

Лингвистика и гендерные стереотипы

Анастасия Алексеевна Герасимова¹,
Екатерина Анатольевна Лютикова²

^{1,2}МГУ имени М.В. Ломоносова

¹anastasiagerasimova432@gmail.com, ²lyutikova2008@gmail.com

Посвящение

В наше время никого, наверное, не удивляет, что среди лингвистов много женщин, а в некоторых областях лингвистики (причем далеко не самых простых!) их явное большинство. Студенты ОТиПЛа если и ощущают гендерный дисбаланс, то исключительно в отношении нехватки молодых людей. Тем удивительнее звучит сегодня история поступления на отделение Ольги Крутиковой (см. стр. 19–20 наст. сборника — *прим. ред.*), чья увлеченность и настойчивость привели к снятию несуразных гендерных ограничений на новую специальность уже с первого года ее существования в МГУ.

Звегинцев и Успенский создали ОТиПЛ. Ольга Фёдоровна открыла его для всех.

1. Введение

Гендерная смещенность (*gender bias*) представляет собой широко распространенный феномен, который заключается в ненамеренном и зачастую неосознаваемом выборе одного гендера на фоне другого (других). Релевантность проявлений гендерной смещенности для теоретических и прикладных лингвистических работ стала активно обсуждаться в последние десятилетия. В прикладных областях мы обнаруживаем влияние гендерной смещенности на коллекции обучающих текстов и, как следствие, на модели автоматической обработки естественного языка (см. обзор в [Sun et al. 2019]). Вероятно, многие сталкивались с особенностями машинного перевода с языков без родовых различий на языки с грамматикализованным родом или именным классом. Например, Google Translate уже много лет предлагает в качестве переводов грамматически нейтральных английских предложений контра-

стирующие в отношении грамматического рода переводы на русский язык (1); в не так давно появившемся татарско-русском автоматическом переводе результаты аналогичные (2). Широкомасштабные многоязычные исследования в области машинного перевода подтверждают, что и промышленные, и академические разработки характеризуются ошибками, связанными с гендерной смещенностью [Stanovsky et al. 2019].

(1) a. *I worked.*

‘Я работал.’ (перевод Google.Translate)

b. *I danced.*

‘Я танцевала.’ (перевод Google.Translate)

(2) a. *Мин укытучы бул-ып эшлә-де-м.*

я учитель быть-CVB работать-PST-1SG

‘Я работала учителем.’ (перевод Google.Translate)

b. *Мин инженер бул-ып эшлә-де-м.*

я инженер быть-CVB работать-PST-1SG

‘Я работал инженером.’ (перевод Google.Translate)

Гендерная смещенность проявляется и в языковых данных, которые используются в теоретической лингвистике, в частности, в специально конструируемых при грамматическом моделировании примерах. Языковой материал, приводимый в иллюстративных целях в учебниках и статьях, зачастую характеризуется гендерной смещенностью, что подтверждается многочисленными исследованиями, сообщающими о преобладании в языковых примерах персонажей-мужчин и корреляциях гендера и ролевой характеристики персонажа [Macaulay, Brice 1994, 1997; Pabst et al. 2018; Richey, Burnett 2020; Cépeda et al. 2021; Kotek et al. 2021].

Очевидно, что гендерная смещенность лингвистических данных обусловлена в конечном итоге статистическими закономерностями в представленности мужчин и женщин в различных профессиональных и социальных контекстах. Типичность определенных сценариев кодируется в частотно-зависимых параметрах языковых моделей и осознанно или неосознанно используется лингвистами-теоретиками для повышения уровня правдоподобия (англ. *plausibility*) языковых примеров. Поскольку основная цель использования примеров состоит в том, чтобы проиллюстрировать контраст приемлемости, исследователи могут отдавать предпочтение более типичному и частому сценарию, чтобы этот контраст

был заметнее. Степень правдоподобия может зависеть в том числе и от гендера участника ситуации, описываемой в предложении, а именно, от того, соответствует ли он культурно обусловленному **гендерному стереотипу**.

В этой статье мы исследуем влияние гендерных стереотипов на приемлемость, выявляемую в экспериментально-синтаксических исследованиях. Используя сценарии, отобранные на основании предшествующих исследований гендерных стереотипов, мы проверяем, влияет ли изменение гендера протагониста на абсолютную приемлемость предложения и на возможность установления контраста между парами экспериментальных условий.

Статья организована следующим образом. В разделе 2 мы подробнее расскажем о том, как гендерная смещенность проявляется в лингвистических примерах и на какие гендерные стереотипы она опирается. В разделе 3 мы описываем языковой феномен, на основе которого организовано экспериментальное исследование, представленное в разделе 4. Раздел 5 подводит итоги исследования.

2. Гендерные стереотипы в лингвистических примерах

Изучению гендерных стереотипов в лингвистических примерах посвящен целый ряд работ. М. Маколей и К. Брайс впервые представили данную проблему на материале учебников по синтаксису, которые были опубликованы с 1969 по 1994 гг. [Macauley, Brice 1994, 1997]. В двух недавних статьях [Cépeđa et al. 2021; Kotek et al. 2021] проводится количественный анализ сконструированных примеров из шести англоязычных учебников по синтаксису, опубликованных в период с 2005 по 2017 гг., и из статей, опубликованных с 1997 по 2018 гг. В работе С. Ричи и Х. Бёрнет количественное исследование сконструированных примеров воспроизводится на материале лингвистических статей на французском языке, которые вышли в свет в 1969–1971 и 2008–2017 гг. [Richy, Burnett 2020]. Несмотря на то, что названные исследования проводились в разные десятилетия и на материале разных языков, все они подтверждают наличие гендерной смещенности и языковых стереотипов в лингвистических примерах.

В иллюстративном материале отмечается устойчивая тенденция к использованию одушевленных именных групп мужского (грамматического) рода и к предпочтению их одушевленным именным группам женского рода. Одушевленные аргументы мужского рода с большей вероятностью используются в качестве подлежащих по

сравнению с другими грамматическими функциями, одушевленные аргументы женского рода чаще встречаются в неподлежащей функции. Участники разного гендера оказываются неравномерно распределены в отношении их семантических ролей: преобладает использование одушевленных аргументов женского рода в роли пациенса и реципиента, в то время как одушевленные аргументы мужского рода скорее используются в роли агенса и экспериенцера. В отношении лексического выбора фиксируется закономерность, согласно которой мужчины упоминаются в связи с интеллектуальной, перцептивной и насильственной деятельностью. Женщины, напротив, упоминаются в контексте физической и эмоциональной активности. Количественные исследования лингвистических примеров позволяют выявить стереотипные представления о гендерах: мужчины часто имеют дело с автомобилями, насилием, спортом и владением имуществом, женщины скорее упоминаются в контексте работы по дому и романтических отношений. Анализ тональности в [Kotek et al. 2021] также показывает корреляцию между гендером протагониста предложения и эмоциями, которые описываются в предложении: в лингвистических примерах, в которых описываются отрицательные эмоции, с большей вероятностью речь идет о мужчинах, в предложениях, описывающих положительные эмоции — о женщинах.

Как мы уже отмечали выше, причиной подобной смещенности может быть неосознанная склонность автора представить контраст приемлемости предложений на наиболее правдоподобном лексическом материале. Действительно, правдоподобие, определяемое как отсутствие конфликта со знаниями об устройстве окружающего мира, прототипичность или высокая вероятность определенной аргументной структуры [Adger 2003], может влиять на субъективную оценку приемлемости предложения. В исследовании [Schütze 2020] влияние фактора (не)правдоподобия на приемлемость было продемонстрировано в экспериментальных условиях на отдельном примере; в работах [Levelt 1977] и [Häussler, Juzek 2019] показано, что логическая (не)последовательность и высокая образность vs. конкретность также могут повышать или понижать приемлемость предложения вне зависимости от его грамматического статуса. Кроме того, отмечается бóльшая сложность неправдоподобных предложений для анализа, что выражается в задержке в чтении с саморегулирующей скорости при несоответствии по правдоподобию (англ. *plausibility mismatch*), когда

аргумент нарушает селективные ограничения предиката [Tanenhaus et al. 1989; Stowe et al. 1991; Boland et al. 1995; Traxler, Pickering 1996]. В связи с этим в литературе, посвященной методологии проведения экспериментальных исследований, обычно даются рекомендации по составлению лингвистических стимулов таким образом, чтобы те были максимально правдоподобными [Schütze 1996; Schütze, Sprouse 2014; Häußler, Juzek 2019; Huang, Ferreira 2020]. Несмотря на то, что в ряде работ, которые упомянуты выше, был продемонстрирован эффект смещения оценок вследствие (не)правдоподобия, отдельных систематических исследований этого эффекта с применением стандартов экспериментального синтаксиса до сих пор не проводилось. Впрочем, имеющиеся пилотные эксперименты, проводимые с целью проверки качества стимульного материала, скорее подтверждают наличие соответствующего эффекта (напр., [Kizach et al. 2013]).

В этой статье мы задаемся вопросом о том, в какой степени правдоподобие требует соответствия гендерному прототипу, и сосредоточиваемся на исследовании влияния гендерного прототипа и его нарушения на восприятие респондентами грамматического контраста в экспериментальных условиях. Нас будет интересовать влияние гендерных прототипов на оценки приемлемости предложений разной степени грамматичности и на возможность выявления контраста по грамматичности между условиями. Обратимся теперь к явлению, с помощью которого мы задаем в эксперименте грамматический контраст.

3. Языковой феномен

Критерии выбора языкового феномена для экспериментального исследования были сформулированы нами следующим образом. Во-первых, он должен легко комбинироваться с различными типами ситуаций, включающих протагонистов разного гендера, и, соответственно, с глаголами различных семантических классов. На основе анализа данных в [Cépeda et al. 2021; Kotek et al. 2021] мы выделили следующие гендерные стереотипы: агенс предикатов агрессии (м), агенс предикатов профессиональной деятельности (м), экспериенцер стенических предикатов (м), пациенс предикатов агрессии (ж), агенс предикатов ведения домашнего хозяйства (ж), экспериенцер астенических предикатов (ж). Эти типы предикатов должны образовывать с исследуемым феноменом естественные конструкции.

Во-вторых, языковой феномен должен быть полярным, то есть подразумевать четкий контраст приемлемости. Обоснование этого критерия состоит в том, что можно предположить различные сценарии влияния гендерных стереотипов на оценки приемлемости полностью грамматичных и не вполне грамматичных стимулов. С одной стороны, возможно одинаковое снижение оценок за несоответствие гендерному стереотипу для разных по грамматичности стимулов (эффект аддитивности). С другой стороны, возможно, что нарушение гендерных стереотипов по-разному отразится на оценке приемлемости разных по грамматичности стимулов. Например, можно ожидать, что действие гендерных стереотипов будет заметно лишь при оценивании полностью грамматичных стимулов и незначимо в случае неграмматичных стимулов (на вероятность такого сценария указывают, в частности, результаты работ [Kizach et al. 2013; Sprouse 2008]). Еще одна логическая возможность состоит в том, что респонденты окажутся чувствительны к нарушению стереотипов, только если предлагаемый для оценки стимул грамматически безупречен; в таком случае мы ожидаем эффекта супераддитивности во взаимодействии экспериментальных факторов. Последняя возможность служит обоснованием третьего критерия выбора языкового феномена: необходимость разграничивать низкие оценки требует того, чтобы грамматически менее корректный вариант тем не менее не вызывал абсолютной неприемлемости.

Языковой феномен, на котором мы остановили свой выбор — полипредикативные конструкции с придаточным времени, вводным союзом *когда*. Если придаточное предложение находится в препозиции, в главном возможно появление местоименного коррелята *то*, ср. (3а–б). Союз *когда* и коррелят *то* образуют, по разным описаниям, сложный союз, коррелятивную пару или составной коннектор [Русская грамматика 1980; Inkova 2014; Pekelis 2016; Пекелис 2018а, 2018б; Кобозева 2019].

- (3) а. *Когда Ада ушла, дед заплакал.*
[Саша Миндориани. Дедушкины песни // «Волга», 2014]
- б. *Когда бабушка забила в конец гвоздь, то палочка треснула.*
[А.А. Темников. Поведеление родинки во сне (2006) // «Волга», 2011]

Коррелятивная конструкция в (3б) и простая подчинительная конструкция в (3а) имеют не полностью совпадающую дистри-

буцию. Коррелятивная конструкция часто используется в равносубъектных конфигурациях (4), где она лицензирует нулевое подлежащее (*pro*) главной клаузы [Пекелис 2018b; Pekelis 2016; Madariaga 2018].

- (4) *Когда Андрей_i вышел из операционной, то pro_i увидел нескольких иностранцев.* [Лев Дурнов. Жизнь врача. Записки обыкновенного человека (2001)]

Напротив, в разносубъектных конфигурациях коррелятивная конструкция подразумевает контраст ситуаций, описываемых в главной и зависимой части предложения [Пекелис 2015; Podlesskaya 1997] и реинтерпретацию временного отношения как условного или причинного, ср. (5). Если же такая реинтерпретация прагматически неестественна, то равносубъектная коррелятивная конструкция оказывается сомнительной (6). Подчеркнем, что в таких конфигурациях простая подчинительная конструкция (без коррелята *то*) полностью приемлема.

- (5) а. *Когда Гердт смеялся, (то) в мире наступала гармония.*
[З.Е. Гердт. Рыцарь совести (2010)]
- б. *Когда милиционеры остановили машину, (то) находившиеся в салоне четверо боевиков открыли стрельбу.*
[Хроника террора (2003) // «Спецназ России», 15.05.2003]
- (6) *Когда стемнело, (*?то) к даче подкатили два грузовика.*
[К.Л. Селиверстов. 12 с половиной, или Моя жизнь в чистом искусстве // «Волга», 2011]

Мы используем эти свойства простой и коррелятивной разносубъектной временной конструкции для реализации грамматического фактора в экспериментальном дизайне. Уровнями фактора служат наличие или отсутствие коррелята в разносубъектной временной конфигурации. Чтобы исключить условную интерпретацию (которая способствовала бы повышению оценок коррелятивной конструкции), мы используем предельные глаголы совершенного вида в форме прошедшего времени в обеих клаузах. Мы ожидаем, что в этих условиях стимулы с коррелятом должны получать более низкие оценки, чем стимулы без коррелята. В то же время мы ожидаем, что стимулы с коррелятом не будут оцениваться как полностью неприемлемые. Основанием для этого ожидания служит тот факт, что коррелятивная разносубъектная кон-

струкция с временным значением все же засвидетельствована в корпусе, хотя и встречается существенно реже (7).

- (7) *Когда Зоя с Зиной опомнились, то Шмеля и след уже простыл.* [Виктор Кологрив. Медовый луг // «Мурзилка», 2002]

4. Эксперимент

Экспериментальное исследование имело дизайн 2×2 и включало факторы НАЛИЧИЕ ТО и СТЕРЕОТИП. В эксперимент включалось три типа сценариев, которые более стереотипны для мужчин (8a–c), три типа сценариев, которые более стереотипны для женщин (8d–f), а также контрольное условие — сценарии, для которых предположительно нет стереотипа о гендере протагониста (8g).

- (8) a. Агенса предикатов агрессии
(*избить, выгнать, сломать руку*)
- b. Экспериментатор стенических предикатов
(*обрадоваться, оживиться, прийти в восторг*)
- c. Агенса предикатов профессиональной деятельности
(*посадить самолет, созвать собрание, разработать модель*)
- d. Пациента предикатов агрессии
(*уволить, ограбить, оскорбить*)
- e. Агенса предикатов ведения домашнего хозяйства
(*испечь шарлотку, помыть полы, протереть пыль*)
- f. Экспериментатор астенических предикатов
(*разрыдаться, застесняться, испугаться*)
- g. Пациента предикатов изменения состояния и неконтролируемых действий
(*удариться, простудиться, поскользнуться*)

Для каждого типа предикатов было составлено восемь лексикализаций, что в итоге дало 56 предложений, которые были распределены по четырем экспериментальным листам по методу латинского квадрата. Эксперимент также включал 56 филлеров и четыре тренировочных предложения, которые представляли собой конструкции с составными союзами *едва ... как и если ... то и*

были сбалансированы по грамматичности. Грамматичные филлеры были также сбалансированы с точки зрения выраженности коррелята *как* или *то*. Неграмматичность филлеров достигалась за счет использования несуществующих сочетаний указанных союзов и коррелятов **едва ... то* и **если ... как*. Примеры стимулов приведены в (9), примеры грамматичных и неграмматичных филлеров — в (10a–b) и (10c–d) соответственно.

- (9) *Когда завод начал выпуск мотоциклов, ...*
- a. ... *Дарья разработала новую модель двигателя.*
[– ТО; – СТЕРЕОТИП]
 - b. ... *Дмитрий разработал новую модель двигателя.*
[– ТО; + СТЕРЕОТИП]
 - c. ... *то Дарья разработала новую модель двигателя.*
[+ ТО; – СТЕРЕОТИП]
 - d. ... *то Дмитрий разработал новую модель двигателя.*
[+ ТО; + СТЕРЕОТИП]
- (10) a. *Едва вдали показался поезд, как на перроне собрались торговки ягодами.*
- b. *Если в сапоги попала вода, опытные туристы прокладывали их газетой.*
 - c. *Если последняя электричка уже ушла, как студенты ночевали на вокзале.*
 - d. *Едва на дороге появился пешеход, то велосипедисты сбавили скорость.*

В эксперименте на оценку приемлемости по шкале Ликерта 1–7 приняли участие 64 респондента. Эксперимент проводился с помощью платформы PennController for IBEX [Zehr, Schwarz 2018]. Для каждого респондента была подсчитана сумма квадратов отклонения филлеров от стандарта. Из анализа были исключены респонденты, у которых данный показатель превышал два стандартных отклонения от среднего. Также мы исключили тех респондентов, которые больше семи раз не успевали оценить предложение в ходе эксперимента. В результате статистический анализ проводился на материале ответов 54 респондентов (25 женщин и 29 мужчин, средний возраст 39, SD = 13).

Анализ результатов проводился в среде R [R Core Team, 2018] с использованием библиотек lme4 [Bates et al. 2015a] и lmerTest [Kuznetsova et al. 2017]. Исходные оценки были нормализованы, с тем чтобы исключить возможное искажение шкал респондентов. Анализ запланированных контрастов мы начали с составления максимальной модели, которая включала все возможные фиксированные и случайные эффекты (в зависимости от респондента и стимульного предложения) для факторов НАЛИЧИЕ ТО и СТЕРЕОТИП и их взаимодействия [Barr et al. 2013]. Контрольное условие в модель регрессии не включалось. После этого было предпринято итеративное сокращение компонентов, входящих в структуру случайных эффектов [Bates et al. 2015b]. Шесть моделей, удовлетворяющих требованию сходимости, сравнивались между собой с помощью теста отношения правдоподобия, а также с помощью информационных критериев Акаике (AIC) и Шварца (BIC). Наилучшей моделью оказалась модель, включающая фиксированные эффекты НАЛИЧИЕ ТО и СТЕРЕОТИП, и случайный угол наклона для фактора НАЛИЧИЕ ТО в зависимости от респондента и стимульного предложения (рисунок 1). Важно отметить, что принятая процедура сравнения моделей регрессии отдает предпочтение моделям с большим количеством компонентов, что позволяет модели лучше объяснить наблюдаемую дисперсию, даже если эти компоненты не являются значимыми. Как показано на рисунке 2, в итоговой модели значимым оказался только фактор НАЛИЧИЕ ТО, фактор СТЕРЕОТИП оказался незначимым.

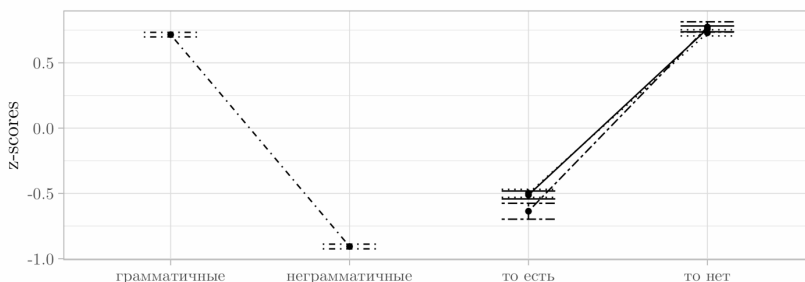
	Эффекты фиксированные	Эффекты случайные	npar	AIC	BIC	logLik	χ^2	χ^2 df	p-value
1	СТЕРЕОТИП + ТО	+(1 + ТО респ.) +(1 + ТО предл.)	10	4731.5	4789.5	-2355.7		2	
2	СТЕРЕОТИП + ТО	+(1 + ТО респ.) +(1 предл.)	8	4734.2	4780.6	-2359.1	6.71	1	0.035
3	ТО	+(1 + ТО респ.) +(1 предл.)	7	4732.2	4772.9	-2359.1	6.79	1	0.079
4	СТЕРЕОТИП + ТО	+(1 респ.) +(1 предл.)	6	4990.4	5025.3	-2489.2	266.97	1	≪ 0.0001 *
5	СТЕРЕОТИП	+(1 респ.) +(1 предл.)	5	5361.9	5391.0	-2676.0	640.47	0	≪ 0.0001 *
6	ТО	+(1 респ.) +(1 предл.)	5	4988.5	5017.6	-2489.3	267.06	0	≪ 0.0001 *

Рисунок 1. Результаты теста отношения правдоподобия для наилучшей модели (строка 1) и всех других исследованных моделей

Модель	z-оценка ~ СТЕРЕОТИП + ТО + (1 + ТО респ.) + (1 + ТО предл.)			
Фиксированные эффекты				
Эффект	Коэффициент	Стандартная ошибка (SE)	t-значение	p-значение
Отрезок	0.737	0.041	17.88	≪ 0.0001 *
СТЕРЕОТИП	0.016	0.036	0.44	0.662
НАЛИЧИЕ ТО	-1.249	0.077	-16.22	≪ 0.0001 *
Случайные эффекты				
Группа	Эффект	Дисперсия	SD	Корреляция
респондент	Отрезок	0.046	0.214	
	НАЛИЧИЕ ТО	0.237	0.487	-0.96
стимул	Отрезок	0.02	0.142	
	НАЛИЧИЕ ТО	0.072	0.268	-0.50
Остатки регрессии		0.354	0.595	
Количество наблюдений:		2461	предложение: 192	респондент: 52

Рисунок 2. Данные о коэффициентах фиксированных эффектов в выбранной модели вместе с данным о случайных эффектах. Степени свободы оценены с помощью метода аппроксимации Саттервайта (пакет lmerTest)

Попарные сравнения оценок для разных условий проводились в пакете emmeans [Lenth et al. 2019] с помощью критерия Тьюки (степени свободы оценены с помощью метода аппроксимации Кенварда-Роджера). Результаты сравнений показывают, что оценки респондентов были значимо ниже для тех предложений, которые включали частицу *то* (p-value < 0.001), причем фактор стереотипности описанной ситуации не оказал влияния на приемлемость предложений (p-value = 0.6651) (рисунок 3).



Условие ··· филлер — стереотип ····· не стереотип -·-· контрольное

Рисунок 3. Нормализованные оценки приемлемости на графике взаимодействия факторов НАЛИЧИЕ ТО и СТЕРЕОТИП в сравнении с филлерами.

На графике изображены интервалы, задаваемые стандартной ошибкой

5. Заключение

Результаты экспериментального исследования свидетельствуют об отсутствии взаимосвязи между приемлемостью и правдоподобием, определяемым в терминах стереотипных гендерных сценариев. Использование участников разных гендеров в стимулах с гендерно-стереотипными ситуациями не повлияло ни на абсолютный уровень приемлемости целевых предложений, ни на относительный контраст между условиями. Отсюда мы делаем вывод, что устранение гендерно-стереотипных формулировок из иллюстративных материалов и стимулов экспериментальных исследований не приводит к неправдоподобию и снижению достоверности данных.

Насколько нам известно, это первое исследование, в котором изучается влияние гендерных стереотипов на правдоподобие языковых выражений. Для целей нашего эксперимента мы определили правдоподобие как вероятность/прототипичность некоторой ситуации с учетом гендерного противопоставления. Однако результаты этого исследования не предполагают, что эффекта правдоподобия не существует в принципе. Полученные нами результаты не применимы к случаям, когда эффект неправдоподобия возникает, например, из-за нарушения аргументом семантических селективных ограничений предиката, таких как требование множественного или одушевленного аргумента.

Следует подчеркнуть, что предикаты, связанные с гендерными стереотипами, довольно часто используются в стимульном материале экспериментальных исследований и в иллюстративном материале научных работ и учебных пособий по лингвистике: предикаты профессиональной деятельности и эмоций удобны тем, что задают контекст для однозначно определяемой семантической роли протагониста (агенс, экспериенцер), а предикаты агрессии позволяют поместить в одну предикацию участников с максимально различными семантическими ролями (агенс и пациенс). Предпринятое нами исследование показывает, что внесение гендерного разнообразия в такие материалы не повлечет их несоответствия своему предназначению, но при этом поможет заинтересованному лингвисту избежать ненамеренной поддержки гендерных стереотипов.

Список условных сокращений

1SG — 1 лицо единственного числа; CVB — конверб; PST — прошедшее время.

Литература

- Кобозева И.М. Глава 3. Коннекторы, выражающие отношение непосредственного предшествования одного события другому // Инькова О.Ю. (ред.). Структура коннекторов и методы ее описания. Москва: ТОРУС ПРЕСС, 2019. С. 87–117.
- Пекелис О.Е. Показатель *то* как средство акцентуации имплицативного отношения (на примере союза *если... то*) // Вопросы языкознания. 2015. № 2. С. 55–96.
- Пекелис О. Е. Корреляты. Материалы для проекта корпусного описания русской грамматики (<http://rusgram.ru>). На правах рукописи. М., 2018а.
- Пекелис О.Е. Эллипсис подлежащего главной клаузы в русском языке и его связь с типологией коррелятов // Вопросы языкознания. 2018b. № 6. С. 31–59.
- Русская грамматика: В 2 т. Т. I. Шведова Н.Ю. (гл. ред.). М.: Наука, 1980.
- Adger D. Core syntax: A minimalist approach. Oxford: Oxford University Press, 2003.
- Barr D.J., Levy R., Scheepers C., Tily H.J. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*. 2013. Vol. 68. P. 255–278.
- Bates D., Maechler M., Bolker B., Walker S. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*. 2015a. Vol. 67. P. 1–48.
- Bates D., Kliegl R., Vasishth S., Baayen H. *Parsimonious mixed models*. 2015b. Vol. arXiv:1506.04967. Access mode: <https://arxiv.org/abs/1506.04967>.
- Boland J.E., Tanenhaus M.K., Garnsey S.M., Carlson G.N. Verb argument structure in parsing and interpretation: Evidence from wh-questions. *Journal of Memory and Language*. 1995. Vol. 34. P. 774–806.
- Cépeda P., Kotek H., Pabst K., Syrett K. Gender bias in linguistics textbooks: Has anything changed since Macaulay & Brice 1997? *Language*. 2021. Vol. 97. № 4. P. 678–702.
- Juzek T. S., Häussler J. Semantic Influences on Syntactic Acceptability Ratings. Gattnar A., Hörnig R., Störzer M., Featherston S. (eds.). *Proceedings of Linguistic Evidence*. Tübingen: University of Tübingen, 2018.
- Huang Y., Ferreira F. The Application of Signal Detection Theory to Acceptability Judgments. *Frontiers in Psychology*. 2020. Vol. 11. № 73.
- Inkova O. *La corrélation en russe. Structures et interprétations*. Berne: Peter Lang, 2014.
- Kizach J., Nyvad A.M., Christensen K.R. Structure before Meaning: Sentence Processing, Plausibility, and Subcategorization. *PLoS ONE*. 2013. Vol. 8. № 10.
- Kotek H., Dockum R., Babinski S., Geissler C. Gender bias and stereotypes in linguistic example sentences. *Language*. 2021. Vol. 97. № 4. P. 653–677.
- Kuznetsova A., Brockhoff P.B., Christensen R.H.B. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 2017. Vol. 82(13), P. 1–26.
- Lenth R., Singmann H., Love J., Buerkner P., Herve M. *emmeans: Estimated marginal means, aka least-squares means*. 2019. R package version 1.4.5. <https://cran.r-project.org/web/packages/emmeans/index.html>

- Levelt W.J., Van Gent J.A. W.M., Haans A.F.J., Meijers A.J.A. Grammaticality, paraphrase, and imagery. Greenbaum S. (ed.). *Acceptability in language*. 1977, The Hague: Mouton Publishers. P. 87–101.
- Macaulay M., Brice C. Gentlemen prefer blondes: A study of gender bias in example sentences. *Cultural performances: Proceedings of the Third Berkeley Women and Language Conference*. 1994. P. 449–461.
- Macaulay M., Brice C. Don't touch my projectile: Gender bias and stereotyping in syntactic examples. *Language*. 1997. Vol. 73. № 4. P. 798–825.
- Madariaga N. Diachronic change and the nature of pronominal null subjects: The case of Russian. A synchronic and diachronic perspective. Cognola F., Casalicchio J. (eds.). *Null Subjects in Generative Grammar*. Oxford Univ. Press, 2018. P. 171–198.
- Pabst K., Cépeda P., Kotek H., Syrett K., Donelson K., McCarvel M. *Gender bias in linguistics textbooks: Has anything changed since Macaulay & Brice (1997)?* Paper presented at the 92nd annual meeting of the Linguistic Society of America, Salt Lake City. 2018.
- Pekelis O.E. Correlative markers, contrastiveness and grammaticalization: A comparative study of conditional correlatives in Russian and Italian. *Italian Journal of Linguistics*. 2016. Vol. 28. №2. P. 143–180.
- Podlesskaya V.I. Syntax and semantics of resumption: some evidence from Russian conditional conjuncts. *Russian Linguistics*. 1997. Vol. 21. P. 125–155.
- R Development Core Team. R: A Language and Environment for Statistical Computing 3.4.4. R Foundation for Statistical Computing: Vienna, Austria, 2018.
- Richy C., Burnett H. Jean does the dishes while Marie fixes the car: A qualitative and quantitative study of social gender in French syntax articles. *Journal of French Language Studies*. 2020. Vol. 30. P. 47–72.
- Schütze C. *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Berlin: Language Science Press, 1996.
- Schütze, C., Sprouse, J. Judgment data. Sharma D., Podesva R. (eds.). *Research methods in linguistics*. Cambridge: Cambridge University Press, 2014. P. 27–50.
- Schütze C.T. Acceptability ratings cannot be taken at face value. Schindler S., Drożdżowicz A., Bröcker K. (eds.). *Linguistic intuitions: Evidence and method*. Oxford: Oxford University Press, 2020. P. 189–214.
- Sprouse J. The differential sensitivity of acceptability judgments to processing effects. *Linguistic Inquiry*. 2008. Vol. 39. №4. P. 686–694.
- Stanovsky G., Smith N.A., Zettlemoyer L. *Evaluating gender bias in machine translation*. 2019. Vol. arXiv:1906.00591. Access mode: <https://arxiv.org/abs/1906.00591>.
- Stowe L.A., Tanenhaus M.K., Carlson G.N. Filling gaps on-line: Use of lexical and semantic information in sentence processing. *Language and speech*. 1991. Vol. 34. №4. P. 319–340.
- Sun T., Gaut A., Tang S., Huang Y., El Sherief M., Zhao J., Mirza D., Belding E., Chang K.-W., Wang W. Y. *Mitigating gender bias in natural language processing: Literature review*. 2019. Vol. arXiv: 1906.08976. Access mode: <https://arxiv.org/abs/1906.08976>.

Tanenhaus M.K., Carlson G., Trueswell J.C. The role of thematic structures in interpretation and parsing. *Language and cognitive processes*. 1989. Vol. 4. №(3/4). P. 211–234.

Traxler M.J., Pickering M.J. Plausibility and the Processing of Unbounded Dependencies: An Eye-Tracking Study. *Journal of Memory and Language*. 1996. Vol. 35. P. 454–475.

Zehr J., Schwarz F. *PennController for Internet Based Experiments (IBEX)*. 2018.