

# О применении систем автоматического перевода устной речи

М. В. Кругликова

Санкт-Петербургский государственный электротехнический университет  
«ЛЭТИ» им. В.И. Ульянова (Ленина)

k.maria.21@mail.ru

**Аннотация.** В настоящем исследовании обсуждается прогресс, достигнутый за последнее время в области систем автоматического перевода речи, и рассмотрены возможные способы их применения. Выявлены ограничения современных программ автоматического перевода речи, такие как недостаточный учет контекста и культурных особенностей, а также невозможность генерации естественных переводческих трансформаций. Сделан вывод о необходимости сотрудничества разработчиков с профессиональными переводчиками с целью улучшения качества автоматического перевода.

**Ключевые слова:** автоматический перевод речи; автоматическое распознавание речи; синтез речи; машинный перевод

## I. СИСТЕМЫ АВТОМАТИЧЕСКОГО ПЕРЕВОДА УСТНОЙ РЕЧИ

В настоящее время системы автоматического перевода играют важную роль в преодолении языковых барьеров и обеспечении более эффективного общения между людьми, говорящими на разных языках. Благодаря быстрому развитию технологий и искусственного интеллекта, системы автоматического перевода становятся все более совершенными и доступными и находят широкое практическое применение: голосовые помощники, программы автоматического субтитрования, верификации говорящего по голосу и речи и др.

В области автоматического перевода речи регулярно проводятся научные исследования и разработки, благодаря чему замечен ощутимый прогресс в развитии технологий. Современные системы перевода речи способны осуществлять перевод устной речи с одного естественного языка на другой в режиме реального времени. Однако вопрос качества перевода остается актуальным, что подчеркивает важность исследований в этой области.

Цель данного исследования – обсудить возможное применение систем автоматического перевода речи, выявить причины недостаточно эффективной работы этих систем.

В контексте автоматического перевода речи выделяют как область научных исследований, так и разработку специальных программ, систем и сервисов, позволяющих переводить устную речь с одного естественного языка на другой, а также непосредственно автоматический перевод речи с помощью программ, систем и сервисов [1].

Процесс автоматического перевода речи происходит в три основных этапа: автоматическое распознавание

речи, машинный перевод и синтез речи. Каждый из этапов характеризуется своими сложностями, требующими оптимального решения. В частности, если мы говорим о системах перевода речи, они должны быть способны правильно распознавать речь в различных условиях, в том числе при наличии шума в окружении говорящего. Также для распознавания речи могут представлять сложность диалекты, акценты, какие-либо индивидуальные особенности речи говорящего.

Разработчики постоянно совершенствуют архитектуру систем перевода речи и пробуют различные подходы. Долгое время процесс автоматического перевода речи включал в себя этап преобразования речи в текст с целью машинного перевода полученного текста и дальнейшего синтеза речи. Однако компания «Google» смогла разработать технологию, способную обрабатывать речевые сигналы напрямую, без преобразования речи в текст («end-to-end direct speech-to-speech translation»), что в свою очередь позволило учитывать и воспроизводить при переводе интонацию говорящего и характерные особенности звучания голоса [3].

Системы автоматического перевода речи имеют очевидные преимущества, обуславливающие их применение в различных областях. Во-первых, они помогают в преодолении языкового барьера между представителями разных культур, позволяя им общаться в реальном времени. Во-вторых, системы перевода обеспечивают возможность доступа к различным источникам информации людям, не владеющим иностранными языками.

Одна из возможных сфер применения технологий автоматического перевода речи – здравоохранение. В недавнем исследовании П. Хадельсон и Ф. Шапюю были изучены возможности использования приложений автоматического перевода речи при проведении консультаций медицинских специалистов с пациентами, говорящими на разных языках [2]. Исследователи предложили специалистам Женевской университетской больницы (Geneva University Hospitals) провести консультации со своими иностранными пациентами, используя одну из двух программ перевода речи («Microsoft Translate» или «Pocketalk W»). После каждой консультации специалисты и пациенты должны были пройти опрос и оценить их опыт использования программ во время консультаций. В ходе исследования 14 специалистов провели 52 консультации с пациентами, говорящими на 13 разных языках. Специалистам удалось успешно провести 43/52 консультаций (82.7%), однако они были удовлетворены тем, как прошла коммуникация

с пациентами, только в 28/52 случаев (53.8 %). Специалисты достаточно хорошо понимали пациентов в 37/52 случаях (71.2 %). 87.8 % пациентов положительно оценили опыт общения с помощью программ автоматического перевода, и 77 % пациентов ответили, что пользование программами, а не услугами переводчиков, является более предпочтительным для обсуждения деликатных вопросов с медицинскими работниками. Также сравнивая работу программ и профессиональных переводчиков, пациенты указали на большую доступность программ перевода речи, обеспечивающих мгновенный доступ к переводу без необходимости ожидания назначенного времени. Большинство опрошенных (73 % специалистов и 84 % пациентов) хотели бы продолжить использовать системы автоматического перевода в дальнейшем во время консультаций.

В итоге, медицинские специалисты и пациенты отметили, что в процессе общения трудности возникали, в первую очередь, из-за недостатка их опыта работы с программами, а также в связи с необходимостью выбирать более простые конструкции при выражении мыслей для наиболее точного автоматического перевода. Специалисты в области здравоохранения отметили, что использование систем автоматического перевода речи наиболее уместно для консультаций, предполагающих обмен фактической информацией (медицинские осмотры, простые обследования, повторные приемы) [2].

Тем самым, автоматический перевод речи, даже в медицинской сфере, весьма определенной в плане терминологии, имеет как очевидные преимущества, так и существенные недостатки. Как показывает проведенное исследование, системы перевода в настоящий момент уже могут использоваться для облегчения общения между людьми, говорящими на разных языках, в частности во время проведения медицинских консультаций и осмотров. Эти системы пока еще не способны полностью заменить профессиональных переводчиков, так как не могут обеспечить качественный перевод: программы не прибегают к переводческим трансформациям, позволяющим более деликатно выразить ту или иную мысль, недостаточно адаптированы в лингвокультурном плане и не учитывают контекст, что зачастую приводит к серьезным ошибкам перевода.

Успешная техническая реализация высокоэффективных переводческих автоматов невозможна без тренировки нейросетевых моделей, для чего необходимо наличие объемных корпусов данных с

аудио- и видеофайлами. Для английского языка, например, существует достаточное количество открытых корпусов данных, чего нельзя сказать про менее распространенные языки. Соответственно, требуются дополнительные усилия по сбору данных, что может усложнять процесс разработки систем перевода. В связи с этим сотрудничество разработчиков с лингвистами приобретает особую значимость, так как экспертные знания лингвистов необходимы для анализа языковых данных, создания корпусов и метрик оценки качества автоматического перевода.

## II. ЗАКЛЮЧЕНИЕ

Проведенный анализ позволяет сделать следующие выводы:

1. Системы автоматического перевода речи находят широкое применение в различных областях, в частности, в медицине, что подтверждает их практическую пользу и указывает на необходимость проведения дальнейших исследований с целью улучшения систем перевода и расширения их функциональности.
2. Несмотря на очевидный прогресс, достигнутый за последнее время в области автоматического перевода речи, разработчикам следует продолжать работу над качеством перевода совместно с профессиональными переводчиками.

## СПИСОК ЛИТЕРАТУРЫ

- [1] Сухоручкина И.Н. Сервисы автоматического перевода речи при языковых контактах / И.Н. Сухоручкина, А.А. Сухоручкина // Военно-гуманитарный альманах: Материалы XVI Международной научной конференции по актуальным проблемам языка и коммуникации, Москва, 22–24 июня 2022 года / Под общей редакцией Н.В. Иванова. Москва: Военный университет, 2022. С. 115-123. URL: <https://elibrary.ru/item.asp?id=49801724> (дата обращения: 17.03.2024).
- [2] Hudelson P., Chappuis F. Using Voice-to-Voice Machine Translation to Overcome Language Barriers in Clinical Communication: An Exploratory Study // Journal of General Internal Medicine. 2024. URL: [https://www.researchgate.net/publication/378173785\\_Using\\_Voice-to-Voice\\_Machine\\_Translation\\_to\\_Overcome\\_Language\\_Barriers\\_in\\_Clinical\\_Communication\\_An\\_Exploratory\\_Study](https://www.researchgate.net/publication/378173785_Using_Voice-to-Voice_Machine_Translation_to_Overcome_Language_Barriers_in_Clinical_Communication_An_Exploratory_Study) (дата обращения: 18.03.2024).
- [3] Jia Ye et al. Translatotron 2: High-quality direct speech-to-speech translation with voice preservation. International Conference on Machine Learning (2022). URL: <https://arxiv.org/pdf/2107.08661.pdf> (дата обращения: 18.03.2024).