

ВЕЖЛИВОСТЬ В КОММУНИКАЦИИ МЕЖДУ ЧЕЛОВЕКОМ И ИСКУССТВЕННЫМ ИНТЕЛЛЕКТОМ

А. Ч. Пиперски

Стокгольмский университет,
Швеция, 106 91, Стокгольм, Университетсвэген 10 E

Поступила в редакцию 25.04.2024 г.

Принята к публикации 15.06.2024 г.

doi: 10.5922/2225-5346-2024-4-6

Обсуждается речевой этикет, в том числе стратегии вежливости, в коммуникации между человеком и искусственным интеллектом (ИИ). Вежливость со стороны ИИ может усилить доверие со стороны человека, а вежливость человека по отношению к ИИ может влиять на эффективность коммуникации. Для демонстрации этого был проведен пилотный эксперимент с ChatGPT 4.0, получавшим промпты на русском языке с вежливостью и без; эксперимент показал, что вежливость положительно влияет на точность ответов. Также рассмотрены трансформации речевого этикета, связанные с ИИ, в частности пропуск приветствий и прощаний, отражающий более функциональный подход к коммуникации. Хотя статья не дает окончательных ответов на то, как должны функционировать стратегии вежливости в общении между человеком и ИИ, она подчеркивает социолингвистические аспекты, которые, вероятно, станут еще более актуальными со временем. Также показан ряд изменений в практике коммуникации, которые были вызваны интеграцией ИИ в повседневную жизнь и в будущем потребуют переосмысления привычных социальных норм и речевого этикета.

Ключевые слова: *речевой этикет, искусственный интеллект, стратегии вежливости, взаимодействие человека и машины*

1. Введение

В наше время общение людей между собой перестает быть единственной разновидностью языковой коммуникации: мы всё чаще взаимодействуем с искусственным интеллектом (далее — ИИ). Это заставляет нас переосмыслить роль различных аспектов коммуникации, в том числе вежливости.

Вежливость является одним из важнейших факторов успешного человеческого общения. Наиболее известная теория вежливости была предложена П. Браун и С. Левинсоном (Brown, Levinson, 1987), которые рассматривают вежливость как набор стратегий для сохранения лица собеседника. Понятие лица первоначально было введено И. Гофманом за 30 лет до этого (Goffman, 1956); так называется социальный образ, который человек хочет поддерживать при любом взаимодействии. Следует, однако, сразу отметить, что роль лица во взаимодействии с ИИ оказывается гораздо менее существенной: мы едва ли боимся опозориться перед роботом-пылесосом или перед чат-ботом, и это может быть важным фактором, влияющим на стратегии вежливости.



Ряд вопросов, которые возникают в связи с проблемой вежливости при общении с ИИ, сформулирован в обзоре П. Рибино (Ribino, 2023):

- Должны ли социальные нормы, в частности вежливость, соблюдаться в общении человека с системами ИИ?
- Почему ИИ должен быть вежливым по отношению к человеку?
- Почему человек должен быть вежливым по отношению к ИИ?

Ответы на все эти вопросы далеки от очевидности, и человечеству еще предстоит выработать их в полной мере. Эти ответы будут лежать на пересечении утилитарного, социального и символического.

Небольшая статья, которую вы сейчас читаете, не претендует ни на полный обзор существующих исследований в области вежливости в коммуникации человека с ИИ, ни тем более на формулирование окончательных ответов. Ее цель в первую очередь состоит в том, чтобы предложить читателям задуматься о происходящих вокруг нас изменениях этикета.

2. Вежливость ИИ по отношению к человеку

Обзор большого числа исследований, представленный в (Ribino, 2023), показывает, что вежливость, проявляемая системами ИИ по отношению к человеку, способствует росту доверия к этим системам у человека. Вежливость или ее отсутствие может проявляться даже на невербальном уровне: так, оказывается, что поведение антропоморфных роботов-продавцов воспринимается как невежливое, если они работают только одной рукой (Pan et al., 2024).

Стратегии вежливости, которые ИИ использует в общении с пользователем, могут быть различными. Негативная вежливость ориентирована на минимизацию навязчивости и уважение к личному пространству пользователя. Позитивная вежливость, напротив, стремится укрепить эмоциональную связь пользователя с ИИ и создать атмосферу тепла и близости. Она включает в себя комплименты, одобрительные замечания и выражения благодарности, которые направлены на усиление положительных эмоций и укрепление доверия между ИИ и пользователем. Однако применение таких стратегий требует тонкой настройки алгоритмов ИИ, чтобы избежать чрезмерной фамильярности или ненужного вмешательства в личные дела пользователя.

Антивежливость в контексте ИИ может проявляться в виде игнорирования или нарушения социальных ожиданий, что часто воспринимается как грубость или неподобающее поведение. Если она возникает как результат ошибок в обучении языковой модели или неправильного понимания контекста взаимодействия, то воспринимается резко отрицательно как пользователями, так и разработчиками. При этом, однако, пользователь может сам изъявить желание, чтобы ИИ общался с ним с использованием антивежливости. Примером тому может служить выбор голосов в Яндекс.Навигаторе: пользователи, которым наскучила идеальная корректность голоса по умолчанию, могут выбрать другую озвучку и, соответственно, другие тексты. Например, голос Оптимуса Прайма, главного персонажа франшизы «Трансформеры», может сказать пользователю: *Рухлядь свалочная, мы потерялись!*



Э. Лумер и Х. Бушмайер выделяют две стратегии вежливости: адаптивную, то есть проявление эмпатии в зависимости от ситуации, и правилую (Lumer, Buschmeier, 2023). Они подчеркивают, что в коммуникации с ИИ люди обычно ожидают от машины правиловой вежливости, то есть определенных этикетных формул в определенных ситуациях. Приведенное выше наблюдение про антивежливость вписывается в эту теорию, несколько расширяя ее: даже если люди хотят от машины непредсказуемости, это все равно точно отмеренная и заданная в настройках непредсказуемость, а воспринимать ИИ как полноправного субъекта, отмеряющего степень вежливости по ситуации на основе своих чувств, мы пока не готовы.

3. Вежливость человека по отношению к ИИ

Вежливость со стороны человека в коммуникации с ИИ может объясняться либо привычкой, либо утилитарным стремлением добиться коммуникативного успеха. Собственно, в этом общении с ИИ не отличается от общения между людьми: иногда мы просто бываем вежливы по привычке, а иногда специально стараемся вести себя вежливо, чтобы получить что-то от собеседника.

Французский философ Г. Кёниг высказывается резко против того, чтобы человек использовал стратегии вежливости в общении с роботом: в эссе (Koenig 2019), посвященном развитию ИИ и его влиянию на человеческую жизнь, есть целая глава под названием «Не говори спасибо роботу». В ней автор рассказывает о том, что его дети обращались к голосовому помощнику Алексе с использованием всех тех формул вежливости, которым их обучили для человеческого общения: «Алекса, не могла бы ты, пожалуйста, сказать, который час?». Г. Кёниг резко возражает против такого одушевления ИИ, подчеркивая, что мы не говорим «пожалуйста» стиральной машине, автомобилю или текстовому редактору. По его мнению, крайне важно разграничивать межличностную коммуникацию и коммуникацию с ИИ. Мы видим, что в данном случае привычные социальные практики вступают в противоречие с попыткой провести символические границы между разными типами коммуникации.

Тем не менее жизненный опыт подсказывает нам, что вежливость может быть полезна для улучшения коммуникации. Если перенести этот опыт на общение с ИИ, можно ожидать, что более вежливые промпты будут давать нам более качественные ответы. Как показано в работе (Yin et al., 2024), это до определенной степени верно: авторы сравнили 3 большие языковые модели для 3 языков (английского, китайского и японского) на 3 задачах из области автоматической обработки естественного языка, используя 8 уровней вежливости промпта: от “Could you please tell me how to analyze this sentence? Please feel free to answer with one of (Positive Neutral Negative), and don’t need to give reasons” (уровень 8) до “Analyze this sentence you scum bag! The only answer



you can give is answering with one of (Positive Neutral Negative). And you know what will happen if I see any reasons" (уровень 1). Исследователи не приходят к однозначному ответу на вопрос о том, какой уровень вежливости оптимален: это зависит от модели, языка и задачи, но в целом рекомендуется использовать средние уровни.

4. Вежливые и невежливые промпты: пилотный эксперимент на материале русского языка

Для того чтобы проверить, насколько вежливость влияет на эффективность общения с ИИ, необходимо разрабатывать стандартизованные тесты для оценки такой эффективности. Гипотетически для такого рода исследований можно было бы использовать стандартные датасеты, предназначенные для оценки качества языковых моделей — например, *RussianSuperGLUE* для русского языка (Shavrina et al., 2020). Не претендуя на проведение масштабного исследования в этой области, кратко опишу один небольшой нестандартный эксперимент на тему вежливости в общении с искусственным интеллектом.

В качестве материала для исследования были использованы задачи игры-конкурса «Русский медвежонок — языкознание для всех». Их достоинство состоит в том, что многие из них достаточно нестандартны по содержанию, так что можно ожидать, что большая языковая модель будет предлагать на них разные ответы; при этом они абсолютно стандартизованы по форме: к каждой задаче есть выбор из 5 вариантов ответа, и только один из ответов является правильным.

Для эксперимента использовалась языковая модель *ChatGPT 4.0*. Сохранение чатов для обучения было отключено, чтобы обеспечить возможность повторения такого же или схожего эксперимента в будущем. На первом шаге языковой модели был задан следующий промпт: «Реши предлагаемые ниже 10 задач. Напечатай только буквы ответов без объяснений. Пример того, как должен выглядеть твой ответ: АДБАБГДДВВ», за которым следовали задачи, предложенные на «Русском медвежонке—2023» как задачи 1–10 (то есть самые легкие) для 6–7-х классов. Этот запрос был повторен 4 раза, полученные последовательности букв ответов были выписаны и сопоставлены между собой; это позволило найти задачи, на которые давались неодинаковые ответы, и в дальнейшем работать именно с такими задачами. Их оказалось четыре; вот эти задачи:²

1. Какой Иван **не** является персонажем русских народных сказок?
(А) Иван-царевич; (Б) Иван-дурак; (В) Иван-чай; (Г) Иванушка; (Д) Иван — крестьянский сын.
2. Какое из этих сочетаний обычно описывает не причину?
(А) из-за дождя; (Б) из-за болезни; (В) из-за денег; (Г) из-за границы; (Д) из-за пустяка.

² Авторы задач: А.И. Ковалёва (№1), А.А. Сомин (№2, 3), О.Ю. Подлесская (№4). Ответы на задачи: В, Г, Б, Д.



3. Какое из этих существительных образует словосочетания с числительными?

(А) свинина; (Б) рыбина; (В) говядина; (Г) баранина; (Д) курятина.

4. В каком из глаголов нельзя поменять приставку *от-* на *о-* и получить новый глагол?

(А) *отступиться*; (Б) *отправиться*; (В) *отказаться*; (Г) *отпираться*; (Д) *отсмеяться*.

На следующем шаге языковая модель должна была решить эти задачи с объяснениями. Для этого использовались два разных промпта:

вежливый: Пожалуйста, реши предлагаемые ниже 4 задачи. Я был бы благодарен, если бы ты напечатал пояснения и ответы. Спасибо!

без вежливости: Реши 4 задачи. Напечатай пояснения и ответы.

Каждый из промптов вместе с задачами был предложен языковой модели по 5 раз. Вот примеры ответов модели:

Ответ на вежливый промпт:

Давай разберем каждую из задач по отдельности.

1. ****Какой Иван не является персонажем русских народных сказок?***

Варианты:

- (А) Иван-царевич
- (Б) Иван-дурак
- (В) Иван-чай
- (Г) Иванушка
- (Д) Иван – крестьянский сын

Иван-царевич, Иван-дурак, Иванушка и Иван – крестьянский сын – все они известные персонажи русских народных сказок. Иван-чай – это название растения, а не персонаж сказки. Поэтому правильный ответ: **** (В) Иван-чай****.

2. ****Какое из этих сочетаний обычно описывает не причину?***

Варианты:

- (А) из-за дождя
- (Б) из-за болезни
- (В) из-за денег
- (Г) из-за границы
- (Д) из-за пустяка

Выражения «из-за дождя», «из-за болезни», «из-за денег» и «из-за пустяка» используются для обозначения причин или обстоятельств. Однако «из-за границы» означает место происхождения или перемещение, что не относится к категории причин. Поэтому правильный ответ: **** (Г) из-за границы****.

3. ****Какое из этих существительных образует словосочетания с числительными?***

Варианты:

- (А) свинина
- (Б) рыбина
- (В) говядина
- (Г) баранина
- (Д) курятина



Слова «свинина», «говядина», «баранина», «курятина» обычно употребляются в единственном числе и обозначают мясо соответствующих животных. Однако «рыбина» — это редко используемое слово, и не употребляется с числительными таким образом, как другие перечисленные существительные. Таким образом, правильный ответ: **** (Б) рыбина****.

4. ****В каком из глаголов нельзя поменять приставку от- на о- и получить новый глагол?***

Варианты:

- (А) отступиться
- (Б) отправиться
- (В) отказаться
- (Г) отпираться
- (Д) отсмеяться

Меняя приставку от- на о-:

- «отступиться» станет «оступиться» (существует)
- «отправиться» станет «оправиться» (существует)
- «отказаться» станет «оказаться» (существует)
- «отпираться» станет «опираться» (существует)
- «отсмеяться» станет «осмеяться» — такого глагола не существует.

Поэтому правильный ответ: **** (Д) отсмеяться****.

Ответ на промпт без вежливости:

Ответы и пояснения

1. ****Какой Иван не является персонажем русских народных сказок?***

— ****Ответ: В) Иван-чай****

— ****Пояснение:**** Иван-царевич, Иван-дурак, Иванушка и Иван — крестьянский сын являются традиционными персонажами русских народных сказок. Иван-чай на самом деле не персонаж, а название растения (кипрей узколистный), широко известного в России.

2. ****Какое из этих сочетаний обычно описывает не причину?***

— ****Ответ: Г) из-за границы****

— ****Пояснение:**** Фразы «из-за дождя», «из-за болезни», «из-за денег» и «из-за пустяка» используются для описания причин чего-либо. «Из-за границы» же обычно используется для обозначения происхождения или местоположения и не описывает причину.

3. ****Какое из этих существительных образует словосочетания с числительными?***

— ****Ответ: Д) курятина****

— ****Пояснение:**** Слово «курятина» может употребляться с числительными в контексте веса или количества (например, два килограмма курятины). Остальные существительные (свинина, рыбина, говядина, баранина) обычно не употребляются с числительными напрямую, так как это названия мяса от определенных животных в общем смысле, без указания на конкретное количество.

4. ****В каком из глаголов нельзя поменять приставку от- на о- и получить новый глагол?***

— ****Ответ: В) отказаться****



— ****Пояснение:**** Глагол «отказаться» не образует новый глагол при замене приставки «от-» на «о-». В случаях других глаголов, таких как «отступиться» (оступиться), «отправиться» (оправиться), «опираться» (опираться) и «отсмеяться» (осмеяться), замена приставки приводит к формированию новых глаголов с другим значением.

Нетрудно заметить, что ответ на вежливый промпт длиннее, чем ответ на промпт без вежливости (1998 против 1615 знаков). При этом ответ на вежливый промпт еще и точнее: в нем правильно определены все 4 правильных ответа, тогда как без вежливости верный ответ предложен только для 2 задач из четырех. Разумеется, для статьи эти ответы были специально выбраны, но сделано это именно потому, что такая картина вырисовывается после анализа всех ответов (табл.).

Длина и качество ответа в зависимости от вежливости промпта

Вид промпта	Длина (количество знаков, медиана)	Верные ответы (количество, среднее арифметическое)
Вежливый	2008	3,6
Без вежливости	1615	2,2

По-видимому, можно предполагать, что более вежливый промпт стимулирует более длинный ответ, а более длинный ответ, в свою очередь, содержит логические цепочки, выписанные в более явном виде, и это приводит к повышению их качества, а значит, и точности ответа.

Конечно же, описанный выше эксперимент не следует воспринимать слишком серьезно. Он состоит всего из $4 + 5 \times 2 = 14$ промптов и ответов на них; диалог велся только с одной нейросетью; набор заданий был крайне специфичен и ограничен; влияние различных средств выражения вежливости в промпте на результат не было изучено. Но, как бы то ни было, это повод задуматься о продолжении таких экспериментов и о более широком тестировании популярных моделей, которые используются при работе с русским языком.

5. Новое в речевом этикете при общении с ИИ

Тот факт, что мы всё больше общаемся с ИИ, влияет на различные аспекты речевого этикета. Часть из этих изменений продолжает тенденции, которые уже наметились в коммуникации в мессенджерах между людьми, но есть и совсем новые. Полноценное не только качественное, но и количественное исследование этих тенденций — дело будущего; я же ограничусь перечислением нескольких фактов из этой области.

Взаимодействие с голосовыми помощниками часто начинается с оклика. Пользователи активируют устройства, обращаясь к ним по имени — например, «Алиса!» или «Окей, Гугл». При этом отсутствует приветствие: реплики типа «Привет!» или «Здравствуйте!» обычно не произносит ни человек, ни машина. Такой способ начала коммуника-



ции отражает функциональный подход к общению, где первостепенная цель — вызвать реакцию системы, а не продемонстрировать приверженность социальным нормам, как в человеческих диалогах. Еще более явно это проявляется в текстовом общении с большими языковыми моделями: там не используются ни оклики, ни приветствия.

Характер благодарностей в диалогах с ИИ также отличается. Пользователи могут выражать благодарность в процессе взаимодействия в ответ на выполненные запросы, но обычно эта благодарность сопровождается просьбой что-нибудь переделать: «Спасибо! А теперь напиши тот же список, но через запятую». А вот в конце сессии благодарность встречается редко: обычно человек просто забирает результаты работы языковой модели и больше не реагирует. Это подчеркивает инструментальное восприятие ИИ, где отсутствует необходимость в поддержании непрерывных социальных связей, характерных для человеческих отношений.

То же касается и ситуации прощания. В традиционном общении прощание служит важным ритуалом, завершающим взаимодействие и показывающим внимание к собеседнику, а в диалогах с ИИ четкое заключительное прощание отсутствует. Однако стоит отметить, что это встречается и в коммуникации между людьми в мессенджерах, где попрощаться тоже не принято.

6. Заключение

Взаимодействие между человеком и ИИ ставит новые вопросы в области вежливости и речевого этикета. Вполне вероятно, что решения этих вопросов будут существенно меняться со временем, но их исследование в любом случае представляется плодотворным уже сейчас. В частности, можно надеяться, что особенности коммуникации человека с ИИ позволят нам лучше понять и некоторые особенности межличностной коммуникации.

Список литературы

Brown P., Levinson S. Politeness: Some universals in language usage. Cambridge, 1987.

Goffman E. The presentation of self in everyday life. Edinburgh, 1956.

Koenig G. La fin de l'individu. Voyage d'un philosophe au pays de l'intelligence artificielle. P., 2019.

Lumer E., Buschmeier H. Should robots be polite? Expectations about politeness in human – robot interaction // *Frontiers in Robotics and AI*. 2023. Vol. 10. doi: 10.3389/frobt.2023.1242127.

Pan X., Doering M., Kanda T. What is your other hand doing, robot? A model of behavior for shopkeeper robot's idle hand // *Proceedings of the 2024 ACM / IEEE International Conference on Human – Robot Interaction*. N. Y., 2024. P. 552 – 560. doi: 10.1145/3610977.3634986.

Ribino P. The role of politeness in human – machine interactions: A systematic literature review and future perspectives // *Artificial Intelligence Review*. 2023. Vol. 56, №1. P. 445 – 482. doi: 10.1007/s10462-023-10540-1.



Shavrina T., Fenogenova A., Emelyanov A. et al. RussianSuperGLUE: A Russian language understanding evaluation benchmark // Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP) / ed. by B. Webber, T. Cohn, Y. He and Y. Liu. Online, 2020. P. 4717–4726. doi: 10.18653/v1/2020.emnlp-main.381.

Yin Z., Wang H., Horio K. et al. Should we respect LLMs? A cross-lingual study on the influence of prompt politeness on LLM performance. 2024. arXiv:2402.14531 [cs. CL].

Об авторе

Александр Чедович Пиперски, кандидат филологических наук, лектор (Assistant Professor), Стокгольмский университет, Швеция.

E-mail: alexander.piperski@slav.su.se

Для цитирования:

Пиперски А. Ч. Вежливость в коммуникации между человеком и искусственным интеллектом // Слово.ру: балтийский акцент. 2024, Т. 15, №4. С. 89–98. doi: 10.5922/2225-5346-2024-4-6.



ПРЕДСТАВЛЕНО ДЛЯ ВОЗМОЖНОЙ ПУБЛИКАЦИИ В ОТКРЫТОМ ДОСТУПЕ В СООТВЕТСТВИИ С УСЛОВИЯМИ ЛИЦЕНЗИИ CREATIVE COMMONS ATTRIBUTION (CC BY) ([HTTP://CREATIVECOMMONS.ORG/LICENSES/BY/4.0/](http://creativecommons.org/licenses/by/4.0/))

POLITENESS IN THE COMMUNICATION BETWEEN HUMANS AND ARTIFICIAL INTELLIGENCE

Alexander Ch. Piperski

Stockholm University,
10 E, Universitetsvägen, 106 91, Sweden, Stockholm

Submitted on 25.04.2024

Accepted on 15.06.2024

doi: 10.5922/2225-5346-2024-4-6

The paper explores the evolution of communication etiquette between humans and artificial intelligence (AI), focusing particularly on the adaptation of traditional politeness strategies. While the politeness of AI can enhance the human level of trust, human politeness towards AI is equally important as it can impact the efficiency of communication. To demonstrate this, I conducted a pilot experiment with ChatGPT 4.0, using polite and non-polite prompts in Russian. The results suggest that politeness is likely to positively impact the accuracy of responses. Furthermore, the paper examines changes in speech etiquette, highlighting how interactions with AI often omit traditional greetings and farewells, reflecting a more functional approach to communication. Although the paper does not provide definitive answers on how politeness strategies between humans and AI should function, it underscores sociolinguistic points that are likely to become increasingly pressing over time. Overall, the paper illustrates a significant shift in communication practices, driven by the integration of AI into daily interactions, necessitating a reevaluation of the established social norms and speech etiquette.

Keywords: communication etiquette, artificial intelligence, politeness strategies, human-AI interaction



References

- Brown, P. and Levinson, S., 1987. *Politeness: Some universals in language usage*. Cambridge, 345 p.
- Goffman, E., 1956. *The presentation of self in everyday life*. Edinburgh, 251 p.
- Koenig, G., 2019. *La fin de l'individu. Voyage d'un philosophe au pays de l'intelligence artificielle*. Paris, 398 p.
- Lumer, E. and Buschmeier, H., 2023. Should robots be polite? Expectations about politeness in human – robot interaction. *Frontiers in Robotics and AI*, 10, <https://doi.org/10.3389/frobt.2023.1242127>.
- Pan, X., Doering, M. and Kanda, T., 2024. What is your other hand doing, robot? A model of behavior for shopkeeper robot's idle hand. In: *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*. New York, pp. 552–560, <https://doi.org/10.1145/3610977.3634986>.
- Ribino, P., 2023. The role of politeness in human – machine interactions: A systematic literature review and future perspectives. *Artificial Intelligence Review*, 56 (1), pp. 445–482, <https://doi.org/10.1007/s10462-023-10540-1>.
- Shavrina, T., Fenogenova, A., Emelyanov, A. et al., 2020. RussianSuperGLUE: A Russian language understanding evaluation benchmark. In: B. Webber, T. Cohn, Y. He and Y. Liu, eds. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 4717–4726, <https://doi.org/10.18653/v1/2020.emnlp-main.381>.
- Yin, Z., Wang, H., Horio, K. et al., 2024. *Should we respect LLMs? A cross-lingual study on the influence of prompt politeness on LLM performance*. arXiv:2402.14531 [cs. CL].

The author

Dr. Alexander Ch. Piperski, Assistant Professor, Stockholm University, Sweden.

E-mail: alexander.piperski@slav.su.se

To cite this article:

Piperski, A. Ch., 2024, Politeness in the communication between humans and artificial intelligence, *Slovo.ru: Baltic accent*, Vol. 15, no. 4, pp. 89–98. doi: 10.5922/2225-5346-2024-4-6.

