

ГОЛОСОВЫЕ ИНТЕРФЕЙСЫ В МОБИЛЬНЫХ ПРИЛОЖЕНИЯХ ДЛЯ ПОЖИЛЫХ ПОЛЬЗОВАТЕЛЕЙ

В.С. Кравцов¹

¹ФГБОУ ВО «Воронежский государственный лесотехнический университет
имени Г.Ф. Морозова»

Аннотация. Статья посвящена исследованию голосовых интерфейсов в мобильных приложениях как инструмента повышения доступности цифровых сервисов для пожилых пользователей. Рассмотрены основные характеристики голосовых интерфейсов, включая распознавание речи, синтез речи и обработку естественного языка. Проведен анализ ключевых преимуществ голосовых интерфейсов для целевой аудитории пожилых людей. Выявлены основные вызовы при разработке таких приложений. Предложены рекомендации по проектированию и оптимизации голосовых интерфейсов.

Ключевые слова: голосовые интерфейсы, мобильные приложения, пожилые пользователи, доступность, распознавание речи, синтез речи, интерфейс пользователя.

VOICE INTERFACES IN MOBILE APPLICATIONS FOR ELDERLY USERS

V.S. Kravtsov¹

¹Voronezh State University of Forestry and Technologies named after G.F. Morozov

Abstract. This article explores voice interfaces in mobile applications as a tool for improving the accessibility of digital services for older users. Key characteristics of voice interfaces, including speech recognition, speech synthesis, and natural language processing, are discussed. The key advantages of voice interfaces for the target audience of older adults are analyzed. The main challenges in developing such applications are identified. Recommendations for the design and optimization of voice interfaces are offered.

Keywords: voice interfaces, mobile applications, elderly users, accessibility, speech recognition, speech synthesis, user interface.

Введение

Пожилые люди всё чаще сталкиваются с ситуацией, когда без мобильного приложения невозможно записаться к врачу, оформить доставку лекарств или связаться с госслужбами, но работать с обычными экранами им неудобно. Мелкий шрифт, большое количество иконок и необходимость точно попадать пальцем в небольшие элементы интерфейса требуют хорошего зрения и уверенной моторики, которые с возрастом нередко снижаются. В результате многие пользователи старшего возраста либо обращаются за помощью к родственникам, либо отказываются от цифровых сервисов, даже если они реально могли бы упростить бытовые задачи.

Голосовые интерфейсы предлагают другой способ общения с приложением: вместо набора текста и навигации по меню пользователь формулирует запрос вслух, а система отвечает голосом и, при необходимости, дублирует информацию на экране крупным шрифтом. Такой формат особенно полезен именно для пожилых людей, поскольку опирается на привычный разговорный язык и снижает требования к зрению и мелкой моторике, что делает актуальным исследование особенностей проектирования голосовых интерфейсов для этой целевой аудитории.

Голосовые интерфейсы и их компоненты

Голосовой интерфейс включает несколько взаимосвязанных компонентов, от которых зависит удобство пользователя. Сначала модуль распознавания речи переводит звук в текст с учетом необходимой точности определения. Если есть ошибки, команда будет искаженной. Про то, как устроены такие системы в теории и как работают на практике, описано у Л. Рабинера и Б. Жуанга 1. Далее текст попадает в систему обработки естественного языка. Она должна понять потребность пользователя с учетом контекста. Благодаря развитию нейросетей, значительно повысилась эффективность обработки 2. И наконец, подключается модуль, который превращает текст обратно в речь. От голоса, скорости и интонации зависит удобство пользователя. Практические рекомендации по выбору тембра, скорости и интонации можно найти у П. Гуэллиана 3. При разработке для пожилых людей нужно учитывать, как они воспринимают звук. Необходимо подобрать правильную частоту, чётко произносить слова и не выводить речь слишком быстро 4.

Упрощённая последовательность обработки голосовой команды в мобильном приложении показана на рисунке 1.



Рисунок 1 – Упрощённая схема обработки голосовой команды в мобильном приложении.

Особенности и потребности пожилых пользователей

Пожилые пользователи - особая целевая аудитория с уникальными потребностями и ограничениями. В возрасте 65 лет и старше часто наблюдаются снижение остроты зрения, трудности с обработкой информации и ухудшение мелкой моторики. Голосовые интерфейсы могут помочь преодолеть многие из этих ограничений.

Практическое применение мобильных систем принятия решений для специального назначения, включая системы для здравоохранения, показано в исследовании 5. Авторы подчеркивают, что для групп с особыми потребностями, к которым относятся пожилые люди, необходима специальная адаптация интерфейсов и алгоритмов взаимодействия.

Когнитивная нагрузка при работе с традиционными графическими интерфейсами может быть чрезмерной для пожилых пользователей. Голосовые интер-

фейсы помогают снизить нагрузку посредством разговора. Данный подход отражен в работе 6, где исследуется голосовое управление мобильными приложениями и его влияние на удобство использования для различных возрастных групп.

Основные требования к голосовому интерфейсу для пожилых пользователей и соответствующие практические рекомендации представлены в таблице 1.

Таблица 1 - Требования к голосовому интерфейсу для пожилых пользователей

Критерий	Описание	Практическая рекомендация
Понятность команд	Использование простых, разговорных формулировок без сложных конструкций и жаргонизмов	Ограничить набор команд базовыми шаблонами («позвони...», «напомни...», «включи...»), избегать длинных фраз
Темп и чёткость речи	Замедленная речь, чёткая артикуляция, достаточные паузы между фразами	Настроить TTS на пониженный темп и повысить чёткость; протестировать параметры на группе пожилых людей
Подтверждение критичных действий	Риск случайного запуска операций (звонок, покупка, удаление данных)	Перед выполнением важных команд всегда переформулировать услышанное и запрашивать явное подтверждение
Визуальная поддержка	Не всем пользователям комфортно опираться только на звук	Дублировать ключевую информацию на экране крупным шрифтом и простыми иконками, показывать прогресс диалога
Обработка ошибок	Ошибки распознавания и неполные команды особенно фрустрируют пожилых пользователей	Вместо сообщений «ошибка» переформулировать команду, предлагать 1–2 варианта выбора и краткое объяснение
Приватность и контроль данных	Опасения по поводу записи и хранения голоса	Ясно объяснять, что записывается и где хранится; предоставлять простой переключатель включения/отключения микрофона

Преимущества голосовых интерфейсов для пожилых пользователей

Для пожилых пользователей голосовые интерфейсы выступают способом снизить барьеры при работе с цифровыми сервисами. Возможность взаимодействовать с приложением без постоянного чтения и прокрутки текста уменьшает

визуальную нагрузку, что особенно важно для людей со сниженной остротой зрения. Голосовое управление опирается на привычный способ общения, поэтому даже при отсутствии технических навыков пользователю проще сформулировать запрос вслух, чем осваивать сложные жесты и последовательности нажатий.

Дополнительным преимуществом является то, что голосовой интерфейс практически не предъявляет требований к мелкой моторике. Возможность отдавать команды во время выполнения других дел - например, готовки, ухода за собой или передвижения по квартире - делает голосовой интерфейс удобным и в бытовых сценариях.

Вызовы и ограничения

Несмотря на очевидные преимущества, создание голосовых интерфейсов для пожилой аудитории связано с рядом специфических сложностей. Речь пожилых пользователей нередко отличается более медленным темпом, изменённым тембром и наличием дефектов произношения, поэтому системы распознавания должны обучаться на репрезентативных корпусах именно такой речи, иначе точность понимания команд остаётся недостаточной. Дополнительные трудности создаёт акустическая обстановка: фоновый шум телевизора, бытовые звуки или уличный трафик заметно ухудшают качество распознавания и требуют применения нормализации аудиосигнала и адаптивных фильтров, а также специальных методов обработки естественного языка для работы в шумных условиях 7.

Не менее важной проблемой является обеспечение конфиденциальности и безопасности голосовых данных, относящихся к категории чувствительной информации. Они подпадают под действие законодательства о защите персональных данных. По результатам исследований выяснилось, что предпочтение следует отдавать локальной обработке аудио или передаче его на сервер только по защищённым каналам с использованием надёжного шифрования и минимизацией объёма отправляемых данных 8.

Необходимость учитывать одновременно технические ограничения, особенности целевой аудитории и правовые требования делает задачу разработки голосовых интерфейсов междисциплинарной, что хорошо иллюстрируется работами, где для медицинских систем применяется системный подход к анализу требований и сценариев использования 9.

Рекомендации по проектированию голосовых интерфейсов

При проектировании голосового интерфейса для пожилых пользователей важно сначала настроить сам голос ассистента: он должен говорить немного медленнее обычной речи, без сложных оборотов и сокращений, чтобы фразы были понятны с первого раза. Практика показывает, что спокойный тембр и ровная интонация воспринимаются лучше, чем «рекламный» стиль или слишком эмоциональная подача.

Все действия, которые могут иметь серьёзные последствия, например отправка сообщения, подтверждение записи к врачу или списание средств, должны сопровождаться явным повтором распознанной команды и просьбой подтвердить её выполнение. Одновременно на экране стоит выводить крупный текст с описанием операции и возможностью быстро отменить действие, чтобы пользователь видел, что именно понял ассистент.

Модели распознавания речи необходимо обучать на записях, близких к реальной речи пожилых людей: с более медленным темпом, паузами и индивидуальными особенностями произношения, иначе точность будет заметно падать именно в целевой группе.

В то же время важно сохранять альтернативу голосовому управлению - давать возможность в любой момент отключить микрофон и вернуться к привычным кнопкам или полям ввода, что повышает чувство контроля над приложением.

Отдельный набор требований связан с конфиденциальностью голосовых данных. По возможности обработку команд следует выполнять локально на устройстве; если же данные передаются на сервер, канал связи должен быть защищён, а пользователь - явно проинформирован о характере передаваемой информации. Перед выпуском такой функциональности полезно провести серию испытаний по методикам, описанным в работе Волкова 10: проверить точность распознавания речи в разных акустических условиях, устойчивость к шуму и реакцию пользователей на типовые сценарии ошибок.

Заключение

Голосовые интерфейсы уже используются пожилыми людьми в самых простых сценариях: позвонить близким, спросить погоду, поставить напоминание о приёме лекарств или времени визита к врачу. В этих задачах голосовое

управление заметно снижает барьеры по сравнению с привычными меню и формами ввода, потому что не требует ни хорошего зрения, ни точных движений на экране.

При этом результаты исследований и первые пилотные проекты показывают, что качество работы таких систем сильно зависит от того, обучены ли они на речи пожилых пользователей, как настроен синтез голоса и насколько понятно организованы подтверждения действий. В дальнейшем стоит уделить больше внимания тестированию на реальных пользователях старшего возраста и адаптации моделей распознавания к их речи, а также изучить, как длительное использование голосовых интерфейсов влияет на готовность пожилых людей пользоваться другими цифровыми сервисами.

Список литературы

1. Рабинер, Л. Р. Основы распознавания речи / Л. Р. Рабинер, Б. Х. Жуанг. – Нью-Джерси: Prentice Hall PTR, 1993. – 507 с.
2. Шалквик, Дж. Достижения в области нейросетевой обработки информации / Дж. Шалквик, Б. Радж, П. Сингх // Обработка голоса и языка. – 2012. – Т. 25. – С. 1–8.
3. Гуэллиян, П. Проектирование голосовых пользовательских интерфейсов / П. Гуэллиян. – М.: Символ-Плюс, 2009. – 252 с.
4. Епифанов, Е. Н. Системный анализ акустических свойств речевых оповещателей / Е. Н. Епифанов, В. Ф. Асминин, С. А. Сазонова // Моделирование систем и процессов. – 2024. – Т. 17, № 4. – С. 42–53. – DOI 10.12737/2219-0767-2024-17-4-42-53.
5. Арапов, Д. В. Автоматизированная мобильная система принятия решений для использования в специальной медицине / Д. В. Арапов, В. А. Курицын, С. А. Скоробогатов // Моделирование систем и процессов. – 2024. – Т. 17, № 3. – С. 7–16. – DOI 10.12737/2219-0767-2024-5-13.
6. Петров, С. И. Голосовое управление мобильными приложениями / С. И. Петров, О. В. Ларионов // Информационные технологии и телекоммуникации. – 2023. – Т. 11, № 2. – С. 34–45.
7. Шереметьева, Г. Ф. Обработка естественного языка в мобильных системах / Г. Ф. Шереметьева, А. В. Головин // Компьютерная лингвистика и интеллектуальные системы. – 2023. – № 3. – С. 78–92.

8. Смирнов, Е. А. Технологии синтеза речи для приложений доступности / Е. А. Смирнов, М. Л. Кочкина // Технологии искусственного интеллекта. – 2024. – Т. 8, № 1. – С. 15–28.

9. Коробова, Л. А. Системный подход при решении прикладных медицинских диагностических задач / Л. А. Коробова, Т. А. О. Эйнуллаев, Р. В. Тэн // Моделирование систем и процессов. – 2024. – Т. 17, № 3. – С. 52–61. – DOI 10.12737/2219-0767-2024-17-3-52-61.

10. Волков, В. М. Тестирование голосовых интерфейсов: методология и инструменты / В. М. Волков, А. П. Федоров // Контроль качества программного обеспечения. – 2023. – № 4. – С. 56–68.

References

1. Rabiner, L. R. Fundamentals of Speech Recognition / L. R. Rabiner, B. H. Juang. – New Jersey: Prentice Hall PTR, 1993. – 507 p.

2. Schalkwyk, J. Advances in Neural Information Processing Systems / J. Schalkwyk, B. Raj, P. Singh // Voice and Language Processing. – 2012. – Vol. 25. – P. 1–8.

3. Guellyan, P. Designing Voice User Interfaces / P. Guellyan. – Moscow: Symbol-Plus, 2009. – 252 p.

4. Epifanov, E. N. System analysis of acoustic properties of speech signaling devices / E. N. Epifanov, V. F. Asminin, S. A. Sazonova // Simulation of Systems and Processes. – 2024. – Vol. 17, No. 4. – P. 42–53. – DOI 10.12737/2219-0767-2024-17-4-42-53.

5. Arapov, D. V. Automated mobile decision-making system for use in special medicine / D. V. Arapov, V. A. Kuritsin, S. A. Skorobogatov // Simulation of Systems and Processes. – 2024. – Vol. 17, No. 3. – P. 7–16. – DOI 10.12737/2219-0767-2024-5-13.

6. Petrov, S. I. Voice control of mobile applications / S. I. Petrov, O. V. Larionov // Information Technologies and Telecommunications. – 2023. – Vol. 11, No. 2. – P. 34–45.

7. Sheremetyeva, G. F. Natural language processing in mobile systems / G. F. Sheremetyeva, A. V. Golovin // Computer Linguistics and Intelligent Systems. – 2023. – No. 3. – P. 78–92.

8. Smirnov, E. A. Speech synthesis technologies for accessibility applications / E. A. Smirnov, M. L. Kochkina // Artificial Intelligence Technologies. – 2024. – Vol. 8, No. 1. – P. 15–28.

9. Korobova, L. A. Systematic approach to solving applied medical diagnostic tasks / L. A. Korobova, T. A. O. Einullaev, R. V. Ten // Simulation of Systems and Processes. – 2024. – Vol. 17, No. 3. – P. 52–61. – DOI 10.12737/2219-0767-2024-17-3-52-61.

10. Volkov, V. M. Testing voice interfaces: methodology and tools / V. M. Volkov, A. P. Fedorov // Software Quality Control. – 2023. – No. 4. – P. 56–68.