

СТАТЬИ

ЯЗЫКОВАЯ КОМПЕТЕНЦИЯ СКВОЗЬ ПРИЗМУ ГРАММАТИЧЕСКОЙ ВАРИАТИВНОСТИ. ЧАСТЬ 2. КОЛИЧЕСТВЕННАЯ ОЦЕНКА

А.А. Герасимова, Е.А. Лютикова, Л.И. Паско

*Московский государственный университет им. М.В. Ломоносова, Москва,
Россия; anastasiagerasimova432@gmail.com, lyutikova2008@gmail.com,
paskolada@yandex.ru*

Аннотация: Статья посвящена исследованию особенностей распределения оценок приемлемости в суждениях носителей языка о вариативных грамматических феноменах, а именно вопросу о том, насколько согласованы между собой оценки, полученные в контексте синтаксического эксперимента. Предпринимается количественный анализ экспериментальных данных о четырех феноменах вариативности, которые отличаются ожидаемым распределением вариантов и типом грамматического взаимодействия. Количественная оценка проводится по двум параметрам: анализируются единодушие респондентов в языковой общности и индивидуальная последовательность респондентов.

Анализ экспериментальных данных показывает, что доступность в грамматике языка одного или нескольких вариантов определенной конструкции никак не коррелирует с параметрами согласованности. При этом фактическое распределение грамматических профилей носителей языка оказывается много сложнее, чем может следовать из обобщений, сделанных для всей совокупности участников эксперимента. Мы заключаем, что предложенные параметры согласованности играют не менее важную роль в распределении суждений о приемлемости, чем грамматические предикторы, а следовательно, они необходимы для моделирования языковой компетенции как в лингвистических теориях, так и в нейросетевых архитектурах. Выявленные характеристики приемлемости вариативных феноменов также позволят задать точку отсчета при анализе языковой компетенции нейросети и ее сравнении с языковой компетенцией человека.

Ключевые слова: грамматичность; суждения о приемлемости; согласованность; языковая компетенция; вариативность; нейронные сети; русский язык

Финансирование: Работа выполнена при поддержке Программы развития МГУ, проект № 23-Ш02-10 «Языковая компетенция носителей естественного языка и нейросетевых моделей».

doi: 10.55959/MSU0130-0075-9-2024-47-05-9



Для цитирования: Герасимова А.А., Лютикова Е.А., Паско Л.И. Языковая компетенция сквозь призму грамматической вариативности. Часть 2. Количественная оценка // Вестн. Моск. ун-та. Серия 9. Филология. 2024. № 5. С. 111–128.

LINGUISTIC COMPETENCE THROUGH THE LENS OF GRAMMATICAL VARIATION. PART 2. QUANTITATIVE EVALUATION

A.A. Gerasimova, E.A. Lyutikova, L.I. Pasko

*Lomonosov Moscow State University, Moscow, Russia;
anastasiagerasimova432@gmail.com, lyutikova2008@gmail.com,
paskolada@yandex.ru*

Abstract: This paper discusses the distribution of acceptability ratings in native speakers' judgments about grammatical phenomena displaying variability with particular focus on the consistency of the data obtained in the context of a syntactic experiment. We conduct a quantitative analysis of experimental data on four phenomena that vary with respect to the expected distribution of variants and the type of grammatical interaction. The quantitative assessment includes two parameters: the agreement of respondents within the linguistic community and the individual consistency within an experimental trial.

The analysis shows that the availability of one or more variants of a certain construction in grammar does not correlate in any way with the consistency patterns. At the same time, the actual distribution of grammatical profiles of native speakers turns out to be much more complicated than it follows from the generalizations made for the entire population of participants in the experiment. We conclude that the proposed consistency parameters play no less important role in the distribution of judgments than grammatical predictors, and therefore are necessary for modeling linguistic competence both in linguistic theories and in neural network architectures. Finally, the identified characteristics of acceptability in the context of variation provide a baseline for the analysis of the linguistic competence of neural networks and its comparison with the linguistic competence of a human.

Keywords: grammaticality; acceptability judgments; consistency; language competence; variation; neural networks; Russian

Funding: This work was carried out with the support of MSU Program of Development, Project No 23-SCH02-10 "Linguistic Competence of Natural Language Speakers and Neural Network Models".

For citation: Gerasimova A.A., Lyutikova E.A., Pasko L.I. (2024) Linguistic Competence Through the Lens of Grammatical Variation. Part 2. Quantitative Evaluation. *Lomonosov Philology Journal. Series 9. Philology*, no. 5, pp. 111–128.

1. Введение

Статья продолжает исследование основных характеристик приемлемости вариативных феноменов, которые могут использоваться в качестве точки отсчета в задаче сравнения языковых компетенций человека и нейросети. Первая часть работы [Герасимова, Лютикова, Паско 2024] обсуждает исходные теоретические положения исследования, в частности использование оценок приемлемости в качестве диагностик языковой компетенции, предлагает возможные количественные метрики, характеризующие согласованность респондентов в контексте синтаксического эксперимента, а также намечает типологию вариативных феноменов для задачи определения диапазона суждений, допустимых для человека.

При исследовании допустимых моделей согласованности предлагается использовать грамматические феномены, которые различаются уровнем вариативности: от однозначного выбора одного варианта в определенном грамматическом контексте до равной доступности пары грамматических вариантов. При изучении суждений носителей языка о вариативных феноменах мы предложили отказаться от абсолютных характеристик суждений, например, разброса оценок по шкале, средних и медианных показателей, поскольку они могут определяться частотными и признаковыми характеристиками конкретного языкового явления. Вместо этого было предложено оценить сам процесс выноса суждений носителями языка и исследовать параметр согласованности суждений о приемлемости, полученных в экспериментальном контексте (см. [Лютикова и др. 2022; Герасимова 2023: 2.4.4]).

Хотя модель грамматики предопределяет наши ожидания относительно единообразного языкового поведения внутри языкового сообщества, ситуации внутриязыковой вариативности допускают подвижность суждений носителей языка. Именно такого проявления языковой компетенции мы в идеале ожидаем и от нейросети: если языковой феномен характеризуется противоречивыми суждениями о приемлемости со стороны носителей языка, то и в случае нейросетей мы вправе ожидать несопадающих суждений при различных итерациях запуска. Действительно, поскольку нейросетевая модель обучается на массиве текстов, в котором суммируются грамматические профили различных носителей и групп носителей, в поведении нейросети мы должны увидеть соответствующий разброс оценок или выбор вариантов. Отсюда следует возможная методика определения согласованности языковых компетенций человека и нейросети. Можно считать, что языковая компетенция нейросетевой модели отличается от таковой у носи-

телей языка, если суждения нейросети имеют качественно другие характеристики по сравнению с суждениями носителей языка: например, не отражают влияние факторов, предопределяющих вариативность, демонстрируют больший разброс в оценках, чем свойственно феномену в действительности, или напротив, предсказывают бинарное распределение оценок в противоположность градуальности суждений.

Итак, наша цель — изучить возможные распределения оценок приемлемости в суждениях носителей языка о вариативных грамматических феноменах. В данной статье мы проанализируем предварительно собранные экспериментальные данные о приемлемости с точки зрения согласованности суждений носителей языка. Будут проанализированы два аспекта согласованности: единодушные респондентов в языковой общности и индивидуальная последовательность отдельных носителей языка. Мы покажем, что вариативные феномены отличаются не только соотношением приемлемости доступных вариантов, но и степенью уверенности респондентов в приемлемости альтернатив. И если вывод первой характеристики в языковой модели возможен посредством оценки частотности одного из вариантов, то поиск коррелятов индивидуальной последовательности представляет собой отдельную нетривиальную задачу.

2. Рассматриваемые феномены

Для количественного анализа мы привлекли экспериментальные данные о четырех феноменах, ранее представленных в [Герасимова, Лютикова, Паско 2024]. В зависимости от уровня вариативности феномены разделяются на две группы: с высоким и низким уровнем варьирования. Из каждого эксперимента для анализа согласованности мы отобрали данные о паре экспериментальных условий. Обсудим подробнее соответствующие стимулы.

В качестве первого феномена, для которого характерна низкая вариативность, мы привлекаем **падежное варьирование в русских событийных номинализациях**. В экспериментальном исследовании [Pereltsvaig et al. 2018] противопоставляются генитив и инструменталис внешнего аргумента номинализации, при этом только генитивная форма считается грамматичной:

- (1) а. *Это ювелирное изделие есть не что иное как подражание дилетанта/дилетантом мастеру.*
- б. *Важным инструментом развития личности становится мечтание подростка/подростком о будущем.*

Далее мы обращаемся к **предикативному согласованию с сочиненным подлежащим, определяемым относительной клаузой**, которое предполагает однозначный выбор стратегии согласования в пользу множественного числа. Соответственно, в экспериментальном исследовании [Pasko 2023] были противопоставлены условия с согласованием матричного предиката по признакам единственности и множественного числа:

- (2) а. *У беседки цветет/цветут роза и подсолнух, которые Аня поливала по вечерам.*
- б. *На плите жарится/жарятся лук и морковка, которые Наташа принесла с рынка.*

К феноменам с высоким уровнем вариативности мы относим **предикативное согласование в относительных клаузах, определяющих личное местоимение**, где относительное местоимение *который* находится в позиции подлежащего. Из эксперимента [Лютикова 2023] мы берем для анализа стимулы с согласованием предиката вложенной клаузы по первому и третьему лицу:

- (3) а. *Естественно, что мы, которые постоянно работаем/работают по ночам, хотим хорошо выспаться днем.*
- б. *Удивительно, что я, который отлично разбираюсь/разбирается в технике, не могу починить стиральную машину.*

Также высокая вариативность характерна для **отрицательных конструкций с инфинитивными оборотами, содержащими отрицательное местоимение**, которые были экспериментально исследованы в работе [Lyutikova, Gerasimova 2023]. В случае данного феномена мы отдельно рассматриваем варьирование в контексте равносубъектных и разносубъектных инфинитивов. В ситуации равносубъектных инфинитивов нас интересуют экспериментальные условия, в которых противопоставляются следующие две позиции отрицательного местоимения: перед матричным предикатом и базовая позиция (4). Разносубъектные инфинитивы допускают положение отрицательного местоимения в позиции перед вложенным инфинитивом и в базовой позиции, поэтому будут анализироваться соответствующие стимулы (5).

- (4) а. *Журналист (никому) не пробовал возражать (никому) в комментариях к статье.*
- б. *Консультант (никому) не пробовал помогать (никому) в развитии бизнес-проекта.*
- (5) а. *Чиновник не просил сотрудника (никому) отказывать (никому) в бюджетной поддержке.*
- б. *Председатель не вынуждал ассистента (никому) аплодировать (никому) на партийном съезде.*

Основные характеристики экспериментов, результаты которых были использованы для анализа согласованности суждений о приемлемости представлены в табл. 1. В следующем разделе мы обсудим результаты применения метрик согласованности к представленному пулу вариативных феноменов и установим, насколько деление феноменов на две группы по уровню вариативности подтверждается контрольными характеристиками приемлемости.

Таблица 1

Характеристики экспериментов

Феномен (группа феноменов)	Противопоставляемые условия	Метод	Количество участников эксперимента	Количество наблюдений на условии от одного респондента	Публикация результатов
Падежное варьирование в русских событийных номинализациях (низкая вариативность)	Генитив vs инструменталис внешнего аргумента номинализации	Шкала Ликерта 1–5	73	6	[Pereltsvaig et al. 2018]
Согласование с сочиненным подлежащим с определительным придаточным (низкая вариативность)	Единственное vs множественное число предиката матричной клаузы	Шкала Ликерта 1–7	84	4	[Паско 2023]
Согласование в относительной клаузе с вершиной — личным местоимением 1-го лица (высокая вариативность)	Первое vs третье лицо предиката вложенной клаузы	Шкала Ликерта 1–7	76	5	—
Позиция отрицательного местоимения в конструкциях с равносубъектным инфинитивным оборотом (высокая вариативность)	Позиция перед матричным предикатом vs базовая позиция отрицательного местоимения	Шкала Ликерта 1–7	124	4	[Lyutikova, Gerasimova 2023]
Позиция отрицательного местоимения в конструкциях с разносубъектным инфинитивным оборотом (высокая вариативность)	Позиция перед вложенным инфинитивом vs базовая позиция отрицательного местоимения	Шкала Ликерта 1–7	124	4	[Lyutikova, Gerasimova 2023]

3. Анализ согласованности суждений

С точки зрения внешней согласованности наши ожидания о распределении вариантов подтвердились только частично (табл. 2). В случае событийных номинализаций действительно большая часть респондентов предпочитала генитив инструменталису. Однако для феномена с частичным согласованием большая часть респондентов не делали различий между конкурирующими вариантами. Для феноменов с высокой вариативностью, напротив, выделились группы респондентов, предпочитающих только один вариант из представленных, причем группы неслучайны: они составляют не менее 10 % от общей выборки.

Таблица 2

Анализ внешней согласованности респондентов

Феномен	Условие 1 > Условие 2	Условие 2 > Условие 1	Нет различий между условиями
Падежное варьирование в русских событийных номинализациях	GEN > INSTR	INSTR > GEN	GEN = INSTR
	66 (90,4 %)	0	7 (9,6 %)
Согласование с сочиненным подлежащим с определительным придаточным	SG > PL	PL > SG	SG = PL
	21 (25 %)	8 (9,5 %)	55 (65,5 %)
Согласование в относительной клаузе с вершиной — личным местоимением	1P > 3P	3P > 1P	1P = 3P
	8 (10,5 %)	13 (17,1 %)	55 (72,4 %)
Позиция отрицательного местоимения в конструкциях с равносубъектным инфинитивным оборотом	базовая > производная	производная > базовая	базовая = производная
	15 (11,9 %)	54 (42,9 %)	57 (45,2 %)
Позиция отрицательного местоимения в конструкциях с разносубъектным инфинитивным оборотом	базовая > производная	производная > базовая	базовая = производная
	18 (14,5 %)	25 (20,2 %)	81 (65,3 %)

При анализе последовательности респондентов мы будем выделять три уровня: высокий — при размере среднего стандартного отклонения до 1, средний — при попадании показателя в интервал от 1 до 2 баллов, и низкий — при среднем стандартном отклонении выше 2 баллов. Каждый феномен мы охарактеризуем с точки зрения того, какие уровни последовательности показывают респонденты: будем считать уровень действительным для феномена, если его показывают не менее 15 % от всей выборки. Как видно из табл. 3, исследуемые феномены характеризуются различными уровнями последователь-

ности. Высокую последовательность мы наблюдаем для падежного варьирования в номинализациях. Далее для согласовательных феноменов носители демонстрируют высокий и средний уровень последовательности. Наконец, в случае феноменов, связанных с порядком отрицательных местоимений, мы видим полностью задействованный диапазон варьирования с пиком на среднем уровне.

Таблица 3

Анализ последовательности респондентов

Феномен	Высокая	Средняя	Низкая
Падежное варьирование в русских событийных номинализациях	69 (94,5 %)	4 (5,5 %)	0 (0 %)
Согласование с сочиненным подлежащим с определительным придаточным	36 (42,9 %)	40 (47,6 %)	8 (9,5 %)
Согласование в относительной клаузе с вершиной — личным местоимением 1-го лица	27 (35,5 %)	47 (61,8 %)	2 (2,6 %)
Позиция отрицательного местоимения в конструкциях с равносубъектным инфинитивным оборотом	25 (19,8 %)	75 (59,5 %)	26 (20,6 %)
Позиция отрицательного местоимения в конструкциях с разносубъектным инфинитивным оборотом	31 (25 %)	72 (58,1 %)	21 (16,9 %)

В табл. 4 приведено распределение респондентов по последовательности в привязке к группам с точки зрения единодушия в предпочтениях. Наибольший интерес в этом отношении вызывают те респонденты, которые показывают низкий уровень последовательности. Как в случаях согласовательной вариативности, так и в конструкциях с инфинитивным оборотом большая часть таких респондентов не делает различий между альтернативными вариантами.

Таблица 4

Комбинированный анализ единодушия и индивидуальной последовательности респондентов

Последовательность	Условие 1 > Условие 2	Условие 2 > Условие 1	Нет различий между условиями
Падежное варьирование в русских событийных номинализациях	GEN > INSTR	INSTR > GEN	GEN = INSTR
высокая	65 (89%)	0	4 (5,5 %)
средняя	1 (1,4%)	0	3 (4,1 %)
низкая	0	0	0

Последовательность	Условие 1 > Условие 2	Условие 2 > Условие 1	Нет различий между условиями
Согласование с сочиненным подлежащим с определительным придаточным	SG > PL	PL > SG	SG = PL
высокая	8 (9,5%)	6 (7,1%)	22 (26,2%)
средняя	13 (15,5%)	2 (2,4%)	25 (29,8%)
низкая	0	0	8 (9,5%)
Согласование в относительной клаузе с вершиной — личным местоимением	1P > 3P	3P > 1P	1P = 3P
высокая	2 (2,6%)	4 (5,3%)	21 (27,6%)
средняя	6 (7,9%)	9 (11,8%)	32 (42,1%)
низкая	0	0	2 (2,6%)
Позиция отрицательного местоимения в конструкциях с равносубъектным инфинитивным оборотом	базовая > производная	производная > базовая	базовая = производная
высокая	4 (3,2%)	13 (10,5%)	8 (6,5%)
средняя	9 (7,3%)	36 (29%)	30 (24,2%)
низкая	2 (1,6%)	5 (4%)	17 (13,7%)
Позиция отрицательного местоимения в конструкциях с разносубъектным инфинитивным оборотом	базовая > производная	производная > базовая	базовая = производная
высокая	9 (7,3%)	6 (4,8%)	16 (12,9%)
средняя	8 (6,5%)	18 (14,5%)	46 (37,1%)
низкая	1 (0,8%)	1 (0,8%)	19 (15,3%)

Мы также проанализировали последовательность респондентов в отдельных экспериментальных условиях. Оказалось, что для феномена падежной вариативности в номинализациях ответы респондентов при подсчете стандартного отклонения отдельно для каждого условия характеризуются меньшей устойчивостью: последовательность снижается до среднего уровня. Так, в табл. 5 можно увидеть, что респонденты, предпочитающие генитив внешнего аргумента, при оценке генитива могут показывать как высокий, так и средний уровень последовательности (ср. с табл. 4 и соответствующим уровнем последовательности, посчитанным для обоих условий сразу). Для других феноменов подобных изменений не обнаружено. Такой результат для номинализаций объясняется тем, что для части респондентов оценка одного условия менее стабильна по

сравнению с оценкой другого условия, а среднее стандартное отклонение нивелирует подобные различия.

Таблица 5

**Характеристика последовательности
в отдельных экспериментальных условиях**

Падежное варьирование в русских событийных номинализациях	(1) GEN > (2) INSTR		(2) INSTR > (1) GEN		(1) GEN = (2) INSTR	
	1	2	1	2	1	2
высокая	34 (23,3%)	47 (32,2%)	0	0	2 (1,4%)	2 (1,4%)
средняя	32 (21,9%)	19 (13%)	0	0	5 (3,4%)	4 (2,7%)
низкая	0	0	0	0	0	1 (0,7%)

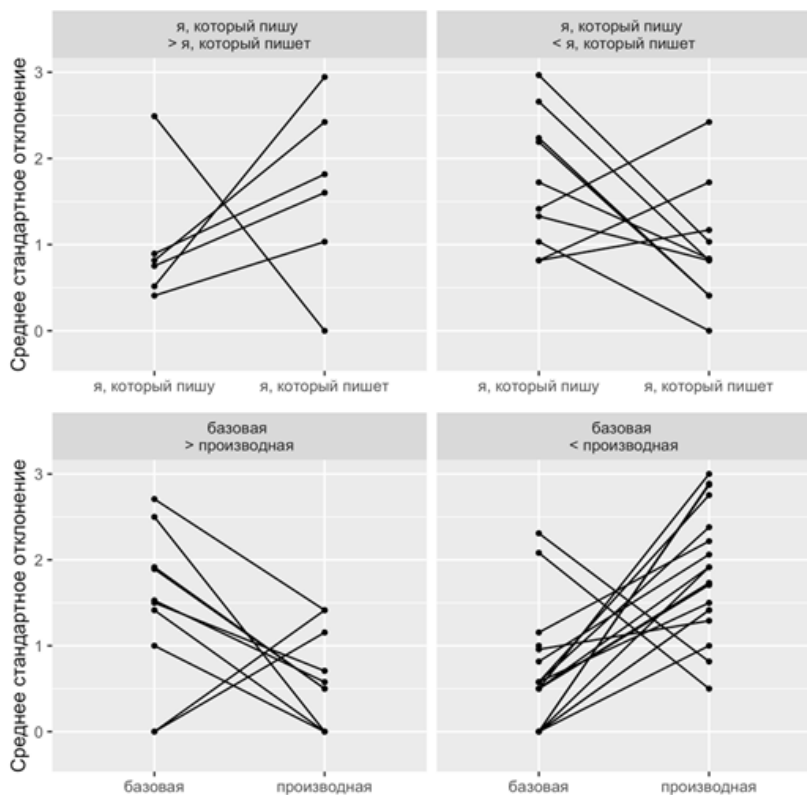


Рисунок 1. Соотношение последовательности респондентов в оценках для альтернативных вариантов. Вверху: согласование в относительной клаузе с вершиной — личным местоимением; внизу: позиция отрицательного местоимения в конструкциях с разнособъектным инфинитивным оборотом

4. Обсуждение

В статье мы ставили цель исследовать диапазон суждений о приемлемости вариативных феноменов с позиций внешней и внутренней согласованности данных. Необходимым предварительным условием для решения задачи по сравнению языковых компетенций человека и нейросети должно быть понимание того, каким в принципе может быть распределение оценок приемлемости. Мы привлекли экспериментальные данные по четырем явлениям вариативности, которые отличаются уровнем варьирования и типом грамматического взаимодействия, и проанализировали согласованность данных.

В результате исследования можно заключить, что четыре рассмотренных феномена демонстрируют широкий диапазон допустимых суждений о приемлемости. Обобщим основные характеристики суждений носителей языка, которые в дальнейшем можно будет использовать в качестве ориентира при анализе приемлемости в нейросетевой модели. Специфичность рассмотренных феноменов проявляется в неодинаковом распределении параметра внешней согласованности: в ситуации внутриязыковой вариативности выделяются группы респондентов с различными грамматическими предпочтениями. При этом для каждого из рассматриваемых феноменов выделяется одна основная группа респондентов, составляющая более 50 % выборки. Примечательно, что в трех случаях мы наблюдаем действительную вариативность: большая часть респондентов не делает различий между условиями. Две другие группы оказываются противопоставлены друг другу и группе большинства.

Соотнесем распределение по внешней согласованности с основными тенденциями выборок. В случае падежного варьирования в событийных номинализациях ответы большей части респондентов соответствуют ожидаемому уровню вариативности и задают значимое различие между сравниваемыми условиями. Для согласования с сочиненным подлежащим с определительным придаточным наблюдается более высокий уровень вариативности, чем предсказывался в теории. Тем не менее, основная тенденция выборки задается группой большинства: различий между приемлемостью согласования по единственному и по множественному числу нет. Аналогично устроена ситуация с приемлемостью в случае согласования в относительной клаузе с вершиной — личным местоимением, хотя здесь наличие групп носителей с различными предпочтениями ожидаемо в отличие от предыдущих феноменов.

В конструкциях с инфинитивными оборотами вариативность также предсказывается на основании структурных обобщений, но

выражается в различном распределении оценок приемлемости. В обоих случаях различие между приемлемостью базовой и производной позиции значимо, и производная позиция оценивается значимо выше, чем базовая. Для конструкций с равносубъектными инфинитивными оборотами основную тенденцию выборки задает одна из групп респондентов с выраженным предпочтением варианта, хотя данная группа соразмерна группе респондентов, не делающих различия между условиями. Примечателен тот факт, что в случае конструкций с разносубъектными инфинитивами группа респондентов, которая последовательно ставит более высокие оценки предложениям с отрицательным местоимением в производной позиции, уступает в размерах группе респондентов, не различающих приемлемость двух альтернатив.

Таким образом, фактическое распределение грамматических профилей носителей языка оказывается много сложнее, чем может следовать из обобщений, сделанных для всей совокупности участников эксперимента. Наше исследование подтверждает обобщения работы [Герасимова, Лютикова 2021] о низком уровне единодушия среди носителей в контексте лингвистического эксперимента, посвященного вариативным феноменам.

Что касается внутренней согласованности, т. е. индивидуальной последовательности, обобщение [Герасимова, Лютикова 2021] о среднем уровне последовательности респондентов и размере среднего стандартного отклонения от 1 до 2 баллов для большей части респондентов подтверждается лишь частично. Контрпримером становится феномен падежной вариативности в номинализациях, для которого даже при подсчетах последовательности отдельно по условиям большинство респондентов демонстрируют высокий уровень последовательности. Причиной тому может быть меньшее количество доступных значений на шкале Ликерта 1–5 по сравнению со шкалой 1–7¹. Однако ввиду того, что другие достоверные данные, которые могли бы подтвердить смещение оценок при изменении размерности шкал, отсутствуют, на текущий момент мы не можем составить однозначное заключение о роли этого фактора.

¹ В [Leung 2011] показано, что психометрические свойства шкал Ликерта с 4, 5, 6 и 11 делениями совпадают, в том числе совпадают основные характеристики распределения — средние и стандартные отклонения, хотя большее количество делений и снижает асимметрию распределения оценок. Учитывая это наблюдение, мы включаем в рассмотрение эксперимент со шкалой Ликерта 1–5. Кроме того, мы предполагаем, что носители пользуются шкалами приемлемости единообразно, ориентируясь главным образом на ярлыки, присвоенные числам, и допускают равную вариативность в оценках для одного условия вне зависимости от количества делений на шкале.

Для всех других рассмотренных феноменов, действительно, большинство участников эксперимента показывает средний уровень последовательности, однако не менее 40 % респондентов при этом характеризуются высокой или низкой последовательностью. Иными словами, мы видим существенную вариативность данного параметра при переходе от феномена к феномену. Теоретически ожидаемый уровень варьирования для грамматического явления также не является предиктором последовательности: из большей вариативности не обязательно следует меньшая уверенность в оценках. Так, для контекста относительной клаузы с вершиной — личным местоимением 1-го лица респонденты совсем не показывают низкого уровня последовательности, хотя в конструкциях с инфинитивными оборотами доля респондентов с низким уровнем последовательности достаточно велика.

Обратимся теперь к вопросу о связи единодушия респондентов с их индивидуальной последовательностью. Важный результат исследования состоит в том, что в группе, не делающей различия между вариантами, наблюдается самый низкий уровень последовательности. Вероятно, уверенность в оценке уменьшается, если в идиолекте носителя доступна альтернатива. При этом, однако, собственно распределение вариантов по приемлемости никак не коррелирует с большей или меньшей последовательностью респондента в оценках. Мы наблюдали как случай большей последовательности для (высоких оценок) более приемлемого варианта, так и большую уверенность в (низких) оценках менее приемлемой альтернативы.

Итак, однозначной корреляции между доступностью в грамматике языка одного или нескольких вариантов определенной конструкции и единодушием и последовательностью респондентов в оценках приемлемости не наблюдается. Феномены показывают различный уровень вариативности не только в отношении предпочтения одного из доступных вариантов, но и в отношении уверенности респондента в собственных предпочтениях.

Закономерно возникает вопрос о том, насколько исследованные параметры согласованности данных о приемлемости значимы для нейросетевых моделей языка. В отношении параметра единодушия мы можем ожидать, что языковые модели столкнутся с трудностями в ситуации вариативности, и это выразится в том, что в модели вероятностное предпочтение будет всегда отдаваться одному из двух возможных вариантов. Поскольку задача предсказания оценок приемлемости с помощью нейронных сетей ранее решалась только в бинарной разметке (см., например, [Warstadt et al. 2019]), доступный, но оцененный как менее вероятный вариант неизбежно получает статус

неграмматичного, что совсем не соответствует действительности. В свете проблемы моделирования грамматики результаты проведенного исследования подтверждают градуальный характер приемлемости и ставят под сомнение бинарно-категориальные архитектуры грамматики, которые не способны уловить сложное распределение приемлемости как с точки зрения единодушия респондентов, так и с точки зрения целостности грамматического представления одного индивида. Таким образом, языковая компетенция нейросетей также должна учитывать градуальность оценок приемлемости.

Отдельно стоит проблема неразличения уровня приемлемости в ситуации вариативности. Нейросетевые решения основываются на предварительном обучении на больших текстовых массивах. Приемлемость в контексте языковой модели при этом можно считать проекцией метрик частотности или вероятности появления определенного варианта в корпусе [Студеникина 2023]. Предположим, что грамматические альтернативы оказываются в равной степени маловероятными. Как показывает проведенный нами анализ оценок приемлемости, за общей тенденцией неразличения вариантов могут стоять разные причины. С одной стороны, возможно, что неразличение альтернативных вариантов свидетельствует о конкуренции групп носителей, имеющих различные грамматические предпочтения, причем основная тенденция выборки далеко не всегда совпадает с грамматическим профилем большей части носителей. С другой стороны, малая встречаемость в корпусе текстов, равно как и низкие или срединные оценки по шкале приемлемости, могут свидетельствовать о периферийном характере конструкции: вероятно, что подобные языковые выражения в целом следует рассматривать в контексте невыразимости (ср. *ineffability* [Pesetsky 1997]). Проблема состоит в том, что на текущем этапе развития больших языковых моделей отличить одну ситуацию от другой вряд ли возможно.

Обратимся теперь к параметру последовательности, который задает больший или меньший разброс оценок по шкале приемлемости. Чем выше последовательность, тем более однозначна оценка. Однако мы видим, что разные феномены характеризуются разными уровнями последовательности, а кроме того, разные уровни последовательности накладываются на разделение носителей по группам предпочтений. Встает вопрос о том, какие требования данный параметр может накладывать на нейросетевые модели. Мы предполагаем, что распределение носителей по индивидуальной последовательности в отношении феномена характеризует его статус через призму грамматических изменений. Поскольку избыточность кода ведет к спецификации отдельных вариантов, в конечном счете ожидается выбор только какой-то одной грамматической стратегии

в конкретном контексте. В то же время в синхронном аспекте последовательность в большей степени характеризует речь конкретного индивида. Нейросетевые архитектуры порождают тексты как некоторый обобщенный говорящий, в речи которого суммируются и ранжируются особенности всех текстов обучающего корпуса. Задача индивидуализации речи и имитации грамматического профиля носителя языка в настоящее время широко не ставится. Тем не менее, для решения такой задачи необходим учет диапазона речевых характеристик носителей языка, которые в том числе выражаются в разных уровнях последовательности.

Коррелятом оценок приемлемости в больших языковых моделях становятся метрики частотности или вероятности, посчитанные по большим корпусам текстов [Lau et al. 2017]. Поскольку и для учета параметра единодушия, и для учета параметра последовательности мы видим необходимость в индивидуализации языковых закономерностей, решением могло бы стать разделение текстов по авторству и отслеживание индивидуальных речевых характеристик. Такое решение могло бы приблизить языковую компетенцию нейросети к языковой компетенции человека.

Завершить статью мы хотели бы следующим вопросом: должна ли нейросеть в принципе рассматривать грамматические варианты как действительные альтернативы и подобно грамматике моделировать факторы, способствующие выбору того или иного варианта? Данный вопрос особенно актуален в свете разворачивающейся дискуссии об объяснительной силе больших языковых моделей, которые противопоставляются достижениям теоретической лингвистики вообще и генеративной грамматики в частности (см., напр., [Piantadosi 2023]). Сторонники апофеоза нейросетей предполагают, что вероятностно-статистический характер языковой модели достаточен сам по себе для моделирования языковой компетенции. Противники расширения технологий работы с текстом и речью до уровня лингвистических теорий исходят из того, что нейросетевые модели только подтверждают существование семантико-синтаксического интерфейса, но при этом не способны предоставить фундаментального понимания структуры языка (см. [Marcolli 2023]).

В связи с этой дискуссией мы можем отметить, что ни нейросетевые решения, ни лингвистические теории в настоящий момент никак не моделируют согласованность представлений о приемлемости в контексте вариативных феноменов. Фактически же параметры внешней и внутренней согласованности играют не менее важную роль в распределении суждений, чем грамматические предикторы, а следовательно, необходимы для моделирования языковой компетенции.

СПИСОК ЛИТЕРАТУРЫ

1. Герасимова А.А., Лютикова Е.А. Лингвистический эксперимент на платформе Яндекс. Толока: оценка исследовательских возможностей // *Zeitschrift für Slavische Philologie*. 2021. Vol. 78. № 1. P. 175–206.
2. Герасимова А.А. Количественные методы исследования грамматических ограничений (на материале вариативного согласования в русском языке): Дисс. ... канд. филол. наук; Московский государственный университет имени М.В. Ломоносова. М., 2023.
3. Герасимова А.А., Лютикова Е.А., Паско Л.И. Языковая компетенция сквозь призму грамматической вариативности. Часть 1. Теоретические и методологические соображения // *Вестник Московского университета. Серия 9. Филология*. 2024. № 4. С. 9–22.
4. Лютикова Е.А., Герасимова А.А., Белова Д.Д., Студеникина К.А., Лютиков А.С. Три аспекта непротиворечивости языковых данных: оценка и интерпретация // *Интеллектуальные системы. Теория и приложения*. 2022. Т. 26. № 1. С. 196–202.
5. Студеникина К.А. Parametrizing Number Variation in Russian Noun Phrases with Experimental Studies and Language Modeling // *Научный результат. Вопросы теоретической и прикладной лингвистики*. 2023. Vol. 9. № 1. P. 192–205.
6. Lau, J.H., Clark A., Lappin S. Grammaticality, acceptability, and probability: A probabilistic view of linguistic knowledge // *Cognitive Science*. 2017. Vol. 41. № 5. P. 1201–1241.
7. Leung S.O. A Comparison of Psychometric Properties and Normality in 4-, 5-, 6-, and 11-point Likert scales // *Journal of Social Service Research*. 2011. Vol. 37, no. 4. P. 412–421.
8. Marcolli M. Public Address on Generative Linguistics. YouTube Lecture, 2023. URL: <https://www.youtube.com/watch?v=-gx3SK7FvKk> (accessed : 23.10.2023).
9. Pesetsky D. Optimality Theory and Syntax: Movement and Pronunciation // *Optimality Theory, an Overview / editors : D. Archangeli, D. T. Langendoen*. Blackwell Publishers, 1997. P. 134–170.
10. Piantadosi S.T. Modern Language Models Refute Chomsky's Approach to Language. Ms., 2023. 45 p. URL: <https://ling.auf.net/lingbuzz/007180>.
11. Warstadt A., Singh A., Bowman S.R. Neural Network Acceptability Judgments // *Transactions of the Association for Computational Linguistics (ACL)*. Cambridge, MA: MIT Press, 2019. P. 625–641.

REFERENCES

1. Gerasimova A.A., Lyutikova E.A. Lingvisticheskii eksperiment na platforme Yandeks. Toloka: otsenka issledovatel'skikh vozmozhnopei [A Linguistic Experiment on the Yandex.Toloka Crowdsourcing Platform: An Evaluation of Its Research Capacity]. *Zeitschrift für Slavische Philologie*. 2021, vol. 78, № 1, pp. 175–206. (In Russ.)
2. Gerasimova A.A. Kolichestvennyye metody issledovaniya grammaticheskikh ogranichenii (na materiale variativnogo soglasovaniya v russkom yazyke) [Quantitative Methods of Investigating Grammar (A Case Study of Agreement Variation in Russian)]. Diss. ... kand. filol. nauk. Lomonosov Moscow State University. Moscow, 2023. (In Russ.)
3. Gerasimova A.A., Lyutikova E.A., Pasko L.I. Yazykovaya kompetentsiya skvoz' prizmu grammaticheskoi variativnosti. Chast' 1. Teoreticheskie i metodologicheskie soobrazheniya [Linguistic Competence Through the Lens of Grammatical Variation. Part 1. Conceptual and Methodological Considerations]. *Lomonosov Philology Journal. Series 9. Philology*, 2024, no. 4, pp. 9–22. (In Russ.)

4. Lau, J.H., Clark A., Lappin S. Grammaticality, Acceptability, and Probability: A Probabilistic View of Linguistic Knowledge. *Cognitive Science*. 2017, vol. 41, no. 5, pp. 1201–1241.
5. Leung S.O. A Comparison of Psychometric Properties and Normality in 4-, 5-, 6-, and 11-point Likert Scales. *Journal of social service research*. 2011, vol. 37, no. 4, pp. 412–421.
6. Lyutikova E.A., Gerasimova A.A., Belova D.D., Studenikina K.A., Lyutikov A.S. Tri aspekta neprotivorechivosti yazykovykh dannyykh: otsenka i interpretatsiya [Three aspects of Linguistic Data Consistency: Estimation and Interpretation]. *Intel'lektual'nye sistemy. Teoriya i prilozheniya* [Intelligent Systems. Theory and Applications]. 2022, vol. 26, no. 1, pp. 196–202. (In Russ.)
7. Marcolli M. Public Address on Generative Linguistics. YouTube Lecture, 2023. URL: <https://www.youtube.com/watch?v=-gx3SK7FvKk> (accessed: 23.10.2023).
8. Pesetsky D. Optimality Theory and Syntax: Movement and Pronunciation. *Optimality Theory, an Overview* / editors : D. Archangeli, D. T. Langendoen. Blackwell Publishers, 1997, pp. 134–170.
9. Piantadosi S.T. Modern Language Models Refute Chomsky's Approach to Language. Ms., 2023. 45 p. URL: <https://ling.auf.net/lingbuzz/007180>.
10. Studenikina K.A. Parametrizing Number Variation in Russian Noun Phrases with Experimental Studies and Language Modeling // *Research result: Theoretical and applied linguistics*. 2023, vol. 9, no. 1. P. 192–205.
11. Warstadt A., Singh A., Bowman S.R. Neural network acceptability judgments. *Transactions of the Association for Computational Linguistics (ACL)*. Cambridge, MA: MIT Press, 2019, pp. 625–641.

Поступила в редакцию 02.11.2023

Принята к публикации 16.04.2024

Отредактирована 20.05.2024

Received 02.11.2023

Accepted 16.04.2024

Revised 20.05.2024

ОБ АВТОРАХ

Герасимова Анастасия Алексеевна — кандидат филологических наук, научный сотрудник Научно-исследовательского вычислительного центра МГУ имени М.В. Ломоносова; anastasiagerasimova432@gmail.com

Лютикова Екатерина Анатольевна — доктор филологических наук; профессор кафедры теоретической и прикладной лингвистики филологического факультета МГУ имени М.В. Ломоносова; ведущий научный сотрудник Научно-исследовательского вычислительного центра МГУ имени М.В. Ломоносова; lyutikova2008@gmail.com

Паско Лада Игоревна — студент магистратуры отделения теоретической и прикладной лингвистики филологического факультета МГУ имени М.В. Ломоносова; paskolada@yandex.ru

ABOUT THE AUTHORS

Anastasia A. Gerasimova — PhD, Researcher, Research Computing Center, Lomonosov Moscow State University; anastasiagerasimova432@gmail.com

Ekaterina A. Lyutikova — Prof. Dr. Habil, Professor, Department of Theoretical and Applied Linguistics, Faculty of Philology, Lomonosov Moscow State University; Leading Researcher, Research Computing Center, Lomonosov Moscow State University; lyutikova2008@gmail.com

Lada I. Pasko — MA student, Department of Theoretical and Applied Linguistics, Faculty of Philology, Lomonosov Moscow State University; paskolada@yandex.ru