

**Конструкционный подход к переключению кодов:
сочетания ADJ + N с переключением между русским
и южно-тунгусскими языками**

Н. М. Стойнова
ИРЯ РАН, Москва
stoynova@yandex.ru

РусКонстр-5, ИЛИ РАН, 03-05.10. 2019
Исследование поддержано грантом РФФИ №18-312-00155 мол_а

Введение

- **Переключение кодов** (code-switching, CS):
 - единицы из L1 и L2 в одном тексте / (одном предложении) / ((одном слове)).

Ср. предложение с переключением кодов между ульчским (тунгусо-маньчжурские) и русским:

(1) *dujs tučam̄i ti солдат-sal онять mewtičit* (ulc, oax)

‘Убежишь в лес, а эти **солдаты онятно** будут стрелять’

Введение

- **Конструкционные подходы** (Fillmore, Goldberg, Croft, ...):
 - единица морфосинтаксической структуры - “конструкция” (форма \Leftrightarrow значение);
 - нет жесткого противопоставления между “синтаксисом” и “лексиконом”;
 - нет жесткой иерархии единиц (конструкций);
 - важна степень конвенционализации конструкции (\sim частотность);
 - эмпирически (корпусно-) ориентированные подходы, квантитативные методы.

Введение

Языковые контакты и переключение кодов - не самая освоенная область для грамматики конструкций.

Но такие попытки есть:

- Höder 2012; Ziegler 2015; Hilpert & Östman 2016; Wasserscheidt 2016; Boas & Höder 2018.

Здесь:

- Материал русского языка в контакте с южно-тунгусскими:
 - разные диалекты нанайского и ульчский;
 - в той или иной степени под угрозой;
 - все носители знают и большинство более активно использует русский.

План доклада

1. Именная группа в русском vs. в тунгусо-маньчжурских языках
2. Именная группа в существующих теориях переключения кодов
3. Переключение кодов в ИГ в терминах конструкций
4. Пример анализа в духе конструкционного подхода: конструкция ADJ + N с переключением кодов “южно-тунгусские vs. русский”
5. Результаты и обобщение

1. Именная группа в русском vs. в тунгусо-маньчжурских языках

1. Именная группа в русском vs. в южно-тунгусских языках

	RUS	TUNG
порядок слов	ADJ + N	ADJ + N
вершина: словоизм.	N-CASE.NUM	N-(NUM)-CASE-(PERS)
вершина: род	m, f, n (согласование)	--
адъектив	N-GNR.CASE.NUM	N (неизм.)

masi tawa-do (сильный огонь-DAT)

(на) *сильн-ом огн-е* (сильный-M.LOC.SG огонь-LOC.SG)

2. Именная группа в существующих теориях CS

2. Именная группа в существующих теориях CS

Matrix Language Frame model (начиная с Myers-Scotton 1993):

- L1 и L2 не симметричны (материал L1 вставляется в структуру L2).
- В т.ч. род, число, падеж прилагательного должны приписываться “матричным языком”, и сами показатели должны быть из “матричного языка”.

сорок четверт-ом а́пан-ду 'в 44-ом году' (ulc, aid)

- Почему русский аффикс? почему падеж у прилагательного? откуда род?
- Есть оговорки и дополнения к модели (острова включенного языка, классическое vs. сложное переключение кодов), но...

2. Именная группа в существующих теориях CS

Генеративный подход: минимализм (начиная с McSwan 1999):

- Общая синтаксическая структура, специфичные для L1 и L2 признаки в лексиконе.
- В т.ч. согласование единицы L1 с единицей L2 возможно, если совпадает набор признаков (сущ. из L1 может задать адъективу L2 предл.п., если в L1 и L2 есть предл.п.)

сорок четверт-ом аһан-ду 'в 44-ом году' (ulc, aid)

- Падеж у прилагательного - ок, но откуда род? дефолтный?

кажд-ую һәһһә 'каждую весну' (gld, rchk)

- Откуда род?

2. Именная группа в существующих теориях CS

⇒ Есть проблемные случаи, для которых существующие теории “слишком жесткие”, особенно:

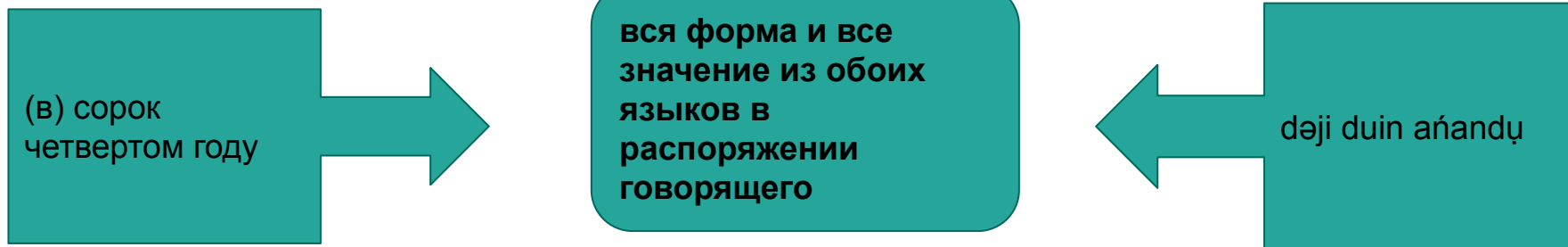
- в языках с богатой морфологией;
- в ситуации языкового сдвига.

См., напр., обзор очень похожих сложных случаев в ИГ при переключении кодов между русским и прибалтийско-финскими в Муслимов (в печати).

Эталон работы о переключении кодов в ИГ:

- какая из теорий лучше справится с новыми фактами, выходящими за их рамки (см., напр., Vanden Wyngaerd (2016): голландский vs. французский).

3. Переключение кодов в ИГ в терминах конструкций



ADJ _{RUS}	INFL _{RUS}	N _{RUS}	INFL _{RUS}
ADJ _{TUNG}	—	N _{TUNG}	INFL _{TUNG}

сорок четверт-ом

ańan-du

4. Пример анализа: конструкция ADJ + N с переключением кодов “южно-тунгусские vs. русский”

План

4.1. Данные: отбор и классификация

4.2. Типы конструкций ADJ + N с переключением кодов

4.3. Словоизменительные характеристики адъектива (согласование по роду, числу, падежу)

4.4. Лексический состав конструкции

4.1. Данные: отбор и классификация

4.1. Данные

Тексты:

- на разных диалектах нанайского и на ульчском, всего ок. 34 ч.
- собраны совместно с С.А. Оскольской в 2011-2019 гг.
- в большинстве в том или ином объеме переключение кодов
- ELAN-файлы с расшифровкой + пометами CS (ок. 2500 случаев)

4.1. Данные

Выборка сочетаний ADJ + N с переключением кодов

- все русские сочетания между тунгусскими (или с одной стороны граница клаузы)

(1) *balana čadu буддийский храма=da biči* ‘раньше там же был буддийский храм’ (gld, ivg)

- (все тунгусские сочетания между русскими)

- все сочетания русский адъектив + тунгусское существительное

(2) *aktə разный aktə aŋgosini* ‘женщины, разные женщины делают’ (gld, anvs)

- все сочетания тунгусский адъектив + русское существительное

(3) *dāi крокодил bilə* ‘большой крокодил, наверное’ (ulc, ok)

Всего: **318** вхождений.

4.1. Данные

NВ адъектив:

- все, что и в том, и в другом языке ведет себя как прилагательное (признаковое слово типа *хороший, большой*)
- + кванторы, детерминанты (*тот, этот*), порядковые числительные
- напр., не количественные числительные (как прилагательное в тунгусских, но не в русском)

NВ переключение кодов:

- включаются все случаи “внутрисловного переключения” (*рыба-ва*)
- случаи с некоторой степенью фонетической адаптации (*храма*)
- все, кроме “старых заимствований” (типа *gumaska* ‘рубль’)

4.2. Типы конструкций ADJ + N с переключением кодов

4.2. Типы конструкций ADJ + N

Самый частотный:

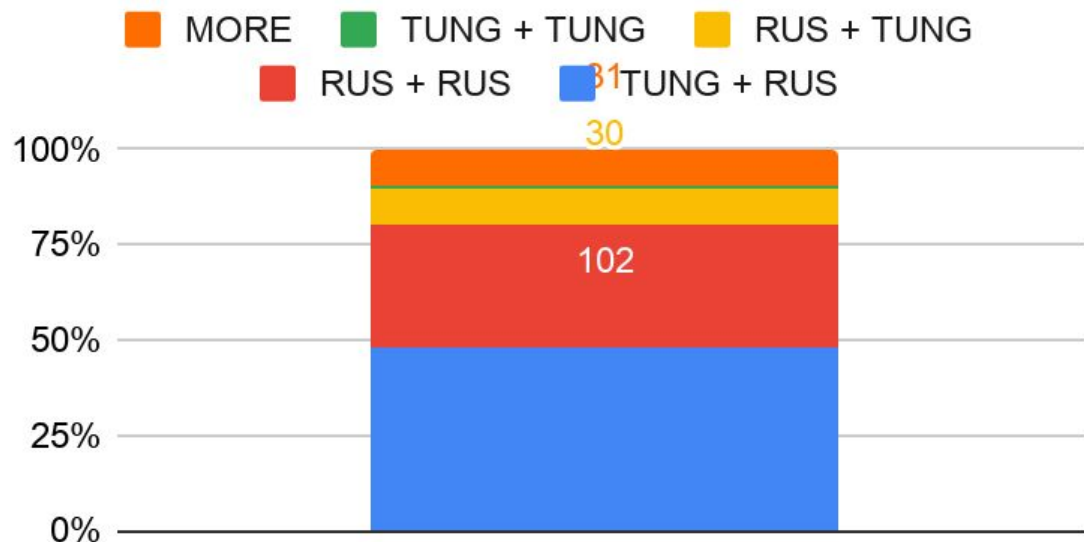
- тунгусский адъектив с русской вершиной.

Самый редкий:

- русский адъектив с тунгусской вершиной.

Не рассматриваем:

- больше двух компонентов,
- чисто тунгусские.



Сочетание ADJ + N: Распределение по типам

4.3. Словоизменительные характеристики адъектива (подробнее о падеже)

4.3. Словоизменение адъектива

ADJ _{RUS}	INFL _{RUS}	N _{RUS}	INFL _{RUS}
ADJ _{TUNG}	—	N _{TUNG}	INFL _{TUNG}

Заполнение тунгусским и русским материалом двух слотов существительного и двух слотов адъектива.

Без предустановленных ограничений.

Ожидания: будут заполняться без очевидных конфликтов между слотами N и ADJ.

Заполнение словоизменительного слота ADJ по тунгусской модели:

- дефолтное число sg, род m, падеж nom (=асс для части парадигмы):
большой (род, число, падеж) / *большая* (число, падеж) / *большие девочками* (падеж)

4.3. Падеж: ADJ_{rus} + N_{tung}

(a) *каждый tūli-du-j* (gld, pnv)
'в каждом дворе.DAT.REFL'

ADJ _{RUS}	INFL _{RUS}	N _{RUS}	INFL _{RUS}
ADJ _{TUNG}	—	N _{TUNG}	INFL _{TUNG}

(b) *тридцать седьмом añan-du* (gld, rchk)
'в 37-м году'

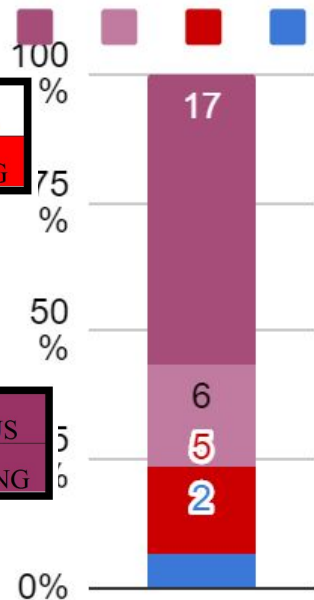
ADJ _{RUS}	INFL _{RUS}	N _{RUS}	INFL _{RUS}
ADJ _{TUNG}	—	N _{TUNG}	INFL _{TUNG}

(c) *стоячий xorqo-sal-ba* (ulc, ulc)
'стоячие комоды.ACC'

ADJ _{RUS}	INFL _{RUS}	N _{RUS}	INFL _{RUS}
ADJ _{TUNG}	—	N _{TUNG}	INFL _{TUNG}

(d) *молодой naj* (gld, itg)
'молодой человек'

ADJ _{RUS}	INFL _{RUS}	N _{RUS}	INFL _{RUS}
ADJ _{TUNG}	—	N _{TUNG}	INFL _{TUNG}



ОЖИДАНИЕ:

ADJ в дефолтной немаркированной форме (nom~acc)
по тунгусской модели (a) **НЕТ**

3.3. Падеж: ADJtung + Nrus

(a) *dāi трактор-ži* (ulc, spk)
‘на большом тракторе.INS’

ADJ _{RUS}	INFL _{RUS}	N _{RUS}	INFL _{RUS}
ADJ _{TUNG}	—	N _{TUNG}	INFL _{TUNG}

(b) **dāi трактором*
‘большим трактором’

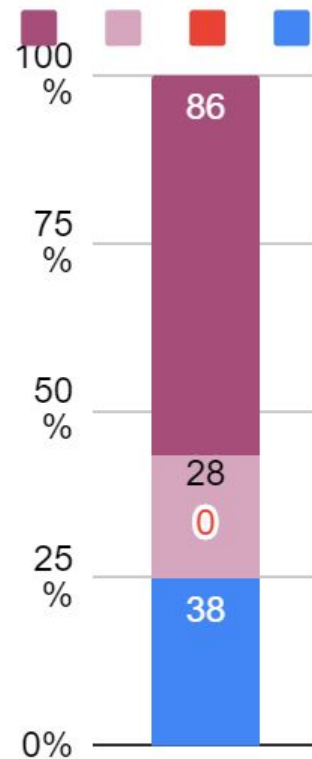
ADJ _{RUS}	INFL _{RUS}	N _{RUS}	INFL _{RUS}
ADJ _{TUNG}	—	N _{TUNG}	INFL _{TUNG}

(c) *goj зарод-wa* (ulc, agk)
‘большой зарод.ACC’

ADJ _{RUS}	INFL _{RUS}	N _{RUS}	INFL _{RUS}
ADJ _{TUNG}	—	N _{TUNG}	INFL _{TUNG}

(d) *təj значок* (gld, kmb)
‘тот значок’

ADJ _{RUS}	INFL _{RUS}	N _{RUS}	INFL _{RUS}
ADJ _{TUNG}	—	N _{TUNG}	INFL _{TUNG}



ОЖИДАНИЕ: тунгусский ADJ может появиться, если русское N оформлено тунгусским падежным аффиксом **НЕ СОВСЕМ**

3.3. Падеж: ADJ_{rus} + N_{rus}

(a) *тридцатого года* (ulc, oab)

ADJ _{RUS}	INFL _{RUS}	N _{RUS}	INFL _{RUS}
ADJ _{TUNG}	—	N _{TUNG}	INFL _{TUNG}

(b) **тридцатый года*

ADJ _{RUS}	INFL _{RUS}	N _{RUS}	INFL _{RUS}
ADJ _{TUNG}	—	N _{TUNG}	INFL _{TUNG}

(c) *железный посудка-du* (gld, i
'в железной посудке.DAT')

ADJ _{RUS}	INFL _{RUS}	N _{RUS}	INFL _{RUS}
ADJ _{TUNG}	—	N _{TUNG}	INFL _{TUNG}

(d) *колхозную картошка-səl* (ulc, oax)
'колхозную картошку.PL'

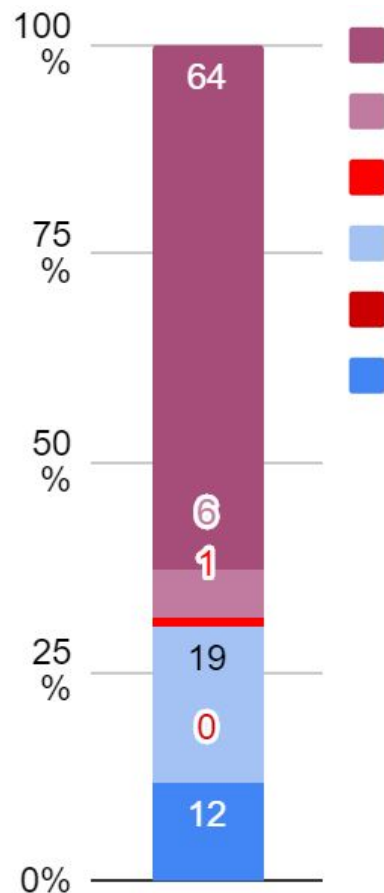
ADJ _{RUS}	INFL _{RUS}	N _{RUS}	INFL _{RUS}
ADJ _{TUNG}	—	N _{TUNG}	INFL _{TUNG}

(e) *сотовый телефон-wa* (ulc, lpd)
'сотовый телефон.ACC'

ADJ _{RUS}	INFL _{RUS}	N _{RUS}	INFL _{RUS}
ADJ _{TUNG}	—	N _{TUNG}	INFL _{TUNG}

(f) *смешанная группа bič'in* (ulc, gir)
'смешанная группа была'

ADJ _{RUS}	INFL _{RUS}	N _{RUS}	INFL _{RUS}
ADJ _{TUNG}	—	N _{TUNG}	INFL _{TUNG}



ОЖИДАНИЕ: тунгусское / русское оформление ADJ
~ тунгусское / русское оформление N **НЕ СОВСЕМ**

3.3. Падеж и род адъектива

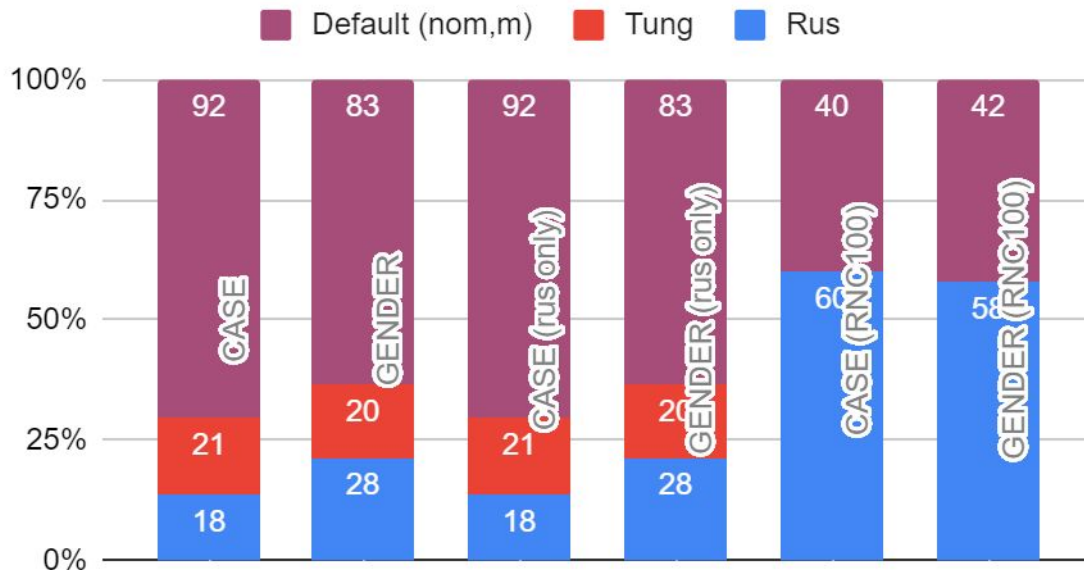
⇒ Переключение кодов

возможно в первую очередь в
“контексте нейтрализации”

т.е. контекст м.р. / им.п.

NB Доля таких контекстов
значимо выше, чем для
обычных ИГ.

(сопоставление с выборкой 100
случайных ИГ в НКРЯ с 1950 г.,
two-tailed ex. Fisher test, $p < 0.01$).



Падеж и род адъектива

3.4. Лексический состав конструкции

3.4. ADJtung + Nrus: типы адъективов

tī сестра-nī ‘эта сестра’ (ulc, aid)

dāi поле ‘большое поле’ (gld, lfb)

min тетя-ηgij ‘моя тетя’ (ulc, epv)

goj-goj стук ‘разный стук’ (ulc, oab)

ñuŋun класс ‘шестой класс’ (ulc, epv)

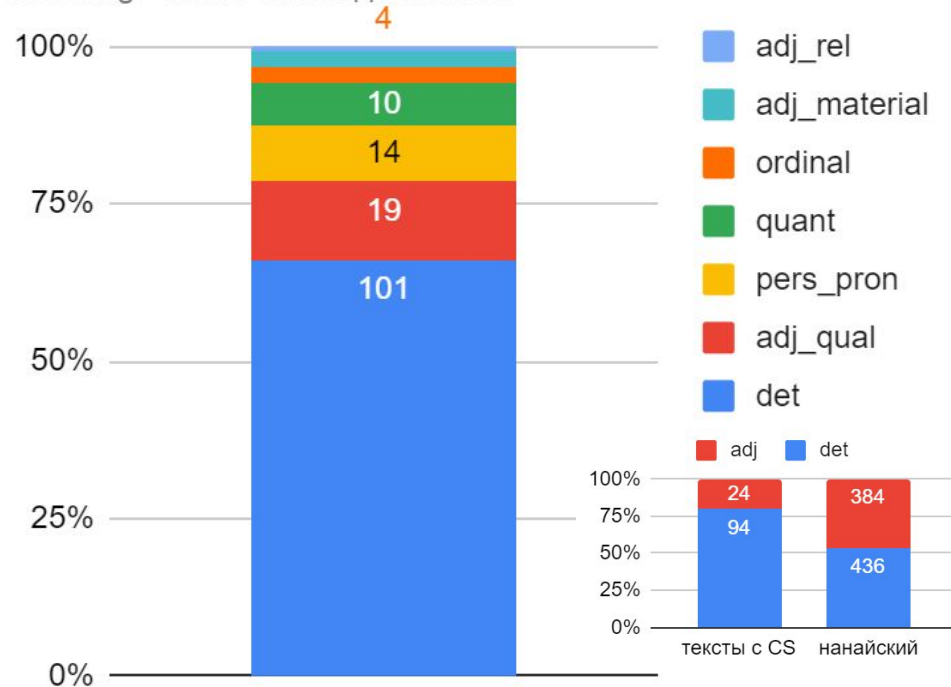
ajsema песок-wa ‘золотой песок.acc’ (ulc, lpd)

nanaj писатель ‘нанайский писатель’ (gld, sds)

⇒ Преобладают детерминанты
сравнение с выборкой нанайских текстов:
sign, two-tailed exact Fisher test.

⇒ Мало относительных (другая к-ция
в тунгусских).

ADJtung + Nrus: типы адъективов



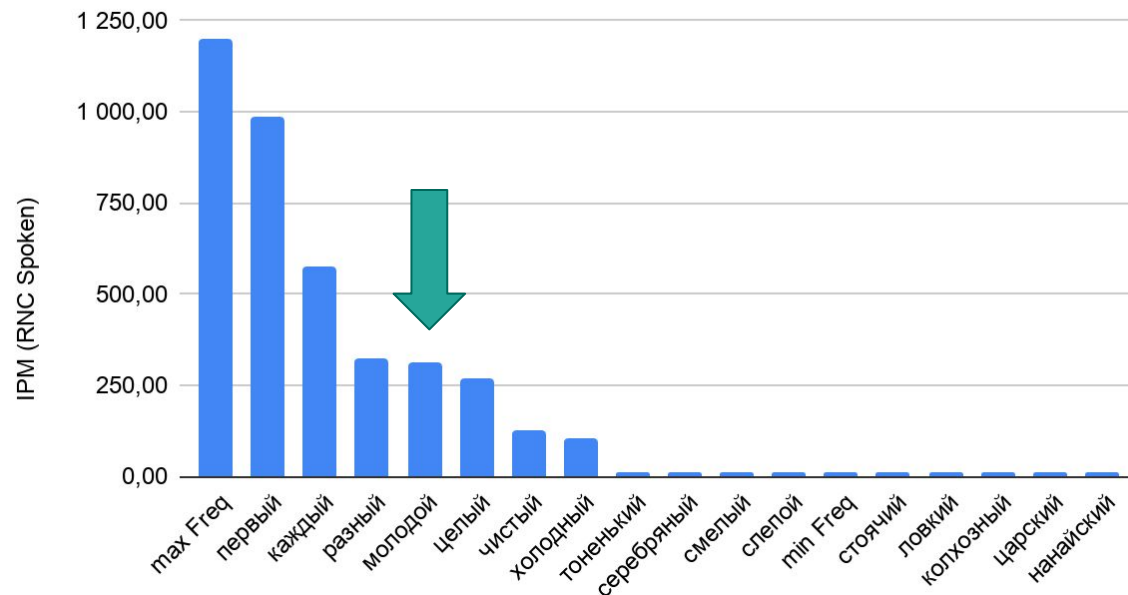
3.4. ADJrus + Ntung: частотность адъективов

ГИПОТЕЗА:

Будут вставляться
низкочастотные
прилагательные.
(Языковой сдвиг,
носители владеют
более частотной
частью лексикона).

ДА

Частотность адъективов, встретившихся в конструкции ADJ + N (по словарю Ляшевская, Шаров 2009), ipm



⇒ Собственно прилагательные действительно низкочастотные (250 ipm и ниже)₂₉

3.4. ADJrus + Nrus: коллокации

- русские реалии (14%)

колхозная поляна, краевое соревнование

- топонимы и собственные (14%)

Охотское море, Найхинский Олема, Николаевский педучилище

- порядковые + *класс / раз / год* (32%)
- прочие прилагательные (23%)

хозяйственное мыло, церковно-приходская школа

ГИПОТЕЗА: ДА

Вставка целой русской ИГ возможно, если она образует устойчивое сочетание.

Степень спаянности релевантна для CS: см. Backus 1999; Myers-Scotton 2002: 140–142), ср. “острова включенного языка” в теории Myers-Scotton.

Грамматика к-ций умеет количественно ее оценивать (collostructional analysis, ср., напр., Stefanowitsch & Gries (2003) и обзор в Wulff 2013).

3.4. ADJrus + Nrus: коллокации

Порядковые + *класс / раз / год* & Прочие

Проверка “взаимного притяжения” по НКРЯ:

LogLikelihood

(Подкорпус с 1950 г., для ADJNUM - со снятой омонимией, только в контексте ADJ + N).

Контекст: ADJNUM + _: LogLikelihood (>1.3)

год 1252.0136

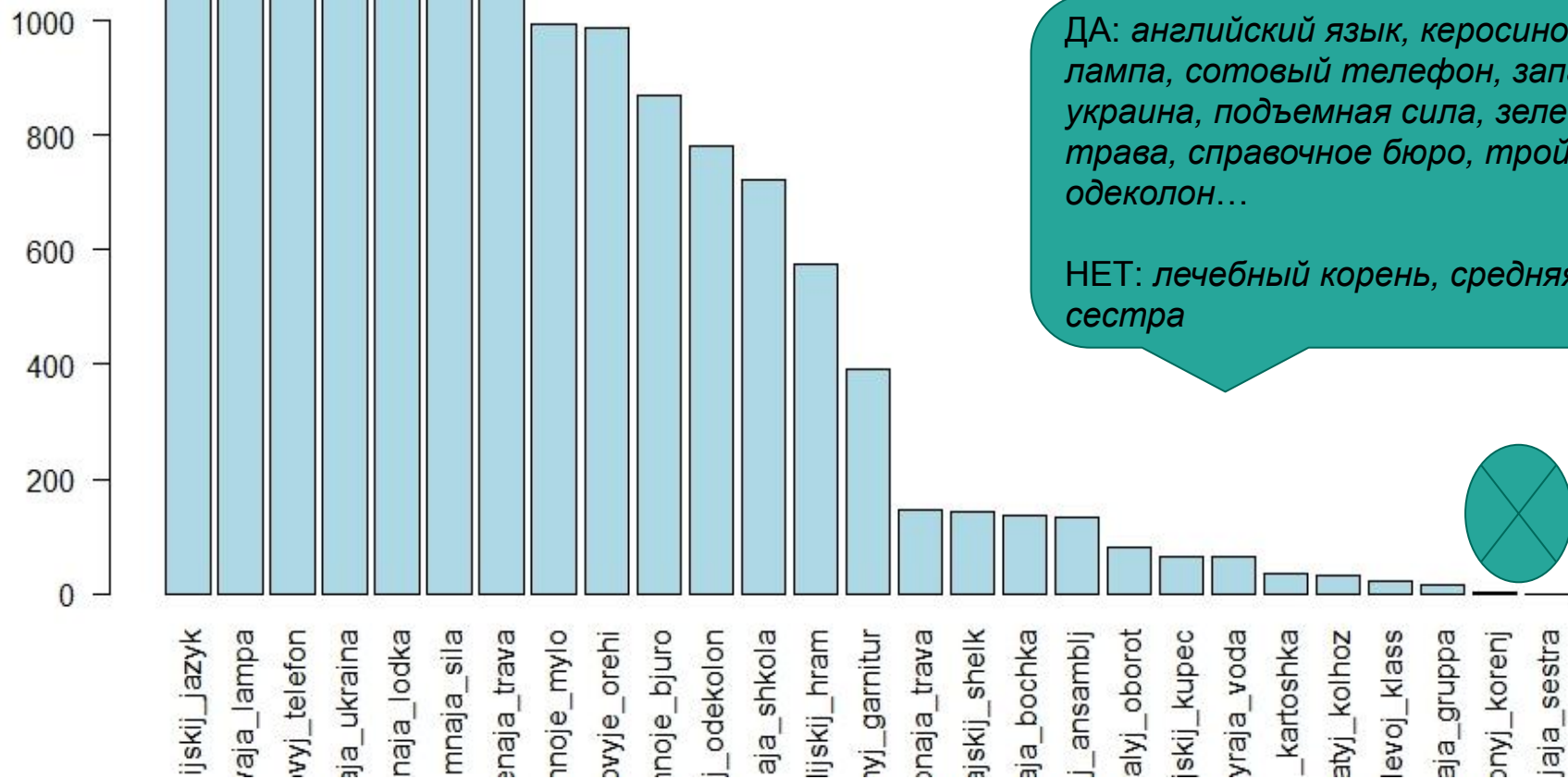
раз 1881.9065

класс 600.9422

⇒ Все отчетливо склонны к употреблению в контексте порядковых числительных.

+ Затруднения у носителей в образовании порядковых числительных и числительных большого количества.

Russian NPs: log-likelihood



4. Результаты и обобщение

4. ADJ + N с переключением кодов: результаты

- **Формальные характеристики:**

- Состав зависимого и вершины (из L1 или L2), род и падеж адъектива.
- Больше **разнообразие** типов, чем можно было бы ожидать a priori, но встречаются с разной частотой.
- Основной аттрактор CS: **дефолтная немаркированная форма**
(совпадение морфологических характеристик, ожидаемых для L1 и L2):

пятый год вставляется лучше, чем *пятом году*.

См. механизм **конгруэнтности**: возможно, единственное или основное, что нужно для описания CS (см. Sebba 2009).

4. ADJ + N с переключением кодов: результаты

- **Лексическое наполнение:**

- **Тунгусский адъектив при русском существительном:** тяготение к определенному типу (детерминанты).

Как это объяснить, не выходя за рамки выбранного подхода? Другая конструкция, а дальше? (Ср. генеративную парадигму: разные места в дереве.)

- **Русский адъектив при тунгусском существительном:** низкочастотные адъективы.

Языковой сдвиг: низкочастотное лучше помнят по-русски.

- **Полная русская ИГ:** устойчивые сочетания (~ коллокации).

Переключение кодов нарушает цельность конструкции.

4. Обсуждение: конструкционный подход к CS

ПЛЮСЫ:

- Отсутствие априорных представлений о механизме CS и предзаданных категорий, в которые нужно “уложить” факты
 - удалось заметить выделенность “дефолтных” форм:

каждый siksə ‘каждый вечер’

каждый - как в русской модели или как в тунгусской?

нет необходимости заранее это решать.

- Удобно для описания CS в ситуации языкового сдвига, с которым более “жесткие” подходы не справляются:

каждую náŋŋə ‘каждую весну’.

4. Обсуждение: конструкционный подход к CS

“ДОПОЛНИТЕЛЬНЫЕ БОНУСЫ”:

- Новый ракурс \Rightarrow новые вопросы, новые акценты, новые методы.
- Обращение к естественным эмпирическим данным.

Ср. экспериментальный метод, принятый в генеративном подходе к CS.

- Анализ в терминах частотности:
 - в частности, обращение к количественным мерам спаянности сочетаний;
 - в частности, обращение к количественным мерам “освоенности” слова.
- NB Но все это на самом деле можно делать, работая в любом фреймворке.

4. Обсуждение: конструкционный подход к CS

МИНУСЫ:

- Те же, что и плюсы.
- Слишком гибкий подход: не всегда позволяет делать предсказания и давать глубокие объяснения.
- Переключение кодов - возможность увидеть морфосинтаксическую структуру “обнаженной”:

например, границы L1 и L2 при CS соответствуют (?) морфосинтаксическим границам, способность к переключению коррелирует со структурным классом и т. п.

- При конструкционном, эмпирическом подходе эта возможность, кажется, не используется в полной мере.