

Взаимосвязь между приемлемостью и вероятностью высказываний: данные предикативного согласования с сочиненным подлежащим в русском языке

© 2024

Ксения Андреевна Студеникина

Московский государственный университет имени М. В. Ломоносова, Москва, Россия;
xeanst@gmail.com

Аннотация: Данное исследование направлено установить, способны ли языковые модели, обученные на неразмеченных текстовых данных, моделировать согласовательную вариативность. Мы провели сравнение суждений о приемлемости, вынесенных носителями языка, и вероятностных метрик, предсказанных двунаправленной языковой моделью ruBERT без тонкой настройки. В качестве лингвистического феномена рассматривалось предикативное согласование с сочиненным подлежащим в русском языке. Мы подробно проанализировали, какие синтаксические, морфологические и семантические факторы оказывают влияние на приемлемость и вероятность высказываний. Использование результатов синтаксических экспериментов позволило выявить роль каждого фактора по отдельности и их взаимодействия. Помимо стандартной метрики для оценки вероятности последовательности, использовались ее вариации, учитывающие либо длину предложения, либо вероятность каждого токена, либо оба этих параметра. Мы предполагали, что модель будет предсказывать наибольшую вероятность той стратегии, которая является наиболее приемлемой с точки зрения носителей языка. Однако гипотеза не подтвердилась: наличие вариативного согласования с сочиненным подлежащим снижает корреляцию между приемлемостью и вероятностью. Линейное расположение элементов предложения — позиция подлежащего и сказуемого, порядок конъюнктов — оказалось единственным фактором, который в равной степени оказывает влияние на приемлемость и вероятность высказывания. Совпадение рода конъюнктов повышает приемлемость предикативного согласования по единственному числу, однако не меняет его вероятность. Одушевленность конъюнктов и симметричность предиката не влияют ни на приемлемость, ни на вероятность высказывания. Наше исследование показывает, что ruBERT не может быть использован для моделирования предикативного согласования с сочиненным подлежащим. Приемлемость высказывания опирается на более тонкие языковые конформансы, которые не значимы при автоматической оценке его вероятности.

Ключевые слова: грамматическая приемлемость, русский язык, синтаксис, согласование, сочинение, эксперимент, языковая модель

Благодарности: Автор выражает благодарность Екатерине Анатольевне Лютиковой, Анастасии Алексеевне Герасимовой и другим участникам Московской группы экспериментального синтаксиса за плодотворное сотрудничество при проведении экспериментальных исследований по изучению вариативного согласования в русском языке. Именно их результаты послужили материалом для данной работы. Исследование выполнено при финансовой поддержке Некоммерческого Фонда развития науки и образования «Интеллект».

Для цитирования: Студеникина К. А. Взаимосвязь между приемлемостью и вероятностью высказываний: данные предикативного согласования с сочиненным подлежащим в русском языке. *Вопросы языкознания*, 2024, 6: 105–132.

DOI: 10.31857/0373-658X.2024.6.105-132

Interaction between acceptability and probability: Evidence from predicate agreement with a coordinated subject in Russian

Ksenia A. Studenikina

Lomonosov Moscow State University, Moscow, Russia; xeanst@gmail.com

Abstract: The study aims to establish whether language models trained on unlabeled text data can parametrize agreement variation. We compared the acceptability judgments made by native speakers and the probability metrics predicted by the language model ruBERT without fine-tuning. As a specific linguistic phenomenon, we considered predicate agreement with a coordinated subject in Russian. We analyzed in detail which syntactic, morphological, and semantic factors influenced sentence acceptability and probability. The experimental data enables us to reveal the role of each factor and their interaction. Besides the standard logarithmic probability, we considered sentence length and unigram probability. We assumed that the model would assign the highest probability to the most acceptable agreement strategy. However, our hypothesis was not confirmed: the correlation between probability and acceptability is lower for sentences with agreement variation than for sentences without variation. The linear position — the subject-predicate order and the conjuncts' order — turned out to be the only factor which equally influences the acceptability and probability of a sentence. If the gender features of conjuncts match, the acceptability of singular agreement increases while the probability does not change. The animacy of conjuncts and the predicate symmetry influence neither acceptability nor probability. Our research demonstrates that ruBERT cannot be used to parametrize predicate agreement with a coordinated subject. The acceptability of a sentence is based on subtle linguistic contrasts which are not significant for the computer evaluation of its probability.

Keywords: acceptability, agreement, coordination, experiment, language model, Russian, syntax

Acknowledgements: The author expresses gratitude to Ekaterina A. Lyutikova, Anastasia A. Gerasimova and other participants of the Moscow Experimental Syntax Group for productive cooperation in conducting experimental studies on agreement variation in Russian. The current paper is based on the results of these studies. This work was supported by Non-Profit Foundation for the Advancement of Science and Education "INTELLECT".

For citation: Studenikina K. A. Interaction between acceptability and probability: Evidence from predicate agreement with a coordinated subject in Russian. *Voprosy Jazykoznanija*, 2024, 6: 105–132.

DOI: 10.31857/0373-658X.2024.6.105-132

1. Введение

Умение оценивать грамматическую правильность предложения на родном языке — ключевое свойство языковой способности человека [Chomsky 2014]. Под грамматичностью подразумевается интуитивная оценка, основанная на врожденной грамматике. Помимо грамматической правильности, на общую приемлемость предложения влияют факторы употребления: частотность слов, длина предложения, семантическая сочетаемость слов, объем рабочей памяти говорящего. Суждения о приемлемости предложений называют «оценками приемлемости». Они показывают, в какой степени высказывание является естественным по мнению носителей языка. В лингвистике оценки приемлемости используются для проверки грамматических теорий. Они выступают как один из основных типов данных для моделирования языковой способности людей и анализа восприятия говорящими определенного языкового выражения. Для извлечения оценок приемлемости проводится синтаксический эксперимент.

Суждения о приемлемости представляют интерес как для бинарно-категориальной, так и для вероятностной теорий грамматичности [Sprouse, Schütze 2014; Sprouse 2015].

Они являются градиентными по своей природе и поэтому не могут быть непосредственно учтены с помощью бинарных ограничений формальной грамматики. Важный вопрос состоит в том, можно ли свести приемлемость высказывания к его вероятности, которая широко исследуется в области компьютерной обработки языка.

Вероятность некоторой цепочки слов можно вычислить с помощью языковой модели. Чтобы она могла корректно рассчитывать вероятности, модель предварительно обучают на неразмеченном корпусе текстов. Алгоритм анализирует, какие комбинации слов и в каком порядке чаще всего встречаются вместе. Чем больше и разнообразнее набор текстов, на которых модель обучается, тем более сложные зависимости она улавливает и воспроизводит на новых данных. Первые языковые модели были построены на основе частотности *n*-грамм. С распространением машинного обучения возросла популярность языковых моделей на основе рекуррентных нейронных сетей и на основе архитектуры Трансформер.

В данной работе мы представим сравнение оценок приемлемости, выставленных носителями языка, и вероятностных метрик, предсказанных нейронной языковой моделью BERT. Материалом для исследования послужили данные синтаксических экспериментов по изучению предикативного согласования с сочиненным подлежащим. Анализ этой конфигурации позволяет выявить синтаксические, морфологические и семантические факторы, которые определяют выбор стратегии согласования при наличии множественного контролера. Сравнение суждений о приемлемости со значениями вероятности показывает, насколько обнаруженные корреляции релевантны для нейронных сетей. Мы намерены ответить на вопрос, способны ли языковые модели, обученные на неразмеченных данных без эксплицитной грамматической информации, усваивать тонкие лингвистические различия, которые важны при наличии согласовательной вариативности.

Статья имеет следующую структуру. В разделе 2 мы рассмотрим ряд предшествующих исследований, которые изучали связь между приемлемостью и вероятностью предложений. Раздел 3 посвящен проблеме вариативного согласования и особенности сбора данных для его изучения. В разделе 4 сосредоточимся на одном из контекстов согласования с несколькими контролерами — предикативном согласовании с сочиненным подлежащим. Раздел 5 содержит технические детали нашего эксперимента, а именно аргументацию выбора языковой модели и набор используемых вероятностных метрик. В разделе 6 мы описываем проведенный эксперимент по сравнению оценок приемлемости носителей и значений вероятности нейронной сети. Раздел 7 подводит итоги исследования.

2. Соотношение понятий «приемлемость» и «вероятность» высказывания

Вопрос о том, каким образом связаны приемлемость высказывания и вероятность его появления в речи, является фундаментальным. Еще Н. Хомский [Chomsky 1956] указывал, что установление соответствия между этими понятиями затруднительно. Так, предложения (1a) и (1b) одинаково невозможно встретить в английской речи, что делает их маловероятными. Однако при отсутствии семантической осмысленности обоих высказываний пример (1a) обладает грамматической правильностью, в отличие от примера (1b), что делает его более приемлемым. Утверждение об одинаково низкой вероятности примеров (1a) и (1b) опровергается в работе Ф. Перейра [Pereira 2000], в которой предлагается использовать биграммную языковую модель на основе классов для подсчета вероятности отсутствующих в обучающем корпусе биграмм. В результате вероятность высказывания (1a) оказывается в несколько раз выше, чем вероятность предложения (1b), что показано в выражении (2).

- (1) a. *Colorless green ideas sleep furiously.*
 b. *Furiously sleep ideas green colorless.*
 ‘Бесцветные зеленые идеи яростно спят’.

$$(2) \frac{P(\textit{Colorless green ideas sleep furiously.})}{P(\textit{Furiously sleep ideas green colorless.})} \approx 2 \times 10^5 \quad [\textit{Pereira 2000: 1245}]$$

Первое более подробное исследование взаимосвязи между приемлемостью и вероятностью представлено в статье Дж. Лау, А. Кларка и Ш. Лаппина [Lau et al. 2017]. На национальном корпусе английского языка были обучены *n*-граммная модель, скрытая марковская модель и рекуррентная нейронная сеть. При помощи краудсорсинга были собраны суждения о приемлемости 2500 предложений: из них 2000 предложений получены циклическим переводом 500 предложений из Британского национального корпуса на китайский, японский, норвежский, испанский и обратно. Суждения о приемлемости предложений собирались по шкалам трех типов: 1–2, 1–4, 1–100. Значения вероятности, порожденные моделью, сравнивались со средними оценками приемлемости носителей с помощью коэффициента корреляции Пирсона. Оказалось, что наблюдается значительная корреляция между предсказанными вероятностями и человеческими суждениями о приемлемости.

Тем не менее использование предложений, полученных с помощью циклического машинного перевода, не совсем корректно, поскольку они не отражают тонкие противопоставления, которые традиционно рассматриваются в экспериментальном синтаксисе. Данный недостаток был исправлен в работе Дж. Спрауза и коллег [Sprouse et al. 2018]. Рассматривалось три датасета, каждый из которых представляет интерес с точки зрения синтаксической теории. Первый набор данных состоял из минимальных пар из журнала «Linguistic Inquiry», в которых изменению подвергался только один грамматический параметр. Во второй датасет были добавлены предложения из учебника по теоретическому синтаксису Д. Аджера [Adger 2003]. Третья выборка содержала 120 предложений, которые получились путем перестановки слов в предложении (1a). Для экспериментов были выбраны *n*-граммная модель и рекуррентная нейронная сеть. Спрауз и коллеги выяснили, что данные языковые модели не способны предсказать вариативность, которая предсказывается категориальными оценками в учебнике Аджера, и не различают 23 % явлений, представленных в журнале «Linguistic Inquiry». Следовательно, при обращении к более сложным синтаксическим феноменам взаимосвязь между приемлемостью и вероятностью становится менее прозрачной.

Важным недостатком рассмотренных нами работ является тот факт, что они не приводят детальное сравнение оценок приемлемости и вероятностных метрик на материале конкретных лингвистических феноменов. Отсутствует информация о том, какие именно факторы учитываются носителями языка при вынесении суждения о естественности предложения и насколько они релевантны для языковых моделей.

Развернутая синтаксическая оценка была осуществлена в работе А. Варштадта и коллег [Warstadt et al. 2020] на материале корпуса лингвистических минимальных пар BLiMP. Он состоит из 67 отдельных наборов данных, каждый из которых содержит 1000 минимальных пар по приемлемости (3). Исследователи проводили эксперименты с *n*-граммной моделью, рекуррентной нейронной сетью и трансформерной моделью GPT-2. Наилучший результат продемонстрировала последняя модель. Было обнаружено, что наиболее точно идентифицируются морфологические контрасты, связанные с согласованием, но более тонкие семантические и синтаксические явления, такие как элементы отрицательной полярности, эллипсис и вопросительные выносы, вызывают трудности.

- (3) a. *The cats annoy Tim.* (приемлемо)
 b. **The cats annoys Tim.* (неприемлемо)
 ‘Кошки раздражают Тима’. [Warstadt et al. 2020: 377]

Таким образом, для простых случаев вроде предложений с предикативным согласованием подлежащего и сказуемого языковые модели довольно успешно предсказывают выбор наиболее приемлемого высказывания, тогда как при наличии дистантных зависимостей точность моделирования падает. Важно отметить, что предложения с предикативным согласованием и вопросительными выносами отличаются не только языковым феноменом, но и сложностью синтаксической структуры.

Интересующий нас вопрос состоит в том, насколько успешным будет моделирование предикативного согласования при наличии нестандартных или множественных контролеров, что делает возможным варьирование согласовательных вариантов. В следующем разделе мы детально рассмотрим проблематику согласования и ее релевантность для анализа взаимосвязи между приемлемостью и вероятностью высказываний.

3. Изучение вариативности при согласовании: проблемы и методы

Согласование (операция Agree) является одной из базовых операций в синтаксисе наряду с операцией соединения элементов (операция Merge) для построения предложения [Chomsky 1995]. В стандартном случае форма мишени согласования однозначно предсказывается на основании признаков контролера согласования, таких как лицо, число, род. Помимо контекстов, в которых выбор согласовательных признаков определяется однозначно, выделяются конструкции, где возможно вариативное согласование. В этом случае наблюдается неоднозначное вычисление согласовательных признаков мишени при фиксированных значениях признаков контролера. Варьирование возникает в ситуациях согласования с неканоническим контролером, таким как управляющий квантификатор (4), и при наличии нескольких потенциальных контролеров, как в сочинительных конструкциях (5).

(4) *Двое из нас придут / придет / *придем в гости.*

(5) *Петя и я идут / идем / *идут / *идет домой.*

Большинство исследований согласования в рамках русской научной традиции рассматривают контексты варьирования как нестандартный вариант, противопоставленный общему случаю. Л. Л. Иомдин [1990] называет их «осложненным согласованием», И. А. Мельчук [1993] — согласованием «по смыслу», А. Е. Кибрик [1999] — отклонением от ядерной модели. Однако простое перечисление случаев варьирования и осложняющих факторов недостаточно для разработки формальной модели согласования, основанной на единой системе признаков.

Эксплицитное моделирование вариативного согласования предполагает повышенные требования к эмпирическим данным. Выбор между согласовательными вариантами обусловлен множеством параметров, таких как лексические характеристики контролера и мишени, их линейный порядок, эффект синкретизма форм. Систематический сбор количественных данных о влиянии конкретных факторов осуществлялся в ряде предшествующих исследований, среди которых частотно-стилистический словарь по материалам газетного корпуса Л. К. Граудиной, В. А. Ицкович и Л. П. Катлинской [2001], монография по русскому синтаксису В. З. Санникова [2008], корпусные исследования конкретных контекстов варьирования [Пекелис 2013а; Кувшинская 2013; Добрушина, Сидорова 2019]. Однако данные источники содержат только описание приемлемых контекстов и не включают перечень неприемлемых случаев согласования. Чтобы разграничить допустимые и недопустимые стратегии согласования, необходимо подтвердить неприемлемость определенного контекста.

Моделирование согласовательной вариативности наиболее эффективно с опорой на данные экспериментальных исследований. В качестве стимулов выступают сконструированные

предложения, включающие набор всех исследуемых факторов, что позволяет учесть и статистически оценить их влияние. Кроме того, экспериментальные данные содержат сведения о допустимости всех типов контекстов, включая неприемлемые. Оценка некоторой конфигурации согласования как недопустимой подтверждает, что данный фактор ограничивает возможную вариативность. Методы экспериментального синтаксиса широко применяются для изучения вариативного согласования на материале различных языков, таких как словенский [Marušič et al. 2015], сербохорватский [Arsenijević, Mitić 2016], французский [An, Abeillé 2019], немецкий [Himmelreich, Hartmann 2023].

Проблематика вариативного согласования в русском языке активно разрабатывается Московской группой экспериментального синтаксиса¹. Ее участники занимаются изучением и формальным моделированием грамматических ограничений на основании экспериментальных данных. Исследованиям научной группы посвящен специальный выпуск журнала «Rhema. Рема» за 2023 год (№ 2). Статья Е. А. Лютиковой и А. А. Герасимовой [2023] рассматривает методологические проблемы, связанные с исследованием согласовательной вариативности. Результатом исследований Д. Д. Беловой и Т. И. Давидюк [2023], Д. Д. Врубель [2023], Л. И. Паско [2023] и К. А. Студеникиной [2023а] являются качественные эмпирические описания конкретных контекстов варьирования, сопровождаемые подробным статистическим анализом.

Результаты исследований Московской группы экспериментального синтаксиса могут выступить в качестве надежной эмпирической базы для анализа корреляции между приемлемостью и вероятностью. В качестве методики для всех экспериментов использовалась оценка приемлемости по шкале Ликерта от 1 до 7 [Likert 1932]. Данный метод является наиболее мощным в сравнении с другими методами извлечения суждений, как показали Дж. Спрауз, К. Шютце и Д. Алмеида [Sprouse et al. 2013] с помощью симуляции выборов. Экспериментальные исследования основываются на факторном подходе к изучению языковых явлений. Контролируя уровни факторов, мы можем оценивать влияние каждого из них по отдельности, а также их взаимодействия. Попарные сравнения позволяют понять, обнаруживаются ли значимые различия в приемлемости экспериментальных условий. Аналогичная процедура статистического анализа может быть применена и для вероятностных метрик. Если контрасты, наблюдаемые на материале оценок приемлемости, также выявляются для значений вероятности, можно говорить, что языковая модель способна успешно предсказывать выбор согласовательной стратегии. В следующем разделе будет подробно описан набор факторов, релевантных для моделирования вариативного согласования при сочинении.

4. Предикативное согласование с сочиненным подлежащим

Для автоматического моделирования варьирования было выбрано предикативное согласование с сочиненным подлежащим. Данный феномен позволяет проанализировать влияние разнообразных синтаксических, морфологических и семантических факторов. Исследователями Московской группы экспериментального синтаксиса накоплен значительный объем эмпирических данных о приемлемости, которые могут быть использованы для сравнения с вероятностными метриками языковых моделей. В подразделе 4.1 мы рассмотрим основные теоретические исследования данного явления на материале русского языка. Подраздел 4.2 посвящен экспериментальным исследованиям, результаты которых были использованы для анализа корреляции между приемлемостью и вероятностью высказываний.

¹ Сайт Московской группы экспериментального синтаксиса: <https://expsynt.com/>.

4.1. Теоретические исследования согласования при сочинении

В том случае, если подлежащее выражено сочиненной конструкцией, наблюдается несколько потенциальных контролеров предикативного согласования. Следовательно, возможны различные стратегии вычисления признаков для сказуемого — мишени согласования. Если подлежащее выражено сочиненными существительными единственного числа, для сказуемого наблюдается варьирование при вычислении признака числа. Если один из конъюнктов является местоимением и признаки лица конъюнктов различаются, глагол демонстрирует вариативность признака лица. В обоих случаях признаки глагола либо просто копируются от одного из конъюнктов, либо вычисляются в соответствии с правилами разрешения (англ. *resolution rules*), то есть некоторым образом выбираются исходя из того, какими значениями обладают конъюнкты [Corbett 1983].

Для конструкций, где подлежащее является сочиненной именной группой, возможны две стратегии согласования: полное или частичное. При полном согласовании признак числа вычисляется как сумма признаков обоих конъюнктов и сказуемое стоит во множественном числе (6). При частичном согласовании глагол получает признак единственного числа путем копирования признака только одного из конъюнктов (7). Рассмотрим факторы, определяющие выбор стратегии согласования в русском языке, на материале исследований В. З. Санникова [2008] и О. Е. Пекелис [2013б].

- (6) *Звук и свет **находятся** на единой шкале электромагнитных колебаний.* [«Волга», 2016]²
- (7) *Это был последний фильм, где звук и видеоряд **записывался** одновременно и вживую!* [Форум: 17 мгновений весны (2005–2010)]

В первую очередь, отмечается роль линейного порядка подлежащего и сказуемого. Преппозиция сказуемого (порядок VS) способствует его согласованию с ближайшим конъюнктом по единственному числу (8), а постпозиция (порядок SV) — согласованию со всей сочиненной группой (9).

[Пекелис 2013б]

- (8) *И вот уже **исчезли** / **исчез** и актер, и женщина.*
- (9) *И женщина, и актер **исчезли** / ***исчез**.*

Вероятность частичного согласования также повышается при совпадении признака рода сочиненных существительных. В этом случае согласовательные формы предиката совпадают для обоих конъюнктов, и любой из них способен выступать контролером согласования. Частичное согласование оказывается приемлемо, если оба конъюнкта имеют признак женского рода (10), однако недопустимо при сочинении существительного женского и мужского рода (11).

[Санников 2008: 158]

- (10) *Белена и крапива **росли** / **росла** прямо под окнами.*
- (11) *Белена и бурьян **росли** / ***росла** прямо под окнами.*

На выбор стратегии согласования влияет семантика глагола и конъюнктов. Для предикатов оказывается важной симметричность — семантическое свойство, которое предполагает аргументную структуру особого типа. Симметричным называется предикат, который обязательно имеет более одного участника с одинаковой семантической ролью. Так,

² Примеры (6)–(7) взяты из Национального корпуса русского языка (<https://ruscorpora.ru/>).

частичное согласование оказывается недопустимо, если сказуемое выражено симметричным предикатом, соотносящим субъекты друг с другом (12). Существительные выстраиваются в иерархию по способности индуцировать частичное согласование в зависимости от одушевленности / неодушевленности и конкретности / абстрактности: имена собственные (13) > одушевленные нарицательные > неодушевленные конкретные (14) > неодушевленные абстрактные (15). Чем выше по иерархии семантический тип существительного, тем выше приемлемость единственного числа глагола.

[Санников 2008: 157–159]

- (12) Там целуются / *целуется Коля и его невеста.
 (13) Отсюда мне видны / *виден Коля и Маша.
 (14) Отсюда мне виден / видны дом и опушка леса.
 (15) Во всем был виден / *были видны точный расчет и удивительная целеустремленность.

Для конструкций, где конъюнкты различаются по признаку лица, также возможно несколько стратегий предикативного согласования. Согласно нормативным грамматикам [РГ 1980; Розенталь и др. 1999], глагол получает признак лица конъюнкта, расположенного выше на личной иерархии $1 > 2 > 3$ [Zwicky 1977]. Если одним из конъюнктов является личное местоимение 1 лица, то согласование будет осуществляться по 1 лицу (16)–(17). Если в состав сочиненного ряда входит личное местоимение 2 лица, а второй конъюнкт имеет признак 3 лица, то «выигрывает» 2 лицо (18). Нарушение иерархии лиц возможно при порядке слов VS. В этом случае предикат копирует признак лица ближайшего конъюнкта (19). Кроме того, при сочинении местоимений 1 и 2 лица возможно согласование по 3 лицу. Данная стратегия называется «дефолтной»: ни один из конъюнктов не обладает признаком 3 лица, поэтому можно считать, что оно применяется по умолчанию.

[РГ 1980: 244]

- (16) Ни я, ни ты не едем.
 (17) Я и он придём.
 (18) Ты и твои родные не едете.
 (19) Остаются учителя старших классов и ты. [Пекелис 2013б]

Рассмотренные работы позволяют выделить релевантные факторы, определяющие выбор согласовательного варианта, однако они описывают отдельные сюжеты без учета полной картины параметров согласования. Чтобы осуществить статистическую оценку значимости факторов и их взаимодействия, были проведены экспериментальные исследования.

4.2. Экспериментальные исследования согласования при сочинении

Исследователи Московской группы экспериментального синтаксиса проанализировали влияние различных факторов на выбор стратегии согласования, среди которых тип конъюнктов (два местоимения / два существительных / одно местоимение и одно существительное), линейная позиция подлежащего и сказуемого (SV / VS), соотношение рода конъюнктов (оба мужского рода / оба женского рода / один мужского рода и один женского), симметричность предиката (симметричный / несимметричный), одушевленность конъюнктов (одушевленные / неодушевленные).

В данной статье мы сравним, в какой степени влияние перечисленных факторов на приемлемость высказывания совпадает с влиянием на его вероятность. Будут использованы

данные экспериментов Д. Д. Беловой, Д. Д. Врубель, Т. И. Давидюк, Л. И. Паско и К. А. Студеникиной, представленные на учебной конференции «Экспериментальные исследования языка» в 2022–2023 гг.

В табл. 1 кратко представлена характеристика указанных работ: рассматриваемый тип конъюнктов и независимые переменные. Эксперименты Д. Д. Врубель [2022], Л. И. Паско [2022] и К. А. Студеникиной [2023б] посвящены предикативному согласованию с именными конъюнктами по числу. Д. Д. Белова [2022] и Т. И. Давидюк [2022] изучали предикативное согласование по лицу и числу с конъюнктами, содержащими местоимение. Во всех исследованиях применялась методика извлечения суждений при помощи шкалы Ликерта от 1 до 7. Количество участников каждого эксперимента составило от 75 до 126 человек.

Таблица 1

Описание используемых экспериментальных исследований

Тип конъюнкта	Согласование	Исследуемый фактор	Источник
существительные	числовое	соотношение рода конъюнктов	[Врубель 2022]
		симметричность предиката, позиция конъюнктов и предиката	[Паско 2022]
		одушевленность конъюнктов, позиция конъюнктов и предиката	[Студеникина 2023б]
местоимения <i>я, ты</i>	лично-числовое	порядок конъюнктов, позиция конъюнктов и предиката	[Белова 2022]
местоимение <i>я</i> , существительное		порядок конъюнктов	[Давидюк 2022]

Помимо стимульных предложений, каждый эксперимент содержал филлеры. Их соотношение к экспериментальным предложениям составляло 1:1. Филлеры не демонстрировали контекстов варьирования и были разделены на два типа: грамматичные представляли собой полностью корректные предложения, неграмматичные содержали ошибку в согласовании (20) или в предложном управлении (21). При анализе оценок стимульных предложений можно опираться на оценки для филлеров в качестве границы (не)грамматичности. Вероятностные метрики для филлеров также могут выступать в качестве бейзлайна: они показывают, насколько успешно модель оценивает невариативное согласование.

[Белова, Давидюк 2023: 73]

(20) *Дима и Витя убираем на кухне.*

(21) *Кирилл и Федя ушли по киоск.*

Влияние рода конъюнктов на выбор стратегии согласования рассмотрела Д. Д. Врубель. Для эксперимента были выбраны конструкции с повторяющимся союзом и порядком слов VS, способствующие возникновению частичного согласования. Пример экспериментального блока представлен в табл. 2 (с. 114). Приемлемость частичного согласования оказывается значимо выше для стимулов с совпадающим признаком рода у обоих конъюнктов.

Значимость симметричности предиката и линейного расположения подлежащего и сказуемого проанализировала Л. И. Паско. Примеры стимульных предложений содержатся в табл. 3 (с. 114). Симметричность предиката не оказала значимого влияния на выбор стратегии согласования. Частичное согласование при порядке SV было оценено ниже, чем при порядке VS, но значимо выше, чем неграмматичные филлеры.

Таблица 2

Пример экспериментального блока для исследования Д. Д. Врубель [2022]

Род конъюнктов	Число предиката	Стимул
м. р. + м. р.	ед. ч.	В длинном проигрыше в песне звучал и орган, и барабан.
	мн. ч.	В длинном проигрыше в песне звучали и орган, и барабан.
ж. р. + ж. р.	ед. ч.	В длинном проигрыше в песне звучала и электрогитара, и скрипка.
	мн. ч.	В длинном проигрыше в песне звучали и электрогитара, и скрипка.
ж. р. + м. р.	ед. ч.	В длинном проигрыше в песне звучала и электрогитара, и орган.
	мн. ч.	В длинном проигрыше в песне звучали и электрогитара, и орган.
м. р. + ж. р.	ед. ч.	В длинном проигрыше в песне звучал и орган, и электрогитара.
	мн. ч.	В длинном проигрыше в песне звучали и орган, и электрогитара.

Таблица 3

Пример экспериментального блока для исследования Л. И. Паско [2022]

Тип предиката	Линейная позиция субъекта и предиката	Число предиката	Стимул
симметричный	SV	ед. ч.	На старой фотографии сливается лицо и фон.
		мн. ч.	На старой фотографии сливаются лицо и фон.
	VS	ед. ч.	Лицо и фон сливается на старой фотографии.
		мн. ч.	Лицо и фон сливаются на старой фотографии.
несимметричный	SV	ед. ч.	На старой фотографии стирается лицо и фон.
		мн. ч.	На старой фотографии стираются лицо и фон.
	VS	ед. ч.	Лицо и фон стирается на старой фотографии.
		мн. ч.	Лицо и фон стираются на старой фотографии.

Роль одушевленности конъюнктов и линейной позиции сказуемого исследовала К. А. Студеникина. Табл. 4 (с. 115) включает примеры стимулов. Одушевленность не повлияла на приемлемость стратегии предикативного согласования. Роль линейной позиции подтвердилась: порядок VS более приемлем при согласовании по единственному числу, а порядок SV — по множественному.

Распределение стратегий предикативного согласования с подлежащим, выраженным сочиненными местоимениями *я* и *ты*, исследовала Д. Д. Белова. Было проведено два эксперимента с одинаковыми факторами, различающиеся порядком субъекта и предиката. Примеры экспериментальных блоков представлены в табл. 5 (с. 115). Нормативное согласование по 1 лицу множественному числу получает наиболее высокие оценки в обоих экспериментах. Дефолтное согласование по 3 лицу множественному числу оказывается приемлемым вне зависимости от порядка конъюнктов при SV и VS. Согласование с ближайшим конъюнктом является допустимым только при инверсии субъекта и предиката. После глагола в форме 1 лица единственного числа сочиненная конструкция *я и ты* получает значимо более высокие оценки, чем *ты и я*. Эта же закономерность обнаруживается с формой 2 лица единственного числа.

Таблица 4

Пример экспериментального блока для исследования К. А. Студеникиной [2023б]

Одушевленность конъюнктов	Линейная позиция субъекта и предиката	Число предиката	Стимул
одушевленные	SV	ед. ч.	Воспитатель и ребёнок присел на скамейку в саду.
		мн. ч.	Воспитатель и ребёнок присели на скамейку в саду.
	VS	ед. ч.	На скамейку в саду присел воспитатель и ребенок.
		мн. ч.	На скамейку в саду присели воспитатель и ребенок.
неодушевленные	SV	ед. ч.	Кабачок и арбуз созрел на огороде у бабушки.
		мн. ч.	Кабачок и арбуз созрели на огороде у бабушки.
	VS	ед. ч.	На огороде у бабушки созрел кабачок и арбуз.
		мн. ч.	На огороде у бабушки созрели кабачок и арбуз.

Таблица 5

Пример экспериментального блока для исследования Д. Д. Беловой [2022]

Линейная позиция субъекта и предиката	Порядок конъюнктов	Форма предиката	Стимул
SV (эксперимент 1)	<i>я и ты</i>	1 л., мн. ч.	Я и ты строим крепость из снега.
		1 л., ед. ч.	Я и ты строю крепость из снега.
		2 л., ед. ч.	Я и ты строишь крепость из снега.
		3 л., мн. ч.	Я и ты строят крепость из снега.
	<i>ты и я</i>	1 л., мн. ч.	Ты и я строим крепость из снега.
		1 л., ед. ч.	Ты и я строю крепость из снега.
		2 л., ед. ч.	Ты и я строишь крепость из снега.
		3 л., мн. ч.	Ты и я строят крепость из снега.
VS (эксперимент 2)	<i>я и ты</i>	1 л., мн. ч.	Крепость из снега строим я и ты.
		1 л., ед. ч.	Крепость из снега строю я и ты.
		2 л., ед. ч.	Крепость из снега строишь я и ты.
		3 л., мн. ч.	Крепость из снега строят я и ты.
	<i>ты и я</i>	1 л., мн. ч.	Крепость из снега строим ты и я.
		1 л., ед. ч.	Крепость из снега строю ты и я.
		2 л., ед. ч.	Крепость из снега строишь ты и я.
		3 л., мн. ч.	Крепость из снега строят ты и я.

Изучением согласования по лицу и числу с сочиненным подлежащим, выраженным местоимением *я* и именем собственным, занималась Т. И. Давидюк. Для всех стимульных предложений был зафиксирован порядок слов VS. Табл. 6 содержит пример экспериментального блока. Наиболее высоко оценивается согласование по 1 лицу множественному числу. Также приемлемым оказывается согласование с ближайшим конъюнктом, однако для порядка *я и X* оно оценивается значимо выше, чем для порядка *X и я*. Согласование

по 3 лицу множественному числу является значимо более приемлемым, чем согласование со вторым конъюнктом, оценки для которого ниже, чем для неграмматичных филлеров.

Таблица 6

Пример экспериментального блока для исследования Т. И. Давидюк [2022]

Порядок конъюнктов	Форма предиката	Стимул
я и X	1 л., мн. ч.	В районном центре живём я и Вася.
	1 л., ед. ч.	В районном центре живу я и Вася.
	3 л., ед. ч.	В районном центре живёт я и Вася.
	3 л., мн. ч.	В районном центре живут я и Вася.
X и я	1 л., мн. ч.	В районном центре живём Вася и я.
	1 л., ед. ч.	В районном центре живу Вася и я.
	3 л., ед. ч.	В районном центре живёт Вася и я.
	3 л., мн. ч.	В районном центре живут Вася и я.

Для автоматического моделирования вариативного согласования результаты экспериментов были объединены в единый датасет. Каждому предложению была присвоена средняя оценка приемлемости на основе ответов всех респондентов. Датасет включает следующую информацию: автора эксперимента, текст предложения, среднюю оценку приемлемости, тип (стимул / филлер), для стимулов — тип субъекта (два существительных / я и ты / ты и я / я и X / X и я), тип согласования (ед. ч. / мн. ч. / 1 л. мн. ч. / 1 л. ед. ч. / 2 л. ед. ч. / 3 л. ед. ч. / 3 л. мн. ч.), линейная позиция (SV / VS), симметричность предиката (симметричный / несимметричный), одушевленность конъюнктов (одушевленные / неодушевленные), род конъюнктов (совпадает / не совпадает), для филлеров — тип (грамматичный / неграмматичный). Фрагмент датасета проиллюстрирован в табл. 7, весь набор данных доступен по ссылке³.

Таблица 7

Фрагмент набора данных для моделирования вариативного согласования

author	sentence	score	type	subject	agree- ment	order	sym- metry	ani- macy	gender	filler- type
Pasko	Земля и небо размывается на далеком горизонте.	5	stimul	noun_ noun	sg	SV	nonsym	NA	nonmatch	NA
Pasko	Земля и небо размываются на далеком горизонте.	5.64	stimul	noun_ noun	pl	SV	nonsym	NA	nonmatch	NA
Pasko	Земля и небо сходится на далеком горизонте.	3.2	stimul	noun_ noun	sg	SV	sym	NA	nonmatch	NA
Pasko	Земля и небо сходятся на далеком горизонте.	6.2	stimul	noun_ noun	pl	SV	sym	NA	nonmatch	NA

³ https://github.com/Xeanst/acceptability_probablitiy_coordination

author	sentence	score	type	subject	agree- ment	order	sym- metry	ani- macy	gender	filler- type
Pasko	Маша ценит доброту и честность в других людях.	6.60	filler	NA	NA	NA	NA	NA	NA	good
Pasko	Петя пишет сочинение и диктант на этом уроком.	2.79	filler	NA	NA	NA	NA	NA	NA	bad

Всего в датасет вошло 1660 предложений, среди которых 1404 стимула и 256 филлеров. В следующем разделе мы подробно рассмотрим процедуру выбора модели и подсчета метрик для оценки вероятности примеров из полученного набора данных.

5. Выбор языковой модели и метрик для оценки вероятности высказываний

В настоящее время наилучшие результаты компьютерной обработки текста достигаются с помощью больших языковых моделей на основе архитектуры Трансформер [Vaswani et al. 2017]. Благодаря механизму внимания подобные модели учитывают широкий контекст высказывания и улавливают дистантные синтаксические зависимости. Ожидается, что корреляция между оценками приемлемости и вероятностными метриками для языковых моделей на основе Трансформера должна быть значительно выше, чем для моделей менее сложной архитектуры.

Для современных больших языковых моделей важным становится разделение на однонаправленные (генеративные) и двунаправленные (автоэнкодерные). Исходно языковые модели, начиная с n -граммных, были однонаправленными. Поэтому вероятность P предложения s определялась как произведение вероятности каждого слова w_i с учетом левого контекста $w_{<i}$ (22). При обучении на размеченных текстовых данных для каждого слова предсказывается вероятность встретить его при условии, что известно предыдущее слово: w_2 при условии w_1 , w_3 при условии w_1 и w_2 и т. д. К однонаправленным или генеративным моделям на основе архитектуры Трансформер относятся модели семейства GPT [Radford et al. 2018]. Двунаправленные языковые модели наподобие BERT [Devlin et al. 2019] позволяют учитывать одновременно левый контекст $w_{<i}$ и правый контекст $w_{>i}$ при подсчете вероятности каждого слова w_i (23). В основе обучения лежит алгоритм маскированного языкового моделирования: случайно выбранные слова заменяются на спецсимвол [MASK], модель должна предсказать наиболее вероятный токен на месте маски.

$$(22) \vec{P}(s) = \prod_{i=0}^{|s|} P(w_i | w_{<i})$$

$$(23) \vec{P}(s) = \prod_{i=0}^{|s|} P(w_i | w_{<i}, w_{>i})$$

Корреляция между оценками приемлемости и вероятностными метриками для языковых моделей на основе Трансформера исследовалась в работе Дж. Лау и коллег [Lau et al. 2020]. Одни и те же высказывания рассматривались в трех конфигурациях: (i) изолированно, без контекста; (ii) с подходящим контекстом, в роли которого выступали три предшествующих предложения из исходного текста; (iii) с неуместным контекстом, в качестве которого использовались три случайных предложения. Приемлемые

предложения были взяты из Википедии, неприемлемые были получены с помощью циклического машинного перевода. Для предсказания вероятности последовательности использовались однонаправленная языковая модель GPT-2⁴ и двунаправленная модель BERT⁵. Результаты показали, что корреляция между оценками приемлемости и вероятностными метриками оказывается больше для двунаправленной языковой модели, чем для однонаправленной. Наличие релевантного контекста повышает приемлемость и вероятность предложений.

Сравнение приемлемости и вероятности проводилось также на материале русского языка. В исследовании В. Михайлова и коллег [Mikhailov et al. 2022] на материале корпуса лингвистической приемлемости для русского языка RuCoLA реализуется бинарная классификация предложений по приемлемости. Для применения двунаправленных моделей осуществляется тонкая настройка. Без дообучения применяется только однонаправленная модель ruGPT-3⁶. С помощью кросс-валидации исследователи находят пороговое значение для предсказанных ей вероятностных метрик. Бинарная классификация по приемлемости с помощью ruGPT-3 оказывается практически случайной: доля правильных ответов составила 55.79 %.

Вопрос о том, как связаны оценки приемлемости по шкале Ликерта и значения вероятности генеративной модели ruGPT-3, также был рассмотрен на материале русских именных групп с сочиненными модификаторами [Studenikina 2023]. Рассматриваемая языковая модель предсказывает корректные вероятностные метрики для предложений, где выбор согласовательной стратегии определяется однозначно. Для контекстов свободного варьирования вероятностные метрики также совпадают с оценками приемлемости. Однако предсказанные моделью вероятности не позволяют выявить контрасты в приемлемости, которые демонстрируют оценки носителей при наличии одной предпочтительной стратегии согласования.

Рассмотренные исследования демонстрируют, что вероятностные метрики однонаправленных моделей не вполне коррелируют с оценками приемлемости носителей. Поскольку они учитывают только левый контекст предложения, значительная часть информации о приемлемости теряется. В частности, генеративные модели не позволяют рассчитать, насколько вероятным будет появление определенной формы глагола при сочиненном подлежащем, если сказуемое расположено в препозиции. Поэтому для нашей задачи корректнее будет использовать вероятностные метрики, предсказанные двунаправленной языковой моделью. В нашем исследовании была использована модель ruBERT⁷ в базовой конфигурации, учитывающая регистр. Для удобства вычислений вместо исходной вероятности используется логарифмическая вероятность слов:

$$(24) LP = \sum_{i=0}^{|s|} \log P(w_i | w_{<i}, w_{>i})$$

Важно отметить, что вероятность предложения $P(s)$, подсчитанная с помощью языковой модели, не всегда напрямую связана с его приемлемостью. Влияние оказывают также факторы употребления. Поскольку с увеличением длины предложения его вероятность уменьшается, применяется нормализация по длине $|s|$ с возможным учетом коэффициента масштабирования α . Для учета частотности слов используется униграммная вероятность предложения $P_u(s)$. Она рассчитывалась по русскоязычному корпусу Araneum Russicum⁸ (120 миллионов вхождений).

⁴ <https://github.com/openai/gpt-2>

⁵ <https://github.com/google-research/bert>

⁶ https://huggingface.co/ai-forever/rugpt3medium_based_on_gpt2

⁷ <https://huggingface.co/DeepPavlov/rubert-base-cased>

⁸ <http://aranea.juls.savba.sk/guest/>

Помимо стандартной метрики для оценки вероятности последовательности (*LP*), были опробованы ее вариации, учитывающие либо длину последовательности (*MeanLP*, *PenLP*), либо вероятность каждого токена (*NormLP*), либо оба этих параметра (*SLOR*). Формулы для расчета метрик представлены в табл. 8.

Таблица 8

Метрики для подсчета вероятности предложения,
предложенные Дж. Лау и коллегами [Lau et al. 2020]

Метрика	Формула	Расшифровка
LP	$\log P(s)$	логарифмическая вероятность предложения
MeanLP	$\frac{\log P(s)}{ s }$	логарифмическая вероятность предложения, нормализованная по длине
PenLP	$\frac{\log P(s)}{\left(\frac{5 + s }{5 + 1}\right)^\alpha}$	логарифмическая вероятность предложения, нормализованная по длине с учетом коэффициента масштабирования
NormLP	$-\frac{\log P(s)}{\log P_u(s)}$	логарифмическая вероятность предложения, нормализованная с учетом униграммной вероятности
SLOR	$\frac{\log P(s) - \log P_u(s)}{ s }$	логарифмическая вероятность предложения, нормализованная по длине с учетом униграммной вероятности

В следующем разделе мы продемонстрируем, насколько каждая из рассмотренных метрик коррелирует с оценками приемлемости. Также будет проанализировано, способны ли данные метрики улавливать контрасты в приемлемости стратегий согласования, которые оказываются релевантными для носителей языка.

6. Анализ корреляции между вероятностными метриками
ruBERT и оценками приемлемости

Для анализа величины линейной связи между оценками приемлемости носителей и вероятностными метриками модели ruBERT мы использовали коэффициент корреляции Пирсона. Собранный набор данных содержит как предложения, где наблюдается вариативное согласование (стимулы), так и примеры, где выбор стратегии согласования однозначен (филлеры). Чтобы понять, насколько нейронные сети способны моделировать разные типы согласования, необходимо оценить коэффициент корреляции для каждой из этих групп. В табл. 9 представлены полученные значения. Вероятностные метрики гораздо лучше предсказывают невариативное согласование, значение коэффициента корреляции при наличии вариативности снижается. Для вариативных контекстов наибольшая корреляция наблюдается для метрики SLOR, учитывающей длину и частоту слов. Однозначное согласование лучше моделируется метрикой LP, которая представляет собой логарифмическую вероятность без учета дополнительных параметров.

Таблица 9

Значение коэффициента корреляции Пирсона для каждой из вероятностных метрик

	LP	MeanLP	PenLP	NormLP	SLOR
Наличие вариативности	0.27	0.31	0.30	0.39	0.45
Отсутствие вариативности	0.63	0.49	0.58	0.52	0.51

В ходе экспериментальных исследований предикативного согласования с сочиненным подлежащим были изучены различные факторы: тип конъюнктов (два существительных / два местоимения / местоимение и существительное), порядок слов (SV / VS), соотношение рода конъюнктов (совпадает / различается), тип предиката (симметричный / несимметричный), тип конъюнктов (одушевленные / неодушевленные), порядок конъюнктов (*я и ты, ты и я, я и X, X и я*). Статистический анализ оценок приемлемости, которые выставлены носителями языка, позволяет оценить значимость каждого из них. Мы рассмотрим, насколько данные факторы оказались важными для языковой модели *tuBERT*, осуществив аналогичный статистический анализ для вероятностных метрик. Для выявления значимости использовался многофакторный дисперсионный анализ ANOVA. В подразделе 6.1. мы рассмотрим фактор порядка слов, который был задействован во всех экспериментальных исследованиях. В подразделе 6.2 будут проанализированы факторы, которые влияют на выбор стратегии согласования с именными конъюнктами: соотношение рода, одушевленность существительных и симметричность глагола. Подраздел 6.3 описывает стратегии согласования для местоименных конъюнктов. В подразделе 6.4 представлен анализ предикативного согласования с подлежащим, содержащим местоимение и существительное. Подраздел 6.5 сравнивает результаты для разных вероятностных метрик.

6.1. Влияние порядка слов на выбор стратегии согласования

На рис. 1 (с. 121) представлены средние значения для оценок приемлемости и вероятностных метрик в зависимости от линейного порядка подлежащего и сказуемого. По оценкам можно заметить, что множественное число глагола более приемлемо при любом порядке слов, однако приемлемость единственного числа повышается при обратном порядке подлежащего и сказуемого. Дисперсионный анализ выявил значимое влияние на приемлемость предложений фактора числа ($p < .0001$), порядка слов ($p < .01$) и их взаимодействия ($p < .0001$). Фактор числа также оказывает значимое влияние на вероятность высказываний ($p < .0001$ для всех метрик), однако значимость фактора порядка слов не наблюдается для большинства метрик ($p < .05$ для MeanLP, $p > .05$ для LP, PenLP, NormLP и SLOR). Взаимодействие числа и порядка слов также значимо влияет на вероятность предложений ($p < .0001$ для LP, PenLP, NormLP и SLOR, $p < .001$ для MeanLP).

Разница в приемлемости и вероятности единственного и множественного числа при порядке SV оказывается значимой ($p < .0001$ для оценок носителей и всех вероятностных метрик). При порядке VS различие в приемлемости и вероятности единственного и множественного числа также значимо ($p < .0001$ для оценок, $p < .001$ для SLOR; $p < .01$ для LP, PenLP и NormLP; $p < .05$ для MeanLP). Контексты с единственным числом глагола значимо отличаются по приемлемости и вероятности в зависимости от порядка слов ($p < .0001$ для оценок, LP и MeanLP; $p < .001$ для NormLP; $p < .01$ для PenLP и SLOR). В большинстве случаев значимое различие не наблюдается между прямым и обратным порядком слов при множественном числе глагола ($p < .001$ для LP; $p > .05$ для оценок, MeanLP, PenLP, NormLP и SLOR).

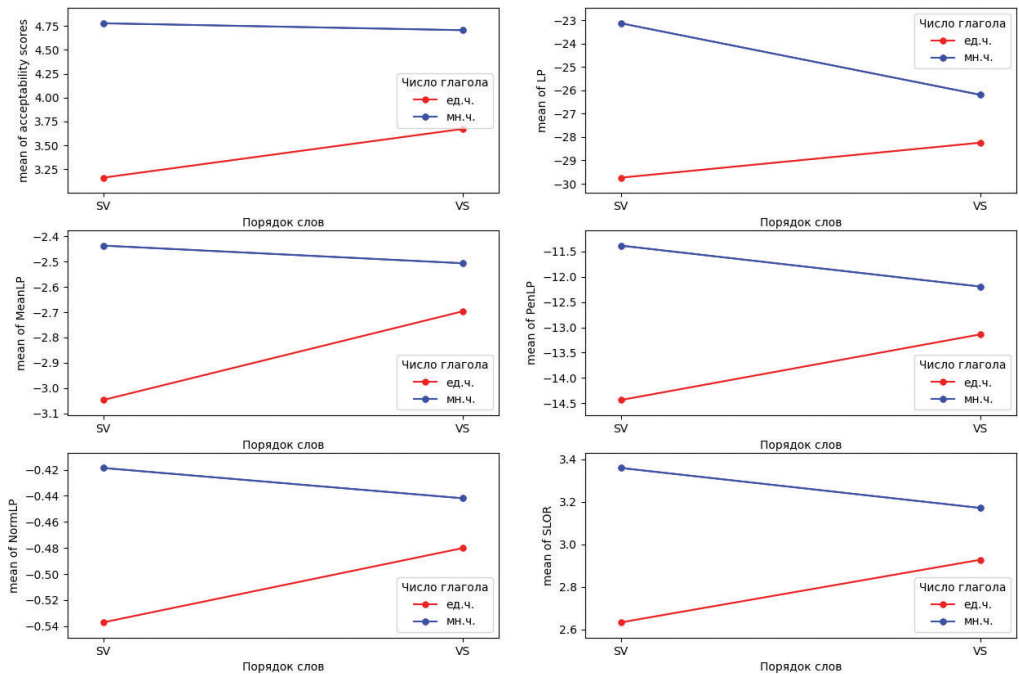


Рис. 1. Средние значения оценок приемлемости и вероятностных метрик в зависимости от порядка слов

Контрасты в приемлемости, наблюдаемые при различном порядке слов, оказываются релевантны и для значений вероятности. Большинство метрик показывают схожий результат, отличие демонстрирует только метрика LP, выявившая избыточный контраст при одном из попарных сравнений. Различия в линейной позиции слов оказываются значимыми при выборе стратегии согласования как для носителей, так и для нейросетевой модели.

6.2. Анализ морфологических и семантических факторов для именных конъюнктов

Рис. 2 демонстрирует средние значения оценок приемлемости и вероятностных метрик с учетом соотношения рода именных конъюнктов. Влияние рода на оценки приемлемости аналогично влиянию линейного порядка: множественное число глагола одинаково приемлемо при любом соотношении признаков рода, приемлемость единственного числа повышается при совпадении рода конъюнктов. Дисперсионный анализ показывает, что приемлемость предложений зависит от числа сказуемого ($p < .0001$), рода подлежащих ($p < .0001$) и взаимодействия этих факторов ($p < .05$). Для вероятностных метрик подтверждается только значимость фактора числа ($p < .01$ для NormLP и SLOR; $p < .05$ для LP, MeanLP и PenLP). На вероятность не оказывают значимого влияния фактор рода конъюнктов ($p > .05$ для всех метрик) и взаимодействие факторов ($p > .05$ для всех метрик).

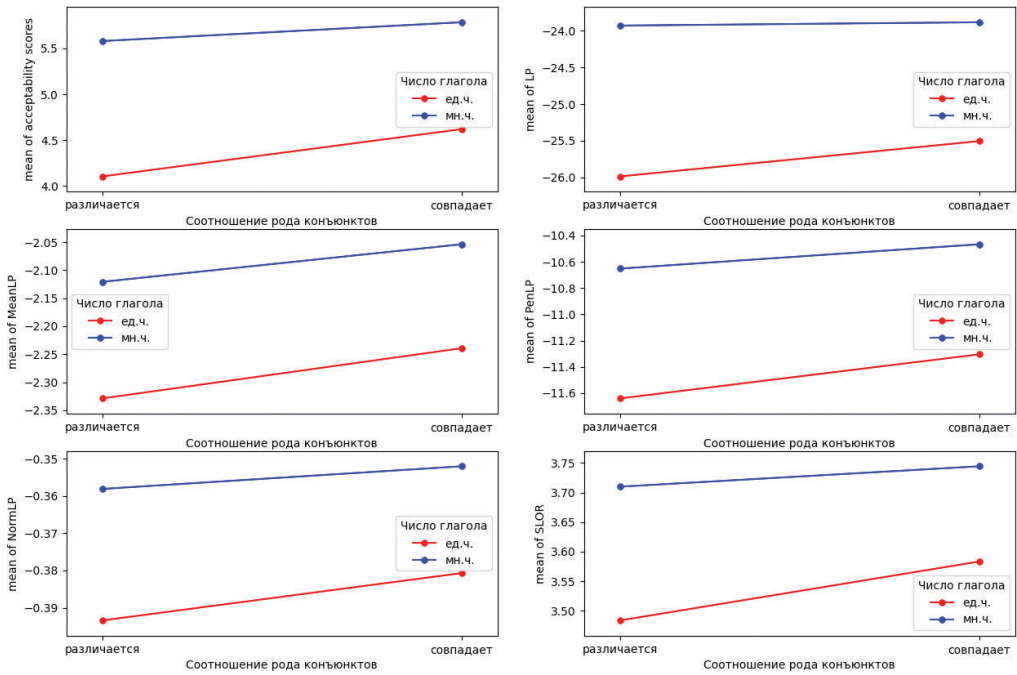


Рис. 2. Средние значения оценок приемлемости и вероятностных метрик в зависимости от соотношения рода конъюнктов

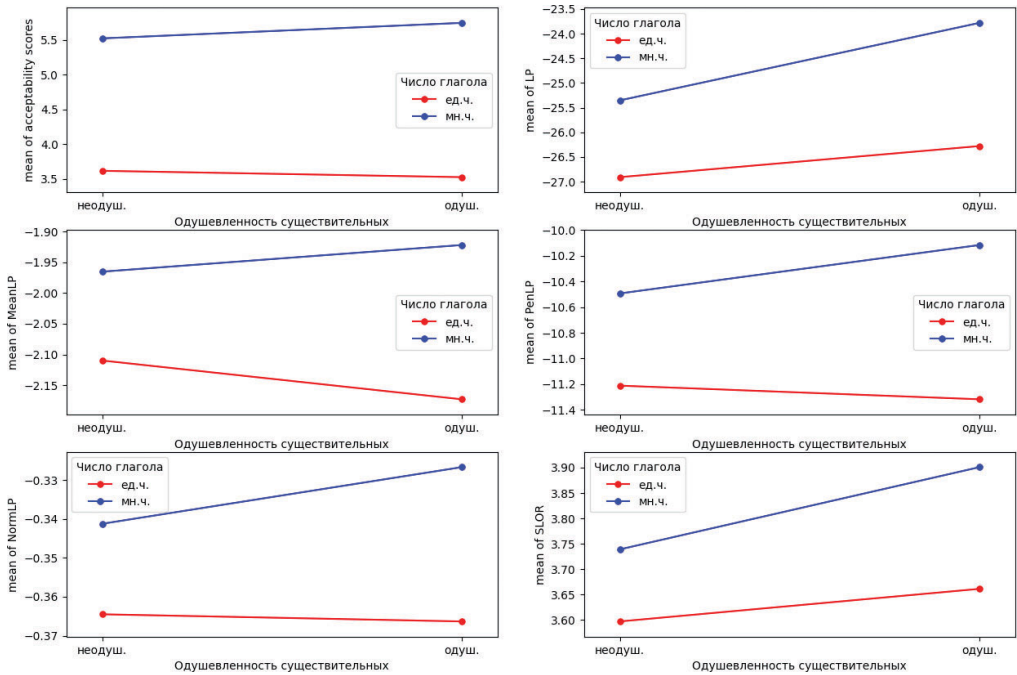


Рис. 3. Средние значения оценок приемлемости и вероятностных метрик в зависимости от одушевленности конъюнктов

Попарные сравнения методом Тьюки выявили значимую разницу в приемлемости единственного и множественного числа при совпадении рода конъюнктов ($p < .0001$) и при различии рода ($p < .0001$). Оценки приемлемости значимо меняются при совпадении и при различии рода, если сказуемое стоит в единственном числе ($p < .0001$), однако не меняются при множественном числе сказуемого ($p > .05$). Что касается контрастов для вероятностных метрик, статистический анализ не выявил значимых различий ни для одной пары условий ($p > .05$ для всех метрик). В отличие от приемлемости, вероятность единственного числа предиката не зависит от соотношения рода конъюнктов. Кроме того, вероятностные метрики не отражают большую приемлемость множественного числа при фиксированном соотношении рода конъюнктов.

На рис. 3 (с. 122) изображены средние значения оценок приемлемости и вероятностных метрик при различном семантическом типе именного подлежащего. Дисперсионный анализ суждений носителей и вероятностных метрик выявил значимость фактора числа ($p < .0001$ для оценок; $p < .05$ для всех метрик). Фактор одушевленности и взаимодействие факторов оказываются незначимыми ($p > .05$ для оценок и всех метрик).

Попарные сравнения выявили значимые различия в приемлемости между единственным и множественным числом глагола при одушевленных ($p < .0001$) и при неодушевленных существительных ($p < .0001$). Однако предложения с одушевленными и неодушевленными конъюнктами значимо не отличаются в приемлемости как в единственном ($p > .05$), так и во множественном числе ($p > .05$). В отношении вероятностных метрик попарные сравнения не выявили значимых контрастов ни для одного из условий ($p > .05$ для всех метрик). Следовательно, одушевленность конъюнктов не влияет ни на приемлемость, ни на вероятность при выборе числа предиката. Большая приемлемость множественного числа глагола при фиксированном семантическом типе существительных не отражается с помощью вероятностных метрик.

Рис. 4 показывает средние значения оценок приемлемости и вероятностных метрик при различном семантическом типе сказуемого. Дисперсионный анализ продемонстрировал

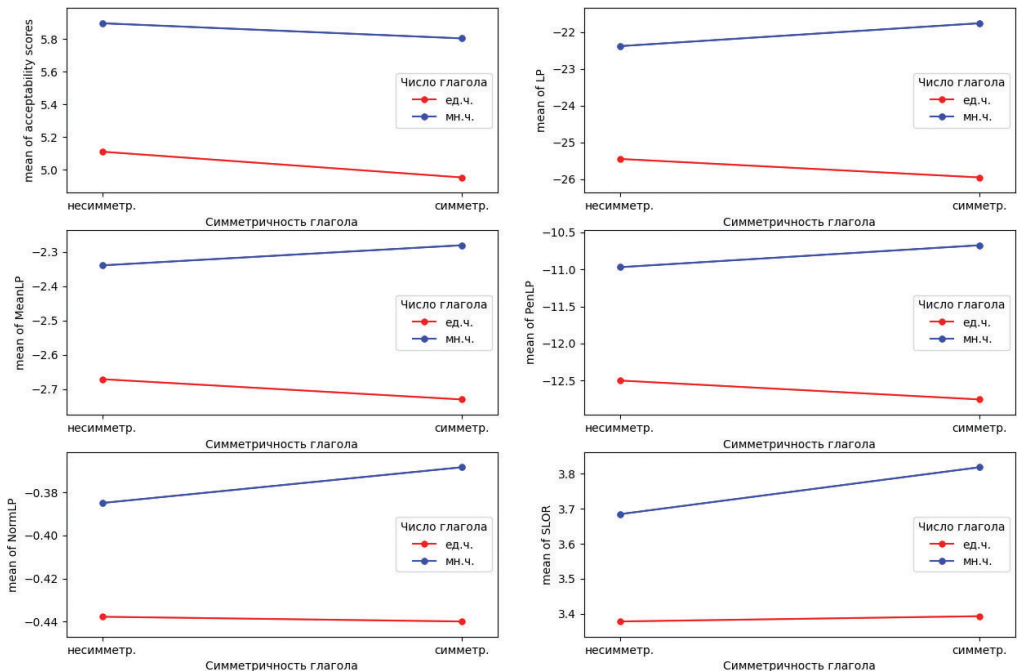


Рис. 4. Средние значения оценок приемлемости и вероятностных метрик в зависимости от симметричности предиката

значимое влияние фактора числа на приемлемость и вероятность высказываний ($p < .0001$ для оценок; $p < .001$ для LP, PenLP, NormLP и SLOR; $p < .01$ для MeanLP). Однако приемлемость и вероятность высказывания не зависят ни от симметричности предиката, ни от взаимодействия этих факторов ($p > .05$ для оценок и всех метрик).

Попарные сравнения показывают, что предложения с единственным и множественным числом глагола получают значимо разные оценки приемлемости при наличии как симметричного ($p < .0001$), так и несимметричного предиката ($p < .0001$). Предложения, которые отличаются только семантическим типом глагола, не различаются по приемлемости ни в единственном ($p > .05$), ни во множественном числе ($p > .05$). Вероятность предложений с симметричными предикатами значимо отличается при единственном и множественном числе глагола ($p < .05$ для всех метрик). Остальные пары условий значимо не различаются по вероятности ($p > .05$ для всех метрик). Таким образом, симметричность предиката не оказывает влияния ни на приемлемость, ни на вероятность той или иной стратегии согласования. Для симметричных предикатов выявляется большая вероятность и большая приемлемость множественного числа. Аналогичный контраст в приемлемости для несимметричных предикатов не подкрепляется при сравнении значений вероятности.

6.3. Стратегии согласования для местоименных конъюнктов

Рассмотрим различные стратегии согласования для местоименных конъюнктов. На рис. 5 изображены средние значения оценок приемлемости и вероятностных метрик при прямом порядке подлежащего и сказуемого. Дисперсионный анализ оценок выявил значимое влияние фактора стратегии согласования ($p < .0001$), однако порядок местоимений

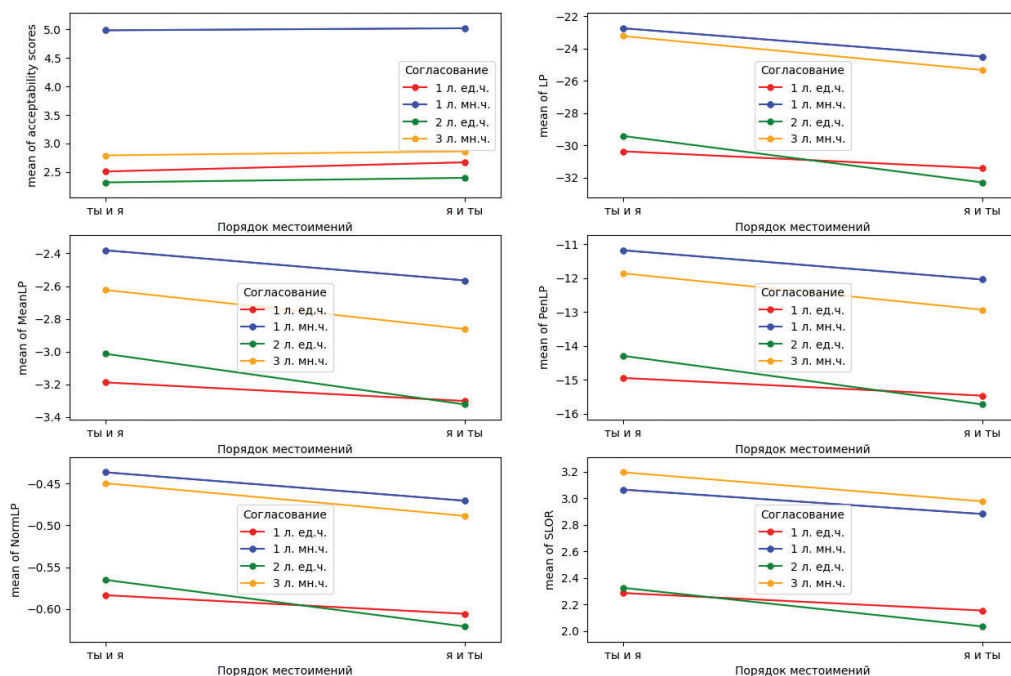


Рис. 5. Средние значения оценок приемлемости и вероятностных метрик для местоимений при порядке слов SV

и взаимодействие факторов оказываются не значимы ($p < .05$). В то же время на вероятность значимо влияют оба фактора: и форма глагола ($p < .0001$ для всех метрик), и тип подлежащего ($p < .01$ для MeanLP, PenLP, NormLP и SLOR; $p < .05$ для LP), тогда как взаимодействие факторов не значимо ($p < .05$ для всех метрик).

Попарные сравнения оценок приемлемости не выявили значимых различий между предложениями, которые отличаются только порядком конъюнктов при фиксированном типе согласования ($p > .05$ для 1 л. мн. ч., 1 и 2 лиц ед. ч., 3 л. мн. ч.). Вне зависимости от порядка местоимений, приемлемость предложений, где глагол принимает форму 1 л. ед. ч., значимо отличается от примеров с остальными типами согласования ($p < .0001$ при попарных сравнениях с формой 1 л. ед. ч., 2 л. ед. ч. и 3 л. мн. ч.). Также значимое различие в приемлемости наблюдается между согласованием по 2 л. ед. ч. и 3 л. мн. ч. для обоих порядков конъюнктов ($p < .01$). Приемлемость остальных стратегий отличается незначимо при любом типе подлежащего ($p > .05$ для форм 1 л. ед. ч. в сравнении с формами 2 л. ед. ч. и 3 л. мн. ч.).

В отношении вероятностных метрик значимые различия между примерами с различным порядком конъюнктов также не наблюдаются ($p > .05$ для всех метрик для каждой из глагольных форм). Для обоих типов сочиненных подлежащих формы 1 л. мн. ч. значимо более вероятны, чем формы 1 л. ед. ч. (*ты и я*: $p < .0001$ для всех метрик; *я и ты*: $p < .0001$ для NormLP, $p < .001$ для LP, MeanLP, PenLP и SLOR) и чем формы 2 л. ед. ч. (*ты и я*: $p < .0001$ для NormLP и SLOR, $p < .001$ для LP и PenLP, $p < .01$ для MeanLP; *я и ты*: $p < .0001$ для всех метрик). Отсутствует значимое различие в вероятности согласования по 1 л. и по 2 л. в ед. ч. (*ты и я*: $p > .05$ для всех метрик; *я и ты*: $p > .05$ для всех метрик), по 1 л. и 3 л. во мн. ч. (*ты и я*: $p > .05$ для всех метрик; *я и ты*: $p > .05$ для всех метрик). Вероятность предложений с формами 3 л. мн. ч. почти всегда значимо выше, чем с формами 1 л. ед. ч. (*ты и я*: $p < .0001$ для NormLP и SLOR, $p < .001$ для LP, $p < .01$ для PenLP, $p < .05$ для MeanLP; *я и ты*: $p < .0001$ для SLOR, $p < .001$ для NormLP, $p < .01$ для LP, $p < .05$ для PenLP, $p > .05$ для MeanLP) и 2 л. ед. ч. (*ты и я*: $p < .0001$ для SLOR, $p < .001$ для NormLP, $p < .01$ для LP, $p < .05$ для PenLP, $p > .05$ для MeanLP; *я и ты*: $p < .0001$ для NormLP и SLOR, $p < .001$ для LP, $p < .01$ для PenLP, $p > .05$ для MeanLP).

При порядке слов SV как приемлемость, так и вероятность той или иной стратегии согласования не зависят от расположения местоименных конъюнктов. Наиболее приемлемым и наиболее вероятным (кроме метрики SLOR) оказывается согласование по 1 лицу множественному числу, описываемое в нормативных грамматиках. Предложения с дефолтным согласованием по 3 лицу множественному числу являются значимо менее приемлемыми, однако различие в вероятности оказывается незначимым. Приемлемость дефолтного согласования незначимо отличается от приемлемости согласования по 1 лицу единственному числу, однако различие в вероятности является значимым (за исключением метрики MeanLP при подлежащем *я и ты*). Ни оценки приемлемости, ни вероятностные метрики не демонстрируют значимых различий между формами 1 и 2 лица единственного числа. Значимым оказывается различие в вероятности и приемлемости при сравнении нормативного согласования (1 лицо множественное число) с согласованием по единственному числу (1 или 2 лица). Дефолтное согласование (3 лицо множественное число) является значимо более вероятным и приемлемым, чем согласование по 2 лицу единственному числу (кроме метрики MeanLP).

Рис. 6 демонстрирует средние значения оценок приемлемости и вероятностных метрик при обратном порядке подлежащего и сказуемого. Значимое влияние на приемлемость оказывает выбор стратегии согласования ($p < .0001$), но не порядок местоимений ($p > .05$). Взаимодействие этих факторов также является значимым ($p < .0001$). Дисперсионный анализ выявил значимое влияние формы предиката только для некоторых вероятностных метрик ($p < .001$ для SLOR, $p < .05$ для NormLP, $p > .05$ для LP, MeanLP и PenLP). Значимого влияния на вероятность высказываний не оказывают порядок конъюнктов, а также взаимодействие факторов ($p > .05$ для всех метрик).

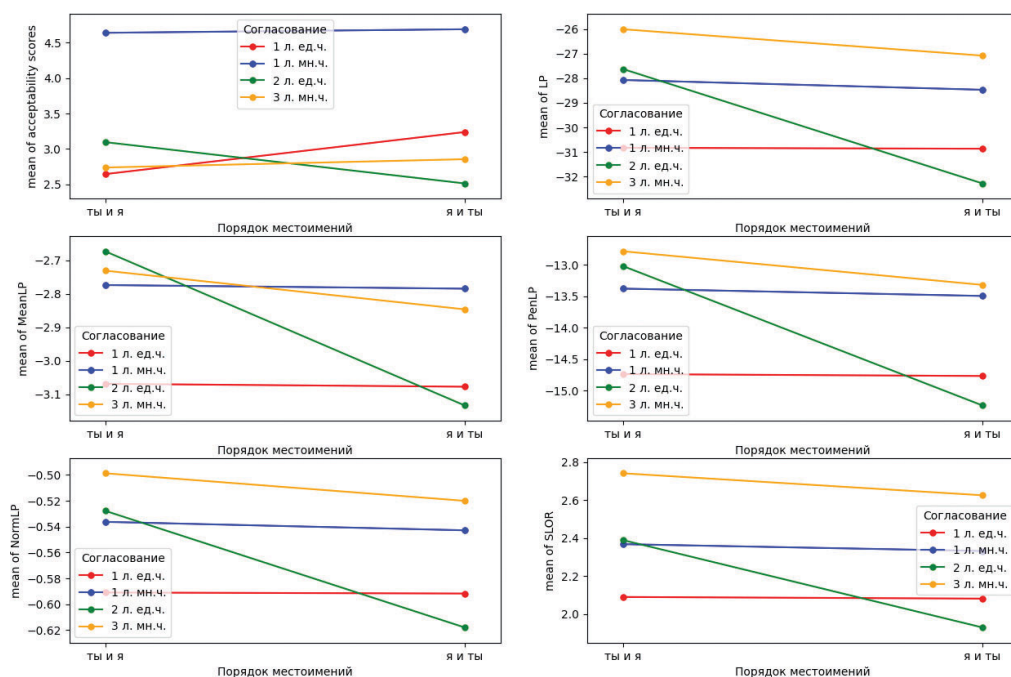


Рис. 6. Средние значения оценок приемлемости и вероятностных метрик для местоимений при порядке слов VS

Попарные сравнения демонстрируют, что приемлемость не зависит от порядка местоимений при нормативном согласовании по 1 л. мн. ч. ($p > .05$) и дефолтном согласовании по 3 л. мн. ч. ($p > .05$). Нормативное согласование все еще значимо более приемлемо, чем все остальные стратегии, при любом порядке конъюнктов ($p < .0001$ при сравнении 1 л. мн. ч. с 1 и 2 лицом ед. ч., 3 л. мн. ч.). В то же время согласование с ближайшим конъюнктом оказывается гораздо более допустимо при порядке VS. Предложения с предикатом в 1 л. и 2 л. ед. ч. значимо различаются по приемлемости (*ты и я*: $p < .01$, *я и ты*: $p < .0001$). Согласование по 2 л. ед. ч. значимо более приемлемо для подлежащего *ты и я* ($p < .0001$), а согласование по 1 л. ед. ч. — для подлежащего *я и ты* ($p < .0001$). При порядке конъюнктов *ты и я* дефолтное согласование значимо менее приемлемо, чем согласование по 2 л. ед. ч. ($p < .05$), но не отличается от согласования по 1 л. ($p > .05$). Для предложений с подлежащим *я и ты* выявляется значимая разница в приемлемости 3 л. мн. ч. по сравнению с ед. ч. 1 л. ($p < .05$) и 2 л. ($p < .05$). В отличие от оценок приемлемости, единственная пара условий, для которой различие в вероятности оказалось значимым, — 2 л. ед. ч. и 3 л. мн. ч. для подлежащего *я и ты* ($p < .05$ для SLOR; $p > .05$ LP, MeanLP, PenLP и NormLP). При попарном сравнении остальных условий вероятностные метрики не демонстрируют значимых контрастов ($p > .05$ для всех метрик).

В двух экспериментах с местоименными конъюнктами были использованы стимульные предложения с одинаковым лексическим содержанием, которые различались только порядком подлежащего и сказуемого (SV / VS). Изменение порядка при сохранении лексического материала оказывает сильное влияние на то, как соотносятся оценки приемлемости и вероятностные метрики. При порядке SV ряд контрастов, выявленных для оценок приемлемости, также демонстрируются при попарном сравнении вероятностных метрик. Однако при порядке VS почти ни для одного из условий, которые различаются по приемлемости, не наблюдается значимого различия по вероятности.

6.4. Согласование с подлежащим, содержащим местоимение и существительное

Рис. 7 демонстрирует средние значения оценок приемлемости и вероятностных метрик для подлежащего, выраженного местоимением 1 лица единственного числа (*я*) и именем собственным (*Х*). Все предложения имели одинаковый порядок слов (VS), но различались порядком конъюнктов и формой глагола. Дисперсионный анализ оценок приемлемости выявил значимое влияние стратегии согласования ($p < .0001$), типа подлежащего ($p < .01$), а также взаимодействия этих факторов ($p < .0001$). В отношении вероятностных метрик значимой оказывается только форма предиката ($p < .001$ MeanLP и PenLP; $p < .01$ LP, NormLP и SLOR), вероятность не зависит от порядка конъюнктов и взаимодействия факторов ($p > .05$ для всех метрик).

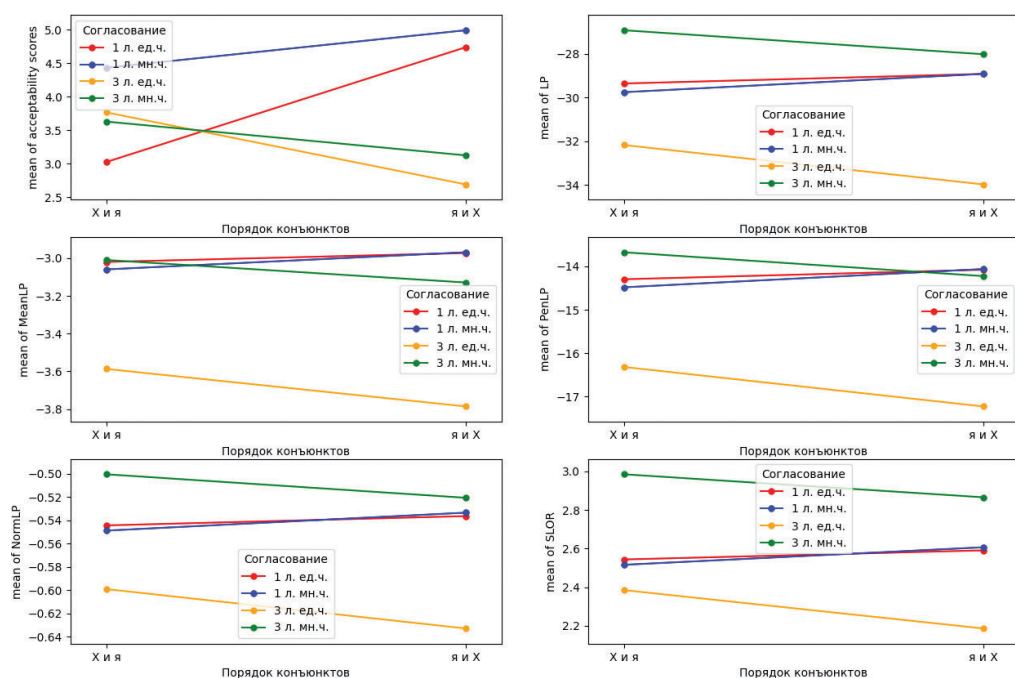


Рис. 7. Средние значения оценок приемлемости и вероятностных метрик для подлежащего *я* и *Х* при порядке слов VS

Попарные сравнения оценок приемлемости показывают, что предложения с различным типом подлежащего значимо отличаются друг от друга ($p < .0001$ для 1 л. и 3 л. ед. ч., $p < .001$ для 1 л. мн. ч., $p < .01$ для 3 л. мн. ч.). При любом порядке конъюнктов нормативная стратегия согласования (1 л. мн. ч.) значимо отличается от согласования по 3 л. ед. ч. (*Х и я*, *я и Х*: $p < .0001$) и 3 л. мн. ч. (*Х и я*, *я и Х*: $p < .0001$). Для подлежащего *я и Х* согласование по 1 л. мн. ч. и 1 л. ед. ч. одинаково приемлемы ($p > .05$), для подлежащего *Х и я* приемлемость нормативной стратегии значимо выше ($p < .0001$). Приемлемость согласования по 1 л. ед. ч. при подлежащем *Х и я* значимо ниже, чем по 3 л. ед. ч. ($p < .0001$) и 3 л. мн. ч. ($p < .0001$). Для порядка конъюнктов *я и Х* согласование по 1 л. ед. ч., наоборот, значимо выше, чем по 3 л. ед. ч. ($p < .0001$) и 3 л. мн. ч. ($p < .0001$). Предложения с предикатом 3 лица при подлежащем *Х и я* занимают среднюю позицию по приемлемости вне зависимости от числа ($p > .05$). При подлежащем *я и Х* приемлемость 3 л. ед. ч. значимо ниже, чем

3 л. мн. ч. ($p < .05$). Парные сравнения вероятностных метрик выявили отличия только для двух пар условий. Предложения с формой 3 л. ед. ч. при подлежащем *я* и *Х* оказываются значимо менее вероятны, чем с формой 1 л. ед. ч. ($p < .05$ для MeanLP; $p > .05$ для LP, PenLP, NormLP и SLOR) и 1 л. мн. ч. ($p < .05$ для MeanLP; $p > .05$ для LP, PenLP, NormLP и SLOR). Различия в вероятности между остальными условиями являются незначимыми ($p > .05$ для всех метрик).

Вероятностные метрики выявляют низкую приемлемость согласования по 3 лицу единственному числу, однако наибольшая приемлемость нормативного согласования (по 1 лицу множественному числу) не подтверждается высокой вероятностью. Кроме того, предложения с различным порядком конъюнктов отличаются по приемлемости, но не по вероятности.

6.5. Сравнение результатов для разных вероятностных метрик

Результат сравнения оценок приемлемости и значений вероятности представлен в табл. 10. Для всех 64 пар условий было посчитано количество истинных и ложных ответов. Истинно положительными исходами считались те, где контраст выявляется как для приемлемости, так и для вероятности. Истинно отрицательный исход наблюдается, если контраст отсутствует и для приемлемости, и для вероятности. Если условия не различаются по приемлемости, но им присваивается разная вероятность, то такие исходы считаются ложно положительными. Пары, для которых значимое различие есть для оценок приемлемости, но не для вероятностной метрики, являются самыми частотными и названы ложно отрицательными. Чем больше истинных исходов и чем меньше ложных исходов демонстрирует вероятностная метрика, тем лучше она коррелирует с приемлемостью.

Таблица 10

Количество контрастов между парами условий для вероятностных метрик в сравнении с контрастами в приемлемости

Тип исхода	Значимое различие		LP	MeanLP	PenLP	NormLP	SLOR
	приемлемость	вероятность					
истинно положительный	+	+	10	10	10	10	11
истинно отрицательный	–	–	16	18	17	17	17
ложно положительный	–	+	3	1	2	2	2
ложно отрицательный	+	–	35	35	35	35	34

Наибольшее количество истинно положительных ответов и наименьшее количество ложно отрицательных ответов получено для метрики SLOR. Больше всего истинно отрицательных исходов и меньше всего ложно положительных исходов продемонстрировала метрика MeanLP. Следовательно, для более точного отражения контрастов в приемлемости с помощью вероятностных метрик необходимо учитывать количество слов и их частотность (SLOR). Однако нормализация только по длине предложения без учета частоты встречаемости слов также показывает хороший результат (MeanLP).

7. Заключение

В данной статье мы проанализировали, как связаны между собой понятия «приемлемость» и «вероятность» высказываний на материале вариативного согласования в русском языке. Мы подробно рассмотрели, какие синтаксические, морфологические и семантические факторы оказывают влияние на приемлемость высказываний с точки зрения носителей языка, а какие — на вероятность, предсказываемую языковыми моделями. В качестве языкового феномена было выбрано предикативное согласование с сочиненным подлежащим. Использование данных синтаксических экспериментов позволило проанализировать влияние каждого фактора по отдельности, а также их взаимодействия.

Мы оценивали способность двунаправленной модели ruBERT-base моделировать вариативное согласование. Было использовано пять различных вероятностных метрик: логарифмическая вероятность предложения без нормализации (LP), с нормализацией по длине (MeanLP), с нормализацией по длине с учетом коэффициента масштабирования (PenLP), с учетом униграммной вероятности (NormLP), с нормализацией по длине и учетом униграммной вероятности (SLOR).

Ожидалось, что модель будет предсказывать наибольшую вероятность стратегии, которая является наиболее приемлемой с точки зрения носителей языка. Однако предположение не подтвердилось. Для примеров без вариативного согласования вероятностные метрики более точно отражают различия в приемлемости. При наличии вариативного согласования с сочиненным подлежащим корреляция между приемлемостью и вероятностью падает.

Вероятностные метрики наиболее точно отражают контрасты в приемлемости при изменении линейного порядка подлежащего и сказуемого. Согласование по единственному числу оказывается более вероятно и приемлемо при порядке VS, тогда как приемлемость и вероятность множественного числа не зависят от линейной позиции предиката.

Для именных конъюнктов значимые различия в вероятности не выявляются, если приемлемость зависит от морфологических характеристик подлежащего. Согласование по единственному числу более приемлемо при совпадении рода конъюнктов, однако на вероятность высказываний этот фактор не влияет. Семантический тип конъюнктов и предиката не оказывает влияния ни на приемлемость, ни на вероятность высказываний.

Для местоименных конъюнктов при порядке SV наиболее приемлемым и вероятным оказывается нормативное согласование (1 л. мн. ч.). Дефолтной стратегии (3 л. мн. ч.) приписывается слишком высокая вероятность при среднем уровне приемлемости. При порядке VS повышается приемлемость согласования с ближайшим конъюнктом (2 л. ед. ч. для *ты* и *я*, 1 л. ед. ч. для *я* и *ты*), однако эта тенденция не выявляется при анализе значений вероятности.

Если подлежащее выражено местоимением *я* и именем собственным, при порядке VS наиболее приемлемо согласование по 1 л. мн. ч. Также растет приемлемость согласования с ближайшим конъюнктом по 1 л. ед. ч. при подлежащем *я* и *Х*. Предсказанные языковой моделью значения вероятности не отражают эти закономерности.

При наличии вариативного согласования приемлемость и вероятность высказывания в значительной степени зависят от линейного расположения элементов предложения: позиции подлежащего и сказуемого, порядка конъюнктов. Морфологические факторы оказывают влияние на выбор наиболее приемлемой стратегии согласования, однако не являются значимыми при сравнении вероятности высказываний. Семантические особенности подлежащего и сказуемого не вызывают различий в приемлемости и вероятности. Наилучший результат показали метрики MeanLP и SLOR.

Можно сделать вывод, что при обучении на большом количестве неразмеченных текстовых данных двунаправленная языковая модель ruBERT не усваивает ряд важных аспектов грамматики и не может быть использована для моделирования вариативного согласования

сказуемого с сочиненным подлежащим без тонкой настройки. Понятие «вероятность» лишь опосредованно связано с понятием «приемлемость»: суждения о приемлемости отражают более тонкие контрасты при выборе стратегии согласования, которые не выявляются с помощью предсказываемых моделью значений вероятности.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Белова 2022 — Белова Д. Д. Предикативное согласование с подлежащим, выраженным местоименными конъюнктами. *Устный доклад на учебной конференции «Экспериментальные исследования языка», МГУ имени М. В. Ломоносова, Москва, 16 июня 2022 г.* [Belova D. D. Predicative agreement with the subject expressed by pronominal conjuncts. *Talk at the student conference “Experimental language studies”, Lomonosov Moscow State Univ., Moscow, June 16, 2022.*]
- Белова, Давидюк 2023 — Белова Д. Д., Давидюк Т. И. Согласование с сочиненным подлежащим, содержащим личное местоимение: экспериментальное исследование на материале русского языка. *Rhema. Рема*, 2023, 2: 53–88. [Belova D. D., Davidyuk T. I. Agreement with coordinated subjects containing a personal pronoun: Experimental data from Russian. *Rhema*, 2023, 2: 53–88.]
- Врубель 2022 — Врубель Д. Д. Синкретичное согласование по роду в конструкциях с повторяющимся союзом *и*. *Устный доклад на учебной конференции «Экспериментальные исследования языка», МГУ имени М. В. Ломоносова, Москва, 16 июня 2022 г.* [Vrubel D. D. Syncretic gender agreement in constructions with a repeated conjunction *i*. *Talk at the student conference “Experimental language studies”, Lomonosov Moscow State Univ., Moscow, June 16, 2022.*]
- Врубель 2023 — Врубель Д. Д. Эффект синкретизма при предикативном согласовании с сочинительными конструкциями с повторяющимся союзом *и*. *Rhema. Рема*, 2023, 2: 104–118. [Vrubel D. D. The role of syncretism in the predicate agreement with coordinate constructions with the Russian correlative conjunction *i...i*. *Rhema*, 2023, 2: 104–118.]
- Граудина и др. 2001 — Граудина Л. К., Ицкович В. А., Катлинская Л. П. *Грамматическая правильность русской речи: Опыт частотно-стилистического словаря вариантов*. М.: Наука, 2001. [Graudina L. K., Itskovich V. A., Katlinskaya L. P. *Grammatical correctness of speech: Towards a frequency and stylistic dictionary of variants*. Moscow: Nauka, 2001.]
- Давидюк 2022 — Давидюк Т. И. Влияние аргументной структуры предиката на лично-числовое согласование с сочиненным подлежащим. *Устный доклад на учебной конференции «Экспериментальные исследования языка», МГУ имени М. В. Ломоносова, Москва, 16 июня 2022 г.* [Davidyuk T. I. The influence of argument structure on predicative agreement in person and number with compound subjects. *Talk at the student conference “Experimental language studies”, Lomonosov Moscow State Univ., Moscow, June 16, 2022.*]
- Добрушина, Сидорова 2019 — Добрушина Е. Р., Сидорова М. И. Число предиката в конструкциях типа «те/все, кто пришел/пришли» и падеж вершины-корпусное исследование. *Вестник Православного Свято-Тихоновского гуманитарного университета. Сер. 3: Филология*, 2019, 59: 22–35. [Dobrushina E. R., Sidorova M. I. Number of the predicate in phrases *te/vse, kto prishel/prishli* ‘those/all who came (sg./pl.)’ and the case of the head: A corpus-based study. *St. Tikhon’s Univ. Review. Series III: Philology*, 2019, 59: 22–35.]
- Иомдин 1990 — Иомдин Л. Л. *Автоматическая обработка текста на естественном языке: модель согласования*. М.: Наука, 1990. [Iomdin L. L. *Avtomaticheskaya obrabotka teksta na estestvennom yazyke: model’ soglasovaniya* [Automatic natural language processing: Agreement model]. Moscow: Nauka, 1990.]
- Кибрик (ред.) 1999 — Кибрик А. Е. (ред.-сост.). *Элементы цахурского языка в типологическом освещении*. М.: Наследие, 1999. [Kibrik A. E. (ed.-comp.) *Elementy tsakhurskogo yazyka v tipologicheskoy osveshchenii* [Elements of Tsakhur in a typological perspective]. Moscow: Nasledie, 1999.]
- Кувшинская 2013 — Кувшинская Ю. М. Согласование сказуемого с подлежащим, выраженным именной группой с количественным значением (по данным НКРЯ за 2000–2010 гг.). *Русский язык в научном освещении*, 2013, 26: 112–151. [Kuvshinskaya Yu. M. Predicate agreement with the quantifier phrase in Russian (according to the data of the National Russian Corpus for the period of 2000–2010). *Russkij jazyk v nauchnom osveshchenii*, 2013, 26: 112–151.]

- Лютюкова, Герасимова 2023 — Лютюкова Е. А., Герасимова А. А. Исследование вариативного согласования в русском языке: проблемы и методы. *Rhema. Рема*, 2023, 2: 9–27. [Lyutikova E. A., Gerasimova A. A. Studying agreement variation in Russian: Problems and methodology. *Rhema*, 2023, 2: 9–27.]
- Мельчук 1993 — Мельчук И. А. Согласование, управление, конгруэнтность. *Вопросы языкознания*, 1993, 5: 16–58. [Mel'čuk I. A. Agreement, regimen, congruence. *Voprosy Jazykoznanija*, 1993, 5: 16–58.]
- Паско 2022 — Паско Л. И. Предикативное согласование с сочиненным подлежащим и симметричность предиката. *Устный доклад на учебной конференции «Экспериментальные исследования языка», МГУ имени М. В. Ломоносова, Москва, 16 июня 2022 г.* [Pasko L. I. Predicative agreement with compound subject and predicate symmetry. *Talk at the student conference "Experimental language studies", Lomonosov Moscow State Univ., Moscow, June 16, 2022.*]
- Паско 2023 — Паско Л. И. Против АТВ-анализа частичного согласования в русском языке: экспериментальное исследование. *Rhema. Рема*, 2023, 2: 89–103. [Pasko L. I. Against AТВ-analysis of partial agreement in Russian: An experimental study. *Rhema*, 2023, 2: 89–103.]
- Пекелис 2013а — Пекелис О. Е. «Частичное согласование» в конструкции с повторяющимся союзом: корпусное исследование основных закономерностей. *Вопросы языкознания*, 2013, 4: 55–86. [Pekelis O. E. Partial agreement with subjects linked by a correlative conjunction: A corpus-based study of main regularities. *Voprosy Jazykoznanija*, 2013, 4: 55–86.]
- Пекелис 2013б — Пекелис О. Е. Сочинение. *Материалы для проекта корпусного описания русской грамматики*. М., 2013. Рук. [Pekelis O. E. Coordination. *Materials for the project of corpus description of Russian grammar*. Moscow, 2013. Ms.] <http://rusgram.ru/Сочинение>.
- РГ 1980 — Шведова Н. Ю. (ред.). *Грамматика русского языка*. Т. 2. М.: Наука, 1980. [Shvedova N. Yu. (ed.). *Grammatika russkogo yazyka* [Russian grammar]. Vol. 2. Moscow: Nauka, 1980.]
- Розенталь и др. 1999 — Розенталь Д. Э., Джанджакова Е. В., Кабанова Н. П. *Справочник по правописанию, произношению, литературному редактированию*. М.: ЧеРо, 1999. [Rozenal D. E., Dzhandzakova E. V., Kabanova N. P. *Spravochnik po pravopisaniiyu, proiznosheniiyu, literaturnomu redaktirovaniyu* [Handbook of spelling, pronunciation, literary editing]. Moscow: CheRo, 1999.]
- Санников 2008 — Санников В. З. *Русский синтаксис в семантико-прагматическом пространстве*. М.: Языки славянских культур, 2008. [Sannikov V. Z. *Russkii sintaksis v semantiko-pragmaticheskom prostranstve* [Russian syntax in a semantic-pragmatic aspect]. Moscow: Yazyki slavyanskikh kul'tur, 2008.]
- Студеникина 2023а — Студеникина К. А. Об идентичности морфологических признаков при эллипсисе в русском языке: данные именных групп с сочиненными прилагательными. *Rhema. Рема*, 2023, 2: 28–52. [Studenikina K. A. Towards the feature identity for ellipsis in Russian: Evidence from noun phrases with coordinated adjectives. *Rhema*, 2023, 2: 28–52.]
- Студеникина 2023б — Студеникина К. А. Влияние одушевленности конъюнктов и линейной позиции сказуемого на выбор стратегии предикативного согласования. *Устный доклад на учебной конференции «Экспериментальные исследования языка», МГУ имени М. В. Ломоносова, Москва, 16 июня 2022 г.* [Studenikina K. A. The role of animacy of conjuncts and predicate linear position in the choice of predicate agreement strategy. *Talk at the student conference "Experimental language studies", Lomonosov Moscow State Univ., Moscow, June 16, 2022.*]
- Adger 2003 — Adger D. *Core syntax: A minimalist approach*. Oxford: Oxford Univ. Press, 2003.
- Arsenijević, Mitić 2016 — Arsenijević B., Mitić I. Effect of animacy and agentivity on the processing of agreement in Serbo-Croatian. *Studies in languages and mind. Selected papers from the 3rd Workshop in Psycholinguistics, Neurolinguistic and Clinical Linguistic Research*. Halupka-Rešetar S., Martínez-Ferreiro S. (eds.). Novi Sad: Univ. of Novi Sad, 2016, 41–77.
- An, Abeillé 2019 — An A., Abeillé A. Number agreement in French binomials. *Empirical Issues in Syntax and Semantics*, 2019, 12: 31–60.
- Chomsky 1956 — Chomsky N. Three models for the description of language. *IRE Transactions on Information Theory*, 1956, 2(3): 113–124.
- Chomsky 1995 — Chomsky N. *The Minimalist Program*. Cambridge (MA): MIT Press, 1995.
- Chomsky 2014 — Chomsky N. *Aspects of the theory of syntax*. Cambridge (MA): MIT Press, 2014.
- Corbett 1983 — Corbett G. Resolution rules: Agreement in person, number, and gender. *Order, concord and constituency*. Gazdar G., Klein E., Pullam G. K. (eds.). Berlin, Boston: De Gruyter Mouton, 1983, 175–205.
- Devlin et al. 2019 — Devlin J., Chang M. W., Lee K., Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding. *Proc. of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2019, 1: 4181–4186.

- Himmelreich, Hartmann 2023 — Himmelreich A., Hartmann K. Agreement with disjoined subjects in German. *Glossa: a journal of general linguistics*, 2023, 8(1): 1–44.
- Lau et al. 2020 — Lau J. H., Armendariz C., Lappin S., Purver M., Shu C. How furiously can colorless green ideas sleep? Sentence acceptability in context. *Transactions of the Association for Computational Linguistics*, 2020, 8: 296–310.
- Lau et al. 2017 — Lau J. H., Clark A., Lappin S. Grammaticality, acceptability, and probability: A probabilistic view of linguistic knowledge. *Cognitive Science*, 2017, 41(5): 1202–1241.
- Likert 1932 — Likert R. A technique for the measurement of attitudes. *Archives of Psychology*, 1932, 22(140): 5–55.
- Marušič et al. 2015 — Marušič F., Nevins A. I., Badecker W. The grammars of conjunction agreement in Slovenian. *Syntax*, 2015, 18(1): 39–77.
- Mikhailov et al. 2022 — Mikhailov V., Shamardina T., Ryabinin M., Pestova A., Smurov I., Artemova E. RuCoLA: Russian Corpus of Linguistic Acceptability. *Proc. of the 2022 Conference on Empirical Methods in Natural Language Processing*. Goldberg Y., Kozareva Z., Zhang Y. (eds.). Association for Computational Linguistics, 2022, 5207–5227.
- Pereira 2000 — Pereira F. Formal grammar and information theory: together again? *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 2000, 358(1769): 1239–1253.
- Radford et al. 2018 — Radford A., Narasimhan K., Salimans T., Sutskever I. *Improving language understanding by generative pre-training*. Technical Report, OpenAI, 2018.
- Sprouse 2015 — Sprouse J. Three open questions in experimental syntax. *Linguistics Vanguard*, 2015, 1(1): 89–100.
- Sprouse et al. 2013 — Sprouse J., Schütze C. T., Almeida D. A comparison of informal and formal acceptability judgments using a random sample from Linguistic Inquiry 2001–2010. *Lingua*, 2013, 134: 219–248.
- Sprouse et al. 2018 — Sprouse J., Yankama B., Indurkha S., Fong S., Berwick R. C. Colorless green ideas do sleep furiously: gradient acceptability and the nature of the grammar. *The Linguistic Review*, 2018, 35(3): 575–599.
- Sprouse, Schütze 2014 — Schütze C., Sprouse, J. Judgment data. *Research methods in linguistics*. Sharma D., Podesva R. (eds.). Cambridge: Cambridge Univ. Press, 2014, 27–50.
- Studenikina 2023 — Studenikina K. A. Parametrizing number variation in Russian noun phrases with experimental studies and language modeling. *Research Result. Theoretical and Applied Linguistics*, 2023, 9(1): 192–205.
- Vaswani et al. 2017 — Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser L., Polosukhin I. Attention is all you need. *Advances in Neural Information Processing Systems*, 2017, 30: 5998–6008.
- Warstadt et al. 2020 — Warstadt A., Parrish A., Liu H., Mohananey A., Peng W., Wang S. F., Bowman S. R. BLiMP: The benchmark of linguistic minimal pairs for English. *Transactions of the Association for Computational Linguistics*, 2020, 8: 377–392.
- Zwicky 1977 — Zwicky A. Hierarchies of person. *Papers from the 13th Regional Meeting, Chicago Linguistics Society*. Beach W. A., Fox S. E., Philosoph Sh. (eds.). Chicago: Chicago Linguistics Society, 1977, 714–733.