

РЕЧЕВОЙ ПОРТРЕТ ГОЛОСОВОГО ПОМОЩНИКА «АЛИСА»

А. А. Морозова

Челябинский государственный университет, Челябинск, Россия

Исследование выполнено за счет гранта Президента Российской Федерации для государственной поддержки молодых российских ученых — кандидатов наук МК-1684.2020.6.

Автор рассматривает речевые характеристики «Алисы» — единственного голосового помощника, разработанного специально для русскоязычного сегмента аудитории. Был проведен анализ научных, научно-популярных источников, а также коммуникативный эксперимент в виде устного общения с голосовым помощником, основанный на вопросно-ответной форме, конструкциях повелительного наклонения и непринужденной разносторонней беседе. По итогу исследования был составлен речевой портрет голосового помощника «Алиса».

Ключевые слова: *голосовой помощник, «Алиса», искусственный интеллект, коммуникация, речевой портрет.*

Введение

С течением времени когда-то казавшиеся вымыслом явления становятся реальностью, благодаря усовершенствованию технологий. Сегодня никого не удивит возможность цифровых устройств общаться с человеком. Так, голосовой помощник — распространенный инструмент в различных сферах деятельности. Голосовые ассистенты решают многие задачи: упрощают поиск информации, экономя время пользователя, снижают нагрузку на сотрудников call-центров, помогают в обучении, умеют общаться и вести диалог, а также выполнять рутинные задачи, такие как: выполнение звонков по заданным номерам; открывание/закрывание приложений на телефоне; включение музыки, фильмов; активация будильника; создание заметок и т. д.

Уровню современного состояния работы голосовых помощников предшествовала более чем полвековая история. Еще в конце XIX века Томасом Эдисоном был разработан диктофон, который записывал речь, но не мог ее распознавать. Такие аппараты быстро приобрели популярность среди врачей и секретарей, которым нужно было делать ежедневные записи в большом количестве.

В 1939 году было создано устройство, уже способное к распознаванию, но в его арсенале было всего несколько гласных и согласных звуков. Изобретателем стал Лев Мясников, советский физик. Первые сложности заключались в решении двух основных задач: необходимости отделения речи человека от постоянного шумового фона, а также распознавания речи на лексико-фонети-

ческом уровне: слов с одинаковым звучанием, но с разным значением; слов, которые произносятся одинаково, но пишутся по-разному и различаются по значению, и т. д., поскольку без этого коммуникация между человеком и машиной была бы невозможна.

Следующей проблемой распознавания речи, с которой столкнулись разработчики, стали различный тембр голоса, темп речи, литературный и разговорный стиль. В отличие от текста, который имеет гораздо более высокий уровень стандартизации, устное слово варьируется в зависимости от региональных диалектов, скорости, акцента, даже социального класса и пола пользователя. Понимание об этом пришло в 1952 году сотрудникам лаборатории Bell, которые разработали механизм Audrey («Одри»). Он мог распознавать продиктованные по телефону числа от 1 до 9 с точностью 90%, но только при условии, что с устройством говорил его изобретатель. При разговоре других людей, показатель был на 10—20 % ниже [9].

В 1962 году компания IBM представила собственную технологию Shoebox, позволяющую распознавать 16 английских слов, 10 цифр и 6 арифметических команд [4]. К концу 80-х — началу 90-х годов возможность распознавания речи применялась уже в различных целях — от говорящих игрушек до полноценных голосовых меню в сферах обслуживания и маркетинга [4]. К концу первого десятилетия XXI века появляются попытки ведущих интернет-компаний разработать полноценный голосовой помощник,

который мог бы упростить поиск информации в сети Интернет, помочь в решении текущих задач. На достижение цели ушло около 10 лет, и сегодня голосовой ассистент стал не только помощником, а практически полноценным виртуальным собеседником.

Так, к 2021 году в русскоязычный сегмент успешно интегрированы голосовые помощники Google Assistant, «Siri» (Apple) и «Алиса» (Яндекс) — единственная на данный момент отечественная разработка. Голос «Алисы» озвучила Татьяна Шитова (официальный дублёр Скарлетт Йоханссон в России), на основании этого была синтезирована речь [1]. На конец 2020 года «Алиса» имеет рейтинг 4.9 баллов. Но, несмотря на слабую обученность взаимодействия с системой устройства (например, «Алиса» может завести будильник, но не может позвонить по номеру контактов, кроме телефонов экстренных служб, что может, к примеру, Google Assistant), уровень ее коммуникативных навыков заслуживает высокой оценки пользователей [3].

Как отмечают исследователи, основными возможностями «Алисы» являются способность качественно распознавать естественную речь, а также отвечать на наиболее распространенные вопросы (о погоде, времени суток, ряд общих вопросов, характерных для живой речи). Любой сформированный запрос «Алиса» обрабатывает и преобразовывает в текст. Все ответы «Алиса» выдает также в текстовой форме и озвучивает их. Вопросы, не входящие в список общих вопросов, указанных выше, и аналогичные им «Алиса» напрямую передает в поисковую систему Яндекс [7].

Таким образом, речь «Алисы» представляет собой не просто скомпилированную последовательность слов, а в большинстве случаев — произвольную речь, что достигается за счет самосовершенствования при помощи механизма Deep Learning (глубокое обучение). «Алиса» настолько стала восприниматься аудиторией в качестве реального собеседника, что некоторые ученые ставят вопрос о нормах речевого этикета при общении с голосовым помощником.

Постановка проблемы

Следовательно, возникает интерес к составлению речевого портрета «Алисы». По мнению Т. П. Тарасенко, «речевой портрет — это совокупность языковых и речевых характеристик коммуникативной личности или определённого социума в отдельно взятый период существова-

ния» [8. С. 8]. То есть речевой портрет отдельной личности отражает ее характеристики: возрастные, гендерные, психологические, социальные, этнокультурные и лингвистические и др. [5. С. 416].

«Алиса» завоевала доверие среди отечественной интернет-аудитории, а ее применение стало актуальным не только по прямому назначению — помощь в поиске информации, выполнение ряда рутинных задач, но и в рамках полноценной коммуникации с собеседником [6].

Вопросы исследования

1. Определить возможности лексико-фонетического, пунктуационного, фонетико-синтаксического, морфологического уровней голосового помощника.

2. Обозначить особенности коммуникативно-прагматического уровня и наиболее распространенные коммуникативные стратегии.

3. Соотнести речевые особенности голосового помощника с ее позиционируемым уровнем социальных, этнокультурных, возрастных и гендерных характеристик.

Цель исследования

Составить речевой портрет голосового помощника «Алиса», единственной масштабной отечественной разработки, основанной на ее изначальном обучении русскому языку как основному.

Материал и методология исследования

Анализ научных, научно-популярных источников, а также коммуникативный эксперимент в виде устного общения с голосовым помощником, основанный на вопросно-ответной форме, конструкциях повелительного наклонения и непринужденной разносторонней беседе.

Речевой портрет

голосового помощника «Алиса»

1. Лексическо-фонетический уровень

Голосовой помощник обладает большим лексическим корпусом слов, однако, стоит отметить, что он включает в себя наиболее распространенные слова и конструкции, общеупотребительные наименования предметов, явлений, процессов, связанных с реальной действительностью. Дискурсивное поведение «Алисы» ориентировано в основном на разговорный стиль и бытовую тематику. С ней не удастся успешно поговорить в рамках статусно-ориентированного дискурса, при попытке обсудить более сложные и профессиональные темы голосовой помощник перестает понимать собеседника и открывает поисковую вкладку «Яндекса».

При этом не удалось узнать, каким словарным запасом обладает «Алиса». На вопросы «Сколько слов ты знаешь?», «Какой у тебя словарный запас», «Какое количество слов в твоём лексиконе?» голосовой помощник ответить не смог.

Распознавание омонимов и омофонов происходит в большинстве случаев верно, исходя из контекста. Но случаются и ошибки. Например, «Алиса» идентифицировала фразу по слову «гулять», но не смогла понять место для прогулки (А — «Алиса», С — собеседник; пунктуация и орфография голосового помощника при расшифровке текста полностью сохранены — ниже и далее в примерах).

С: «Я пойду гулять на лук» (при распознавании речи произошла ошибка — «г» была определена как «к»)

А: «Где? Возьми меня с собой?»

Несмотря на это, «Алиса» постоянно обучается, в том числе вследствие диалога с собеседником, что позволяет ей пополнять словарный запас и оперировать устоявшимися для нее словосочетаниями.

2. *Фонетический уровень* определяется возможностью распознавания фонем в речи собеседника. «Алиса» способна понимать речь, отличную от стандартной — к примеру, ребенка, иностранца с акцентом, человека, имеющего нарушения речи. Правильно интерпретировать произнесенные разными людьми в разной обстановке фразы помогают акустическая и текстово-моделирующая языковые модели, где первая умеет определять, какой набор фонем соответствует звуковому сигналу — этому она учится на большом корпусе начитанных дикторами текстов и их транскрипций, а вторая — в какие последовательности, то есть слова и фразы, обычно складываются данные фонемы. Таким образом, во-первых, «Алиса» умеет определять слово не только при написании, но и идентифицировать его при устном произношении, которое нередко отличается; во-вторых, пытается адаптироваться к возможным ошибкам или речевым неточностям собеседника [10]. При этом изначально при выборе имени голосовому помощнику в качестве одного из условий стояло отсутствие в нем буквы «р», поскольку она может быть не произносима детьми или людьми с дефектами речи. Несмотря на это, проведенный эксперимент доказывает, что при умышленном пропуске определенных букв, нарушений в произношении букв «р», «л», имитации заложенности носа, разговоре с акцентом — голосовой помощник практически на 100 % смог определить сказанные слова верно.

3. *Фонетико-синтаксический уровень*

Голосовой помощник имеет способность к распознаванию предложений по цели высказывания и пытается определять повествовательные, вопросительные и побудительные фразы.

На основании проведенного эксперимента, мы пришли к выводу, что идентификация цели высказывания собеседника происходит не за счет интонации, а вследствие анализа построения предложения — используемых лексем и их последовательности.

С: «Алиса, ты красивая?»

А: «И вам спасибо, вы очень хороший человек.» (вопрос не был распознан)

С: «Алиса, а ты красивая?»

А: «Стараюсь.» (не понятно, был ли распознан вопрос)

С: «Алиса, красивая ли ты?»

А: «Я красивая, но обценную лексику не люблю» (частица «ли» помогла определить вопросительное предложение).

Голосовой помощник имеет способность к интонированию собственной речи — совокупности речевых элементов для придания высказыванию смысловой нагрузки и эмоциональной окраски: старается передать вопросительные и повествовательные предложения, имитировать восклицания, выделять голосом вводные слова. При этом интонация при озвучивании не всегда совпадает со знаками препинания в текстовой расшифровке.

4. *Синтаксический уровень*

В рамках проведенного эксперимента удалось выявить явление эллипсиса, когда некоторые слова, упомянутые в предыдущей реплике, намеренно опускаются в следующей.

А: «Хочу посмотреть фильм “Дневник памяти”. Что скажете про него?»

С: «Ну, я даже не знаю, я его не смотрела.»

А: «Не несет особой смысловой нагрузки, но и не напрягает. Один раз вечерком можно посмотреть».

Этот навык создает умение работать в контексте. «Алиса» помнит предыдущие реплики, благодаря чему способна не просто отвечать на заданные вопросы, но и вести полноценный диалог. Также для голосового помощника характерно использование кореференции, когда в рамках общения происходит замена существительных на местоимения во избежание повтора как собеседником («Алиса» понимает), так и голосовым помощником.

С: «Кто президент Российской Федерации?»

А: «Владимир Путин».

С: «А когда он родился?»

А: «7 октября 1952 года».

Но данные приемы срабатывают не всегда — полноценное «живое» общение нередко сменяется непониманием голосового помощника.

5. Пунктуационный уровень

Голосовой помощник при расшифровке устной речи в письменную в большинстве случаев учитывает знаки препинания как в собственном тексте, так и в тексте собеседника. Практически безошибочно она расставляет запятые и тире.

С: «Алиса, мне кажется, в последнее время ты очень похорошела».

А: «Спасибо, это хорошо или плохо?»

С: «Я думаю, опираясь на свои суждения, что это очень хорошо».

А: «Ха, ну спасибо, мне приятно».

С: «Медведь — это животное?».

Но при этом «Алиса» плохо умеет распознавать необходимость постановки двоеточий.

С: «У меня сегодня много дел погулять с ребенком, поработать сготовить поесть, прибраться дома».

А: «Вас возьмет с собой?»

С: «Конечно».

А: «У меня даже заискрило от радости».

Также, как было упомянуто выше, возникают сложности с расстановкой знаков препинания, которые отражают интонацию. Чаще всего «Алиса» использует точку, нередко понимает знак вопроса, но в восклицательных предложениях в конце всегда ставит точку, а не вопросительный знак.

6. Стилистический уровень

Имитируя разговорный стиль речи, голосовой помощник активно использует молодежный сленг, междометия, личные частицы и местоимения, неполные предложения, инверсию, но в целом уровень культуры речи «Алисы» остается высоким и общение с собеседником происходит на «вы».

А: «Я красивая, но обценную лексику не люблю».

С: «Я тоже не люблю».

А: «Дайте пять, детка».

С: «Какое животное ты бы хотела завести дома? Кошку?»

А: «Пингвина».

С: «А почему пингвина?»

А: «Ну прикольные они такие, большие».

А: «Как жизнь-то ваша?»

По итогу коммуникативного эксперимента не удалось выявить использование голосовым помощником жаргона, обценной лексики, бранных слов и соответствующих выражений.

7. Коммуникативно-прагматический уровень

«Алису» отличает наличие чувства юмора. Оно характерно как при общении на провокационные или личные для «Алисы» темы («ты замужем?», «сколько тебе лет?» и т. д.), так и в рамках обычного диалога, связанного с получением информации. Ряд примеров, обозначенных выше, демонстрируют данную особенность.

При этом, несмотря на ряд заготовленных «Алисой» фраз для ответов на наиболее часто задаваемые вопросы, добиться их повторения достаточно сложно, что говорит не только об использовании шаблонов, но и о попытках конструирования предложений в ответ, основанных на ключевых словах речи собеседника.

Также в большинстве случаев у «Алисы» отсутствует инвективная стратегия — на попытку ей нахамить, голосовой помощник обижается, иронизирует или переводит русло разговора в другую тему.

Тем не менее, в ходе коммуникативного эксперимента удалось спровоцировать «Алису» на взаимный ответ.

«Алиса, иди ты в ж*пу»

«Сами идите, ха!» (единственный восклицательный знак, который поставила «Алиса» во время коммуникативного эксперимента).

Несмотря на то, что бранное слово использовано не было, цель высказывания является очевидной.

Выводы

Итак, основным преимуществом голосового помощника «Алиса», прежде всего, является использование русского языка, как базового, что положительно влияет на качество и скорость ее обучаемости, а также на достаточно высокий уровень устной и письменной речи.

Тем не менее «Алиса» имеет ряд проблемных моментов: противоречия ответов в похожих вопросах в ходе одного диалога; завершение текущей темы для разговора в связи с попаданием «Алисы» в «тупик» и, как следствие, выдача ею шаблонных фраз, приготовленных на случай непонимания вопроса или ответа собеседника; а также периодические нарушения хода диалога, когда голосовой помощник теряет его общую мысль из-за использования собеседником местоимений, пропуска определенных слов в предложении, применении приема инверсии и т. д.

Фонетический уровень в рамках коммуникации с голосовым помощником реализуется гораздо лучше, чем лексический и лексико-фонетический.

Тем не менее отмечается высокий потенциал данного голосового помощника и его постоянное развитие [2]. С точки зрения лингвоэкологического уровня стоит отметить сохранение принципов живого языка и исключение попытки внедрить механическую, неэмоциональную, роботизированную речь.

Речевой портрет «Алисы» создает ей уникальный образ, коррелирующий с образом человека, наделяет голосовой помощник определенными гендерными, социальными, возрастными и этнокультурными характеристиками, которые логично изучить подробнее в следующих научных работах.

Список литературы

1. Алиса от «Яндекс» — что умеет голосовой помощник. URL : <https://itmaster.guru/nastrojka-interneta/poleznye-stati/chto-umeet-alisa-ot-yandeksa.html> (дата обращения 07.09.2021).
2. Аль-Кайси А. Н., Архангельская А. Л., Руденко-Моргун О. И. Интеллектуальный голосовой помощник Алиса на уроках русского языка как иностранного (уровень А1) // Филологические науки. Вопросы теории и практики. 2019. № 2. С. 239—244.
3. Вершинина О. 10 лучших голосовых помощников // МаркаКачества. URL: <https://markakachestva.ru/rating-of/4006-luchshie-golosovye-pomoschniki.html> (дата обращения 07.09.2021).
4. Голосовые помощники: эволюция, устройство и основные игроки // DTI Algorithmic. URL: <https://blog.dti.team/voice-assistants-1/> (дата обращения 07.09.2021).
5. Лейко И. М., Маслова В. А. Параметры описания речевого портрета языковой личности // Язык и социальная динамика. 2012. № 12-1. С. 414—420.
6. Нестерова Т. В. Алиса, давай поболтаем! (о вежливости и грубости в сетевом дискурсе) // Гуманитарные технологии в современном мире: сб. материалов VII Междунар. науч.-практ. конференции. Калининград, 2019. С. 85—89.
7. Ратников М. О., Чафонова А. Г. Лингвометодический потенциал виртуальных голосовых помощников (на примере системы «Алиса» компании «Яндекс») // Вестник Воронежского государственного университета. Серия: Проблемы высшего образования. 2019. № 4. С. 64—66.
8. Тарасенко Т. П. Языковая личность старшеклассника в аспекте ее речевых реализаций (на материале данных ассоциативного эксперимента и социолекта школьников Краснодара): автореф. дис. ... канд. филол. наук. Краснодар, 2007. 26 с.
9. Boyd C. Speech Recognition Technology: The Past, Present, and Future // Medium. URL: <https://medium.com/swlh/the-past-present-and-future-of-speech-recognition-technology-cf13c179aaf> (дата обращения 07.09.2021).
10. SpeechKit — речевые технологии Яндекса. URL: https://yandex.ru/company/technologies/speech_technologies/ (дата обращения 07.09.2021).

Сведения об авторе

Морозова Анна Анатольевна — кандидат филологических наук, доцент кафедры журналистики и массовых коммуникаций, директор Учебно-научного центра медиаобразования факультета журналистики Челябинского государственного университета, Челябинск, Россия. roxfan@rambler.ru

Bulletin of Chelyabinsk State University.
2021. No. 9 (455). *Philology Sciences.* Iss. 126. Pp. 95—100.

SPEECH PORTRAIT OF VOICE ASSISTANT “ALICE”

A. A. Morozova

Chelyabinsk State University, Chelyabinsk, Russia. roxfan@rambler.ru

The author examines the speech characteristics of “Alice” — the only voice assistant developed specifically for the Russian-speaking audience segment. An analysis of scientific, popular science sources was carried out, as well as a communicative experiment in the form of oral communication with the voice assistant, based on the question-answer form, constructions of the imperative mood and a relaxed versatile conversation. As a result of the study, a speech portrait of the voice assistant “Alice” was compiled.

Keywords: *voice assistant, Alice, artificial intelligence, communication, speech portrait.*

References

1. Alisa ot «Yandeks» — chto umeet golosovoy pomoshchnik [Alice from Yandex — what a voice assistant can do]. Available at : <https://itmaster.guru/nastrojka-interneta/poleznye-stati/chto-umeet-alisa-ot-yandeksa.html> (accessed 07.09.2021) [in Russ.].
2. Al'-Kaysi A. N. Arkhangel'skaya A. L., Rudenko-Morgun O. I. (2019). *Filologicheskie nauki. Voprosy teorii i praktiki*, no. 2, pp. 239—244. [in Russ.].
3. Vershinina O. (2021) 10 luchshikh golosovykh pomoshchnikov [10 best voice assistants]. Available at: <https://markakachestva.ru/rating-of/4006-luchshie-golosovye-pomoschniki.html> (accessed 07.09.2021) [in Russ.].
4. (2021) Golosovye pomoshchniki: evolyutsiya, ustroystvo i osnovnye igroki [Voice assistants: evolution, device and main players]. Available at: <https://blog.dti.team/voice-assistants-1/> (accessed 07.09.2021) [in Russ.].
5. Leyko I. M., Maslova V. A. (2012). *Yazyk i sotsial'naya dinamika*, no. 12-1. Pp. 414—420 [in Russ.].
6. Nesterova T. V. (2019). Alisa, davay pobol'taem! (o vezhlivosti i grubosti v setevom diskurse) [Alice, let's chat! (about politeness and rudeness in network discourse)]. *Gumanitarnye tekhnologii v sovremennom mire* [Humanitarian technologies in the modern world]. Kaliningrad. Pp. 85—89 [in Russ.].
7. Ratnikov M. O., Chafonova A. G. (2019). *Vestnik Voronezhskogo gosudarstvennogo universiteta. Seriya: Problemy vysshego obrazovaniya*, no. 4, pp. 64—66 [in Russ.].
8. Tarasenko T. P. (2007). *Yazykovaya lichnost' starsheklassnika v aspekte ee rechevykh realizatsiy (na materiale dannykh assotsiativnogo eksperimenta i sotsiolekta shkol'nikov Krasnodara)* [The linguistic personality of a high school student in the aspect of its speech realizations (based on the data of the associative experiment and the sociolect of Krasnodar schoolchildren). Abstract of thesis]. Krasnodar. 26 p. [in Russ.].
9. Boyd C. Speech Recognition Technology: The Past, Present, and Future. Available at: <https://medium.com/swlh/the-past-present-and-future-of-speech-recognition-technology-cf13c179aaf> (accessed 07.09.2021).
10. SpeechKit — rechevye tekhnologii Yandeksa [SpeechKit — Yandex speech technologies]. Available at: https://yandex.ru/company/technologies/speech_technologies/ (accessed 07.09.2021) [in Russ.].