИПОЛОГИЧЕСКОГО ИССЛЕДОВАНИЯ**

Научный руководитель: к.ф.н., доцент Резникова Т. И.

8 июня 2023, Школа лингвистики НИУ ВШЭ

СОДЕРЖАНИЕ

- Фреймовый подход: back to basics
- Количественные исследования в лексической типологии
- Исследовательский вопрос: истоки
- Сбор данных
- Анализ и результаты
- Выборки: как выбрать
- Обсуждение результатов

ФРЕЙМОВЫЙ ПОДХОД

(Рахилина, Резникова 2013)

- Изучение значения с опорой на языковое поведение:
анализ сочетаемости
- Источники: словари и корпуса + опрос носителей
- **фреймы** – отдельные ситуации внутри поля, которые объединяют некоторые значения релевантных для противопоставления внутри поля параметров
- Несколько фреймов может покрываться одной и той же лексемой – **колексификация**
- Колексификации **неслучайны** – они свидетельствуют о когнитивной близости фреймов



ВЫБОРКА: ФРЕЙМОВЫЙ ПОДХОД

-  1 Определение границ поля на материале родного языка исследователя
-  2 Пилотное исследование на доступных исследователям языках: работа со словарями и корпусами + опрос носителей
-  3 Финализация фреймовой структуры поля и лексико-типологической анкеты
-  4 Экстенсивный сбор материала разных языков

(Koptjevskaya et al. 2007)

Given the various methodological problems and our general poor knowledge of lexicon-rooted phenomena, a large part of lexical-typological research in the nearest future will, most probably, **have to be restricted to small numbers of languages**. This might be a disappointment for some typologists, but it is undoubtedly better to have high-quality research on just a few languages than poor investigations on scores

МАТЕРИАЛЫ ПРЕДШЕСТВЕННИКОВ

- **Параллельные корпуса:** переводы Библии
(Cysouw, Wälchli 2007) (Östling 2016)

- **DatSemShifts** (Zalizniak et al 2016-2020)
(Fedotova 2022)
Диахроническая семантическая реконструкция: наблюдение за развитием семантических сдвигов в 5 уральских языках

МАТЕРИАЛЫ ПРЕДШЕСТВЕННИКОВ

- **CLICS**: Database of Cross-Linguistic Colexifications (Rzymiski et al. 2019)
- **ASJP**: Automated Similarity Judgment Program (Wichmann, Holman and Brown 2020)

(Gast, Koptjevskaja-Tamm 2018)

Выявлены ареальные закономерности распределения отдельных кластеров колексификации, в частности, терминов родства и частей тела

(Georgakopoulos et al. 2021)

Выявлены ареальные закономерности распределения для глаголов восприятия и познания, совпадающие с ареалами, обычно выделяемыми в типологической литературе

ИССЛЕДОВАТЕЛЬСКИЙ ВОПРОС

(Matras, Sakel 2004)

Matter-borrowing – заимствование фонематических единиц языкового материала.

- заимствования отдельных слов (lexical borrowing)

Pattern-borrowing – заимствование схем структурных и дистрибутивных отношений между этими единицами.

- лексические кальки (loan translations)
- заимствование модели полисемии (loan meaning extension)

ЯЗЫКОВЫЕ КОНТАКТЫ

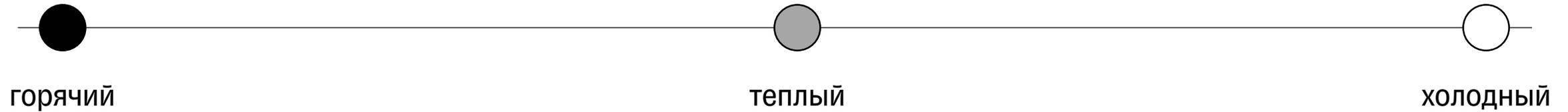
(Бакланов 2021): ‘менять’ в ЭВЕНСКОМ

Словари: особый глагол для фрейма модификации

Узус: доминантная система, все фреймы покрываются глаголом *gajdaj*

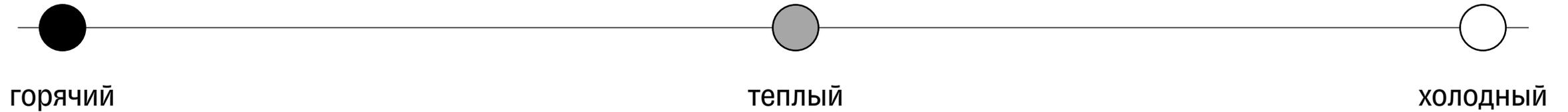
ЯЗЫКОВЫЕ КОНТАКТЫ

(Koptjevskaya 2022): температурные прилагательные в славянских языках



ЯЗЫКОВЫЕ КОНТАКТЫ

(Koptjevskaya 2022): температурные прилагательные в славянских языках



но не в **македонском**: влияние греческого?



БЛИЗКОРОДСТВЕННЫЕ ЯЗЫКИ

представляют особый интерес для лексической типологии: в родственных языках может быть разный набор лексем, разные параметры организации поля и, как следствие, разные типы колексификаций:

- вращение в польском и русском (Рахилина, Прокофьева 2004)
- качание в польском и русском (Рахилина, Прокофьева 2005)
- ‘резать/ломать’ в германских (Majid et al. 2007)
- ‘плавать’ в германских и славянских (Koptjevskaja et al. 2010),
- ‘падать’ в уральских языках (Кашкин 2013)
- ‘литься/сыпаться’ в славянских языках (Федоров 2021)
- ‘менять’ в финском и карельском (Лапшина, Копылова, Матюхина 2022)

ГИПОТЕЗА

Ареальная близость языков влияет на сходство их лексических систем с точки зрения колексификации больше, чем генетическая близость

БОРОТЬСЯ И ИСКАТЬ

Часть 1

ДАННЫЕ

Поле	 'плавать'	 'падать'	 'прятать'
Источник	(Майсак, Рахилина 2007)	(Рахилина и др. 2007)	(Резникова 2022)
Языков	42	36	26
Семей	11	11 + 1 изолят	7 + 1 изолят
Ареалов	5	7	4

ДАННЫЕ

Frame	Glottocode	Area	Family	Genus	tag...2
<chr>	<chr>	<chr>	<chr>	<chr>	<chr>
English	stan1293	Europe	Indo-European	Germanic	loose dominant
Norwegian	norw1258	Europe	Indo-European	Germanic	loose dominant
Modern Greek	mode1248	Europe	Indo-European	Greek	strict dominant
French	stan1290	Europe	Indo-European	Romance	strict dominant
Spanish	stan1288	Europe	Indo-European	Romance	strict dominant
Basque	basq1248	Europe	Isolate	Isolate	strict dominant
Russian	russ1263	Europe	Indo-European	Slavic	2falling-specified
Serbian-Croatian-Bosnian	sout1528	Europe	Indo-European	Slavic	2falling-specified
Finnish	finn1318	Europe	Uralic	Finnic	3higher+detachment
Kazym-Berezover-Suryskarer Khanty	khan1273	Siberia	Uralic	Ugric	2falling-specified
Kazakh	kaza1248	Asia	Altaic	Turkic	2detachment-specified

ИНСТРУМЕНТЫ АНАЛИЗА

Строим модели распределения типов система в зависимости от двух факторов: ареала и семьи –
мультиномиальная логистическая регрессия

Оцениваем, какая модель объясняет данные лучше –
Критерий ANOVA

‘ПЛАВАТЬ’

Статистически значимых различий между моделями не выявлено

Model	Resid. df	Resid. Dev	Test	Df	LR stat.	Pr(Chi)
<chr>	<dbl>	<dbl>	<chr>	<dbl>	<dbl>	<dbl>
as.factor(Area)	228	116.39860		NA	NA	NA
as.factor(family)	180	67.99134	1 vs 2	48	48.40727	0.4563991

Таблица 1. Значения ANOVA для поля ‘плавать’

‘ПАДАТЬ’

Аналогично

Model	Resid. df	Resid. Dev	Test	Df	LR stat.	Pr(Chi)
<chr>	<dbl>	<dbl>	<chr>	<dbl>	<dbl>	<dbl>
as.factor(Area)	162	82.92151		NA	NA	NA
as.factor(Family)	126	54.19522	1 vs 2	36	28.72629	0.8003323

Таблица 2. Значения ANOVA для поля ‘падать’

‘ПРЯТАТЬ’

Удивительно....

Model	Resid. df	Resid. Dev	Test	Df	LR stat.	Pr(Chi)
<chr>	<dbl>	<dbl>	<chr>	<dbl>	<dbl>	<dbl>
as.factor(Area)	154	84.67126		NA	NA	NA
as.factor(Family)	119	61.53811	1 vs 2	35	23.13316	0.9379989

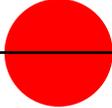
Таблица 3. Значения ANOVA для поля ‘прятать’

ПОПРОБУЕМ ПО-ДРУГОМУ...

Строим модели распределения типов система в зависимости от двух факторов: ареала и семьи – **мультиномиальная логистическая регрессия**

Оцениваем, какая модель =теряет меньше подробностей – **Информационный критерий Акаике (AIC)**

AIC: по дельте значений

Поле	 'плавать'	 'падать'	 'прятать'
Модель с меньшим AIC	генетическая	ареальная	ареальная
Семей	11	11 + 1 изолят	7 + 1 изолят
Ареалов	5	7	4

НЕ НАЙТИ И НЕ СДАВАТЬСЯ

Часть 2

*просчитался,
но...
где?*



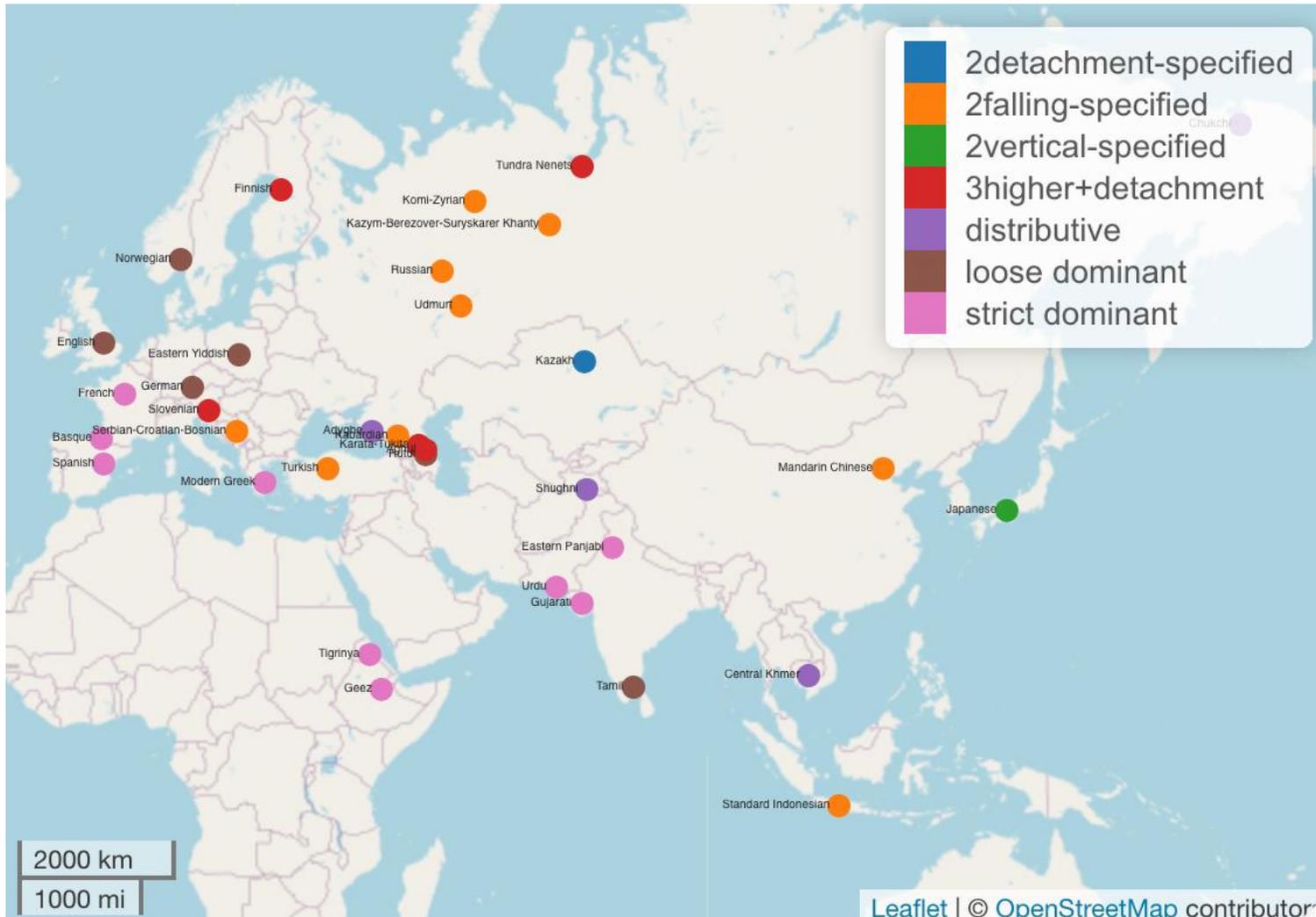
ВЫБОРКА: ФРЕЙМОВЫЙ ПОДХОД

- 1 Определение границ поля на материале родного языка исследователя
- 2 Пилотное исследование на доступных исследователям языках: работа со словарями и корпусами + опрос носителей
- 3 Финализация фреймовой структуры поля и лексико-типологической анкеты
- 4 Экстенсивный сбор материала разных языков



КАК СЛЕДСТВИЕ...

- Convenience bias



Afro-Asiatic	2
Altaic	2
Austroasiatic	1
Austronesian	1
Chukotko-Kamchatkan	1
Dravidian	1
Indo-European	15
Isolate	1
Japonic	1
Nakh-Daghestanian	3
Northwest Caucasian	2
Sino-Tibetan	1
Uralic	5

И СНОВА О ВЫБОРКАХ

(Rijkhoff et al. 1993)

- **Probability samples** нужны для статистической проверки гипотез о наличии статистических тенденций явлений или выведения корреляций между явлениями: должны быть сбалансированы с точки зрения ареальной и генетической близости языков
- **Variety samples** призваны отразить многообразие способов выражения некоторого лингвистической черты, включая наиболее редкие. Могут включать и близкие ареально и генетически языки.

обобщение

данные



гипотеза

данные

проверенная
гипотеза

ВЫВОДЫ

- 0 Проверять, подходят ли данные для ответа на исследовательский вопрос ДО начала исследования
- 1 Данные, собранные в ходе проектов Московской лексико-типологической группы, не годятся в качестве материалов для выявления типологически релевантных статистических тенденций: они не представляют в должной мере языкового разнообразия и не сбалансированы относительно ареальной и генетической близости входящих в них языков
- 2 Однако, эти данные подходят для количественного анализа с целью выявления дополнительных связей между данными: например, обнаружения новых контактных или генетических явлений
- 3 Принцип создания выборки, принятый в рамках фреймового подхода, полностью соответствует принципу создания variety sample
- 4 Данные, собранные во фреймовом подходе, отлично подходят для типологических статистических исследований: так как все единицы собираются по единой анкете, это значительно снижает вероятность «зашумления» данных (ср. CLICS и ASJP)

СПИСОК ЛИТЕРАТУРЫ

1. Бакланов 2021 — Бакланов А.Е. Поле 'менять' в эвенском языке. Доклад на Открытом семинаре по лексической типологии. Москва, НИУ ВШЭ, 2021
2. Кашкин 2013 — Кашкин Е.В. Языковая категоризация фактуры поверхностей (типологическое исследование наименований качественных признаков в уральских языках): дис. ... канд. филол. наук. Москва, 2013.
3. Лапшина, Копылова, Матюхина (рукопись) — Лапшина К.М., Копылова Е. В., Матюхина Е.П. Система глаголов поля менять в финском языке. На правах рукописи.
4. Майсак, Рахилина 2006 — Майсак Т.А., Рахилина Е.В. (ред.) Глаголы движения в воде: лексическая типология. М: Индрик, 2007.
5. Рахилина, Прокофьева 2004 — Рахилина Е. В., Прокофьева И. А. Родственные языки как объект лексической типологии: русские и польские глаголы вращения // Вопросы языкознания. 2004. № 1. С. 60—78.
6. Рахилина, Прокофьева 2005 — Рахилина Е. В., Прокофьева И. А. Русские и польские глаголы колебательного движения: семантика и типология // Топоров В. Н. (ред.). Язык. Личность. Текст. Сб. к 70-летию Т. М. Николаевой. М.: Языки славянских культур, 2005. С. 304—314.
7. Рахилина, Резникова 2013 — Рахилина Е.В., Резникова Т.И. Фреймовый подход к лексической типологии // Вопросы языкознания. 2013. № 2. С. 3—31.
8. Рахилина и др. 2020 — Рахилина Е. В., Резникова Т. И., Рыжова Д. А. Глаголы падения в языках мира: фреймы, параметры и типы систем // Acta Linguistica Petropolitana. 2020. XVII(1). С. 3—43.
9. Резникова 2022 — Резникова Т. И. Глаголы прятания: типология систем // Вопросы языкознания. 2022. Т. 4. С. 66-94. Cysouw, Wälchli 2007 — Cysouw, M., & Wälchli, B. (2007). Parallel texts: Using translational equivalents in linguistic typology. Language typology and universals, 60(2), 95-99.
10. François 2008 — François, A. (2008). Semantic maps and the typology of colexification. From polysemy to semantic change: Towards a typology of lexical semantic associations, 106, 163.
11. Gast, Koptjevskaja-Tamm 2018 — Gast, V. & Koptjevskaja-Tamm, M. (2019). The areal factor in lexical typology. In D. Olmen, T. Mortelmans & F. Brisard (Ed.), Aspects of Linguistic Variation, 43-82
12. Georgakopoulous et al. 2022 — Georgakopoulos, T., Grossman, E., Nikolaev, D. & Polis, S. (2022). Universal and macro-areal patterns in the lexicon: A case-study in the perception-cognition domain. Linguistic Typology, 26(2), 439-487.
13. Haspelmath, Tadmor 2009 — Haspelmath, M., & Tadmor, U. (Eds.). (2009). Loanwords in the world's languages: a comparative handbook. Walter de Gruyter.
14. Koptjevskaja et al. 2010 — Koptjevskaja-Tamm, M., Divjak, D., Rakhilina, E. V., Hasko, V., & Perelmutter, R. (2010). Aquamotion verbs in Slavic and Germanic. New approaches to Slavic verbs of motion, 315-341.
15. Koptjevskaya 2015 — Koptjevskaja-Tamm M.(ed.). The linguistics of temperature. Amsterdam: John Benjamins
16. Koptjevskaya 2022 — Koptjevskaja-Tamm, M. (2022). Talking temperature with close relatives. The typology of physical qualities, P. 133-215.
17. Koptjevskaya et al. 2007 — Koptjevskaja-Tamm, M., Vanhove, M. & Koch, P. (2007). Typological approaches to lexical semantics. Linguistic Typology, 11(1), 159-185.
18. Majid et al. 2007 – Majid, A., Gullberg, M., Staden, M. & Bowerman, M. (2007). How similar are semantic categories in closely related languages? A comparison of cutting and breaking in four Germanic languages. Cognitive Linguistics, 18(2), 179-194.
19. Matras, Sakel 2004 — Matras, Y., & Sakel, J. (2007). Investigating the mechanisms of pattern replication in language convergence. Studies in Language. International Journal sponsored by the Foundation "Foundations of Language", 31(4), 829-865.
20. Östling 2016 — Östling, R. (2016). Studying colexification through massively parallel corpora. The lexical typology of semantic shifts, 58, 157. Rijkhoff et al. 1993 — Rijkhoff, J., Bakker, D., Hengeveld, K., & Kahrel, P. (1993). A method of language sampling. Studies in Language. International Journal sponsored by the Foundation "Foundations of Language", 17(1), 169-203.

ОБСУЖДЕНИЕ

(не только результатов)