

# ВАРИАТИВНОСТЬ В СОГЛАСОВАНИИ ПО РОДУ В РУССКОЙ РЕЧИ НОСИТЕЛЕЙ ТУНГУССКИХ ЯЗЫКОВ: КОЛИЧЕСТВЕННОЕ ИССЛЕДОВАНИЕ

Н. М. Стойнова, ИРЯ РАН & ИЯз РАН, Москва  
[stoynova@yandex.ru](mailto:stoynova@yandex.ru)

Славянские языки в условиях современных вызовов - VII.  
Секция Языковое контактирование. Томск, 18-19.10.2021 г.

Исследование подготовлено в рамках проекта РФФ No. 17-18-01649.

# Введение

→ Согласование по роду в устной русской речи носителей южно-тунгусских языков – нанайского и ульчского:

(1) Тебе приехали **богатый дед** \ сватать - СТАНДАРТНОЕ: контролер - m, **мишень** - m

(2) **Богааатый** \ **старый** \ бабушка - НЕСТАНДАРТНОЕ: контролер - ожидается f, **мишень** - m

→ В нанайском и ульчском языках категория рода отсутствует.

→ В целом очень характерно для русской речи носителей языков России.

Для носителей тунгусских и самодийских языков - см. наше пилотное исследование  
[Khomchenkova et al. 2018]

# Введение: нанайский и ульчский языки

- южно-тунгусские, Хабаровский край
- не передаются детям, все носители старшего поколения  
см. Герасимова 2002; Калинина, Оскольская 2016; Сумбатова, Гусев 2016
- все свободно владеют и активно пользуются русским языком, близким к литературному стандарту
- однако в речи самых пожилых носителей есть черты, связанные с копированием тунгусского образца и/или нестандартным усвоением русской системы

см. Stoynova 2019

# Введение: нестандартное согласование по роду

→ Нестандартные явления в системе рода также отмечались или описывались на русском материале для:

- пиджинов (ср. [Helimsky 2010, Stern 2012: 351 ff.]).
- диалектов (ср. гипотезы о контактной природе нестандартной системы рода в диалектах: [Азарова 1986; Мызникова 2017]).
- русского языка как иностранного (ср. [Taraban, Kempe 1999; Romanova, Gor 2017]).
- детской речи (ср. [Rodina, Westergaard 2012]).
- русского как языка семейного наследия (Heritage Russian) (ср. [Polinsky 2008]).
- речи афатиков (ср. [Akhutina et al. 2001]).
- ...

# Введение

- От чего зависит выбор стандартного vs. нестандартного согласования по роду?
- С какими социолингвистическими характеристиками коррелирует?
- Какими морфосинтаксическими и семантическими параметрами определяется?
- В частности: ориентируется ли двуязычный носитель при выборе рода на семантику (пол) или на морфологию (тип склонения существительного)?
- Количественное исследование на корпусном материале

**Данные**



# Корпус

Корпус русской речи EN | RU | ? ”

## Корпус русской речи

носителей языков Севера Сибири и Дальнего Востока



Запрос

● Слово №1

Слово:

Лемма:

Грамм. тэги:

Синт. тэги:

Морф. тэги:

Лекс. тэги:

Фон. тэги:

Пдп тэги:

Неконт. тэги:



- <http://web-corpora.net/ruscontact/corpus.html>
- в рамках проекта “Динамика языковых контактов в циркумполярном регионе” (О. В. Ханина)
- совм. с И. А. Хомченковой и П. С. Плешак
- устные спонтанные тексты
- автоматическая морфологическая разметка
- ручная разметка “контактных явлений”



# Выборка

→ Размеченные тексты от носителей южно-тунгусских языков: 104 562 сл.-употр.

→ Автоматический поиск: VERB,sing,past,indc и (ADJF|ADJS), sing: **14 693** примеров

→ После ручной фильтрации: **7 697** примеров

- выброшен “мусор” (село - глаг.)
- пошел и увидел, самая красивая → объединены
- дополнительная фильтрация: -conv (на самом деле), -subst (военный) и нек.др.

→ Основная выборка: **1748 примеров** (из них **458** с нестандарт. согл., 26%)

- только от носителей, в речи которых встретилось нестандартное согласование (15 из 53)
- только с выраженным контролером-существительным
- дополнительно удалены некоторые неоднозначные контексты

(1) Уже она **беременный**\ - НЕТ

(2) **Старый**\ о стала - НЕТ



# Параметры



# Параметры

- Социолингвистические
  - год рождения
  - уровень образования
  - “индекс грамматической интерференции”
- Морфосинтаксические и семантические
  - мишень согласования: морфосинтаксис (ч.р., синт. позиция)
  - мишень согласования: линейная позиция
  - контролер согласования: род (ожидаемый)
  - контролер согласования: семантика (шкала одушевленности)
  - контролер согласования: морфологический тип (тип склонения)

# Социолингвистические параметры

НОСИТЕЛИ		с нестандарт.согл.	всего	% с нестандарт.
образование	начальное	9	9	<b>100,00%</b>
	среднее	5	20	25,00%
	высшее	1	9	11,11%
год рождения	<b>1917-1929</b>	5	5	<b>100,00%</b>
	<b>1930-1939</b>	8	17	<b>47,06%</b>
	1940-1957	2	16	12,50%
пол	f	14	35	40,00%
	m	1	3	33,33%
<b>ВСЕГО</b>		15	38	<b>39,47%</b>

# Социолингвистические параметры

→ “Индекс грамматической интерференции”

- для каждого носителя - вычислен по разметке корпуса:

$(N \text{ морфосинтаксических помет - agr\_adj\_gender\&arg\_verb\_gender}) * 1\ 000$

-----  
N словоупотреблений

(1) У дедушки **три\ жена** было - num

(2) Спроси **старикам** сааамым старшим\ **спроси\** - gov

(3) И стала **разбогатеть\** - asp

# Контролер согласования: шкала одушевленности

→ Русский род коррелирует с полом - ЧАСТИЧНО

- Неодушевленные: род = '∅' (стул, табуретка)
- Животные: род ??? 'пол' (бык, корова)
- Лица: род = 'пол' (муж, жена)

(1) **Такой маленький** \ эта деревня была

(2) Девочка **маленький** пошла за водой/

## ОЖИДАНИЯ

выше % рассогласований > ниже % рассогласований  
INANIMATE > ANIMAL > HUMAN > (1,2 PERSON)

# Контролер согласования: морфологический тип

→ Русский род коррелирует с типом склонения - ЧАСТИЧНО

- “морфологически прозрачные” - род (почти) однозначно предсказывается по типу склонения (~ форме nom sg): **стол**, **стена**
- “морфологически непрозрачные” - род не предсказывается или вступает в противоречие с типом склонения (~ формой nom sg)

(1) Это **вечный**/ память\ - ср. старый лапоть

(2) Чтобы **одна** семечко\ **была** - ср. одна веточка

(3) **Лечила**/ папка мене - ср. лечила мамка

(4) Всё\ пропивает, только **один** пенсий\ **остал** - ср. пенсия ж.р. > пенсий м.р.

## ОЖИДАНИЯ

выше % рассогласований > ниже % рассогласований

**NON-TRANSPARENT > TRANSPARENT**

# Контролер согласования: (ожидаемый) род

→ Встретились почти все возможные типы рассогласований:

самый дорогой лекарство N→M

такая зеркало N→F

первый ложка F→M

длинная дом M→F

??? - M, F → N

(1) За **моё** рыбой - речевой сбой?

→ ОЖИДАНИЯ

выше % рассогласований > ниже % рассогласований

**N > F > M**

## Контролер согласования: род

контролер	согласуется по	N	%
N	→N	82	53,95%
	<b>→M</b>	60	<b>39,47%</b>
	→F	10	6,58%
F	→F	440	62,50%
	<b>→M</b>	262	<b>37,22%</b>
	→N	2	0,28%
M	→M	767	85,99%
	<b>→F</b>	123	<b>13,79%</b>
	→N	2	0,22%



# Мишень согласования: морфосинтаксис

→ согласование с прилагательным (ед.ч.): в атрибутивной позиции - в номинативе или в косвенном падеже; в предикативной позиции

(1a) **Эта богатая** дед-то уже умер/

(1b) В **одном** деревне это... жил старик\

(1c) Папа тоже **старая**\

→ согласование с глаголом (прош.вр. ед.ч.)

(2) Папа тоже **рыбачила**\

→[ выбор анафорического местоимения (ед.ч.)]: **он, который**, ... - не учитывалось

(3) Какая-то шапка\ у **него** б́ыла, у папина маме

# Мишень согласования: линейная позиция

→ Мишень согласования (прилагательное, глагол) может быть справа или слева от контролера согласования (существительного):

(1) **Наша** дедушка/ папкина ... отец\ - ПРЕПОЗИЦИЯ

(2) А дедушка **наша** пришел/ - ПОСТПОЗИЦИЯ

## ОЖИДАНИЯ (?)

выше % рассогласований > ниже % рассогласований

**PREPOSITION > POSTPOSITION**

? “Легче” согласовать прилагательное / глагол, если контролер согласования уже введен?

# Статистический анализ



# Описательная статистика

		станд. согл.	нестанд.	% нестандарт.
образование	начальное	954	450	<b>32,05%</b>
	прочее	336	8	2,33%
мишень: тип	<b>глагол</b>	295	96	<b>24,55%</b>
	<b>прил_пред.</b>	78	24	<b>23,53%</b>
	<b>прил_атр.: им.п.</b>	624	277	<b>30,74%</b>
	прил_атр.: косв.п.	293	61	17,23%
мишень: позиция	препозиция	468	155	24,88%
	постпозиция	822	303	26,93%
контролер: род	<b>n</b>	82	70	<b>46,05%</b>
	<b>f</b>	441	263	<b>37,36%</b>
	m	767	125	14,01%
контролер: семантика	<b>неодуш.</b>	687	290	<b>29,68%</b>
	<b>животные</b>	36	23	<b>38,98%</b>
контролер: тип	лица	567	145	20,37%
склонения	<b>непрозрачное</b>	365	248	<b>40,46%</b>
	прозрачное	925	210	18,50%

# Многофакторный анализ

→ Видимо, на выбор согласования по роду влияет сразу несколько факторов

- возможно, по-разному для разных носителей
- возможно, с разной силой
- факторы логически взаимосвязаны

3 скл. - только *m* и *f*; *n* - не лицо, глагол чаще в постпозиции, а прилагательное в препозиции...

- действие одного можно принять за действие другого

⇒ Многофакторный анализ

- биномиальная логистическая регрессия
- модель со смешанными эффектами: функция `glmer()` в R (случайный эффект - носитель)
- см. Baayen 2008: 214–226; Gries 2013: 293–310; Levshina 2015: 253–276

# Логистическая регрессия

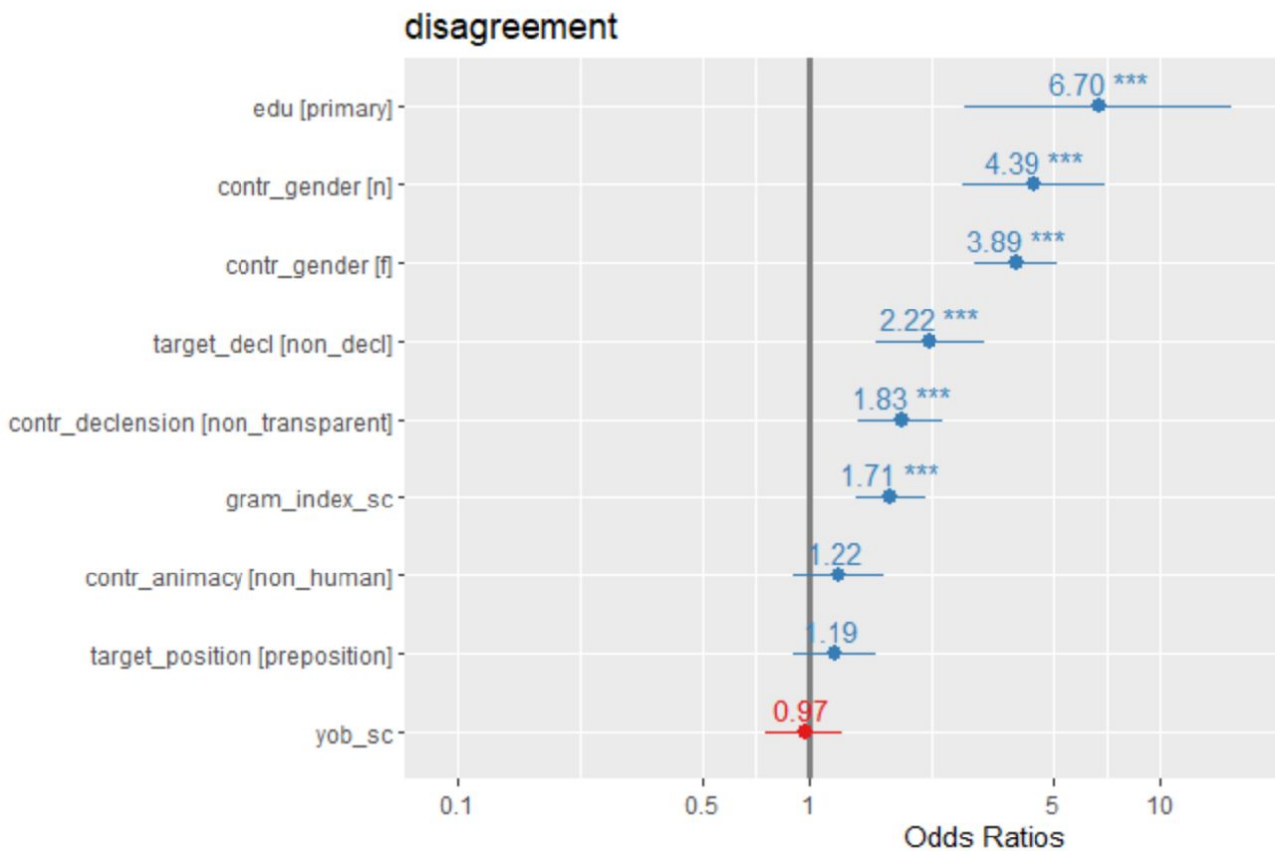
→ Значимые факторы (упорядочены по силе):

- уровень образования [начальное > прочее]
- род существительного-контролера [N > F > M]
- морфосинтаксический тип мишени [прочее > ADJ в косвенном падеже]
- склонение существительного-контролера [морфологически непрозрачные > прозрачные]
- уровень грамматической интерференции в речи носителя
- + положение существительного-контролера на шкале одушевленности [не-лицо > лицо] - статистически не значимо, но нельзя удалить из модели

→ Незначимые факторы:

- линейная позиция мишени согласования
- год рождения (малый разброс в выборке)

# Логистическая регрессия



# Результаты





# Двуязычные носители с нестандартным согласованием по роду

→ Нестандартное согласование - у очень узкой группы носителей:

- примерно до 1940-х гг. р.
- в основном с начальным образованием (иногда ≈ без образования)
- русская речь которых в целом очень нестандартна
- все это носители, для которых нанайский/ульчский первый по времени усвоения

→ Следующие поколения носителей (в речи которых согласование стандартное):

- в той или иной степени сталкиваются с русским еще в дошкольном возрасте
- охвачены всеобщим школьным образованием (на русском языке, как минимум среднее)

⇒ **Нестандартное согласование у тех носителей, которые:**

- **осваивают русский поздно: поздний, последовательный билингвизм**
- **осваивают его (преимущественно) спонтанно: естественный билингвизм**

# Лингвистические факторы, располагающие к нестандартному согласованию

→ ? Семантика (пол) или морфология (тип склонения)?

- **морфология - ДА:** для морфологически прозрачных существительных нестандартного согласования по роду значимо меньше
- **семантика - ДА/НЕТ: неожиданно слабый ориентир** при выборе рода
- для не-лиц (неохарактеризованных по полу) нестандартного согласования больше, но это самый слабый, статистически незначимый фактор

(1) Жена \ умер

(2) А это \ == старший брат / **вышла** / **забрала** / ружье-то

# Лингвистические факторы, располагающие к нестандартному согласованию

→ Морфосинтаксические характеристики мишени согласования:

- не важна линейная позиция
- нет заметной разницы между частеречными классами (с разной согласовательной морфологией): ADJ = VERB
- нет заметной разницы между синтаксическими позициями: ATTR = PRED
- ⇒ ключевая разница - между **прилагательными в косвенных падежах** и всем остальным - почему?

# Типы нестандартного согласования

→ Особый статус **мужского рода**

- “наиболее стабилен”: существительные мр. реже обнаруживают нестандартное согласование
- наиболее “экспансивен”: нестандартное согласование чаще по м.р., очень редко по ср.р.

→ При этом:

- нет эффекта дефолтного согласования по м.р.: хотя и с разной частотой, но встречаются (почти) все возможные типы нестандартного согласования

# Другая перспектива (+)

**Перспектива 1** (здесь): речь билингов как промежуточная система (на фоне стандартной)

- пиджин → наши данные = постпиджинный континуум → стандартный русский
- выбор между “нестандартным” и “стандартным” согласованием - в каких условиях двуязычный носитель ведет себя “как монолингв”, а в каких нет?

**Перспектива 2:** речь билингов как отдельная самодостаточная система

- выбор между *m*, *f*, *n* - чем регулируется в этой системе?

→ Пилотное исследование: выбор между *m* и *f* - только для лиц

(Монолингвы: пол)

→ **Наши данные:** к выбору согласования по ж.р. значимо располагает:

- пол [ж]
- морфология [*nom sg* на гласный (мама), *nom sg* на мягкий согласный (мать)]
- синтаксис [мишень согласования в предикативной позиции]

# Дополнительные слайды



# Выборка (+)

→ Выброшены: контексты с **это**, **эта** и **всё** (неоднозначность)

(1) **Всё** детей я Куруне\ родила = всех детей или всё родила?

(2) Вытаскивает/ **это** грудник = этот грудник или это - частица?

→ Выброшены: локативные и экзистенциальные контексты с **быть**  
(неоднозначность)

(3) У нас амбар/ **был[ъ]** = была ж.р. или было дефолтный ср.р. (как было два амбара)

# Контролер согласования: морфологический тип (+)

→ Этот механизм обсуждался как релевантный для нестандартного согласования:

- в детской речи (Гвоздев 1961; Цейтлин 2006; Rodina, Westergaard 2012)
- в heritage Russian (Polinsky 2008)



# Контролер согласования: морфологический тип (+)

→ Отнесены к морфологически непрозрачным

- а) лица мужского пола на -а (ожид. ж.р.): **дядя, папка, ...**
- б) не-лица на мягкий согл., не j (2 и 3 скл.): **дверь, зверь...**
- в) неодушевленные на безударный гласный: **место, паста, время, + нестандартные: внука, карандаша...**
- г) лица и не-лица несклоняемые - на нестандартный исход: **дзюли, кальдами**
- д) с преобладающим мн.ч.: **сапог(и)...** (+ **дети, люди**) - pl tantum исключены
- е) адъективного скл-я (часто нестандартная форма самого сущ.): **столовая, сохатый**

NB примеры типа **большой щук** - disagreement: yes, contr\_declension: non\_transparent, comment: nonstand

# Контролер согласования: (ожидаемый) род (+)

→ Ср. тенденции к унификации или редукции системы рода в некоторых разновидностях русского языка:

- южнорусские диалекты: m, f, n → m, f, (m/f) (см. обзор в Касаткин 2005: 116-117)
- говорка (Таймырский пиджин): m, f, n → m (см. Stern 2012: 351 ff.)

# Мишень согласования: морфосинтаксис (+)

ОЖИДАНИЯ (?)

выше % рассогласований > ниже % рассогласований  
**VERB > ADJ\_PRED > ADJ\_ATTR\_NOM > ADJ\_ATTR\_DECL**

→ Типологическая иерархия Корбетта (Corbett 1991: 225–260)

**PERS PRON > RELAT PRON > PRED > ATTR**

чем левее, тем выше вероятность “семантического” согласования

Ср. в (литературном) русском:

врач ... она > врач, которая > ?врач пришла > ??наша врач > \*нашей врачу

	Estimate	Std. Error	z value	Pr(> z )	sign.
(Intercept)	-4.99893	0.47877	-10.441	<2e-16	***
<b>edu [primary]</b>	1.90194	0.44711	4.254	2.10e-05	***
yob_sc	-0.02757	0.12845	-0.215	0.830	
<b>gram_index_sc</b>	0.53757	0.11697	4.596	4.32e-06	***
<b>target_morphosynt [non_decl]</b>	0.79757	0.18345	4.348	1.38e-05	***
target_position[preposition]	0.17152	0.13995	1.226	0.220	
<b>contr_gender [n]</b>	1.35815	0.13725	9.896	<2e-16	***
<b>contr_gender [f]</b>	1.47880	0.23736	6.230	4.66e-10	***
contr_animacy [non_human]	0.19973	0.15155	1.318	0.188	
<b>contr_declension [non_transparent]</b>	0.60580	0.14304	4.235	2.28e-05	***

# Сопоставление с пилотным исследованием (+)

<b>Khomchenkova et al. 2018</b>	<b>это исследование</b>	
HUMAN > NON-HUMAN	HUMAN >= NON-HUMAN	НЕ СОВСЕМ: самый слабый фактор
NON-TRANSPARENT=TRANSPARENT	NON-TRANSPARENT>TRANSPARENT	НЕТ: эффект есть
N > F > M	N > F >> M	ДА
ADJ > VERB > PRON	прочее > ADJ_DECL	НЕТ: прилагательные ведут себя по-разному