

Исследование систем распознавания дефектной речи

Автор: З.Х. Магдиева, студент, Университет ИТМО, Санкт-Петербург

Научный руководитель: к.т.н., доцент, П.В. Балакшин, Университет ИТМО

Введение. Природа человека диктует необходимость в социальном взаимодействии, однако это не всегда доступно людям с дефектами речи. Современные реалии способствуют росту числа людей с дисфункцией речи по множеству причин. Увеличение средней продолжительности жизни влечёт за собой прирост пожилого населения, которое подвержено нарушениям речепроизводства и артикуляции ввиду органических поражений мозга. Растущая автомобилизация увеличивает вероятность аварийных ситуаций и физических повреждений, следствием которых являются проблемы с речевым аппаратом. Также нередки случаи врождённых патологий. В связи с этим становится актуальной проблема обеспечения нормального функционирования людей с дефектами речи в обществе. Для решения этой проблемы предлагается использование систем распознавания речи (СРР).

Цель работы. Выявление факторов, влияющих на эффективность распознавания речи, с последующим определением наиболее точной СРР относительно исследуемых дефектов речи.

Базовые положения исследования. Системы распознавания речи подразделяются на несколько категорий, включая:

- потребительские качества: дикторонезависимые, дикторозависимые или диктороадаптивные;
- тип речи: отдельная или слитная речь;
- назначение системы: командная или система диктовки;
- объём словаря;
- используемые алгоритмы и методы.

Ввиду различной программной реализации СРР появляется необходимость в базовом показателе качества распознавания, единым для всех подгрупп. Для этого в исследовании задействуется контрольная группа, не имеющая выраженных речевых патологий. Корреляционный анализ способен выявить зависимость эффективности СРР для людей с дефектами речи от исходных параметров, что определяется через стандартную метрику Word Error Rate (WER). Предполагается, что результаты контрольной группы будут лучше результатов экспериментальной, в особенности при распознавании отдельных слов вне контекста, в силу большей разборчивости речи. Также ожидается выявление обратной зависимости между объёмом словаря и метрикой WER.

Промежуточные результаты. Анализ существующих исследований показал, что эффективность распознавания дефектной речи зависит от единообразия речепроизводства. При этом СРР оказалась более восприимчивой к дефектной речи, чем обычный слушатель. Негативное влияние на точность распознавания оказывали внешний шум, переутомление экспериментальных групп вследствие длительных тренировок СРР и большой объём исходного словаря.

Основной результат. В результате исследования был сделан вывод о целесообразности использования СРР на ежедневной основе группами лиц с дефектами речи, характеризующимися отсутствием динамических изменений или скачков. Для тяжёлых нарушений речи вплоть до отсутствия голоса рекомендовано дополнительное использование детекторов мышечной активности. Учитывая некоторую схожесть дефектной речи и речи человека в состоянии алкогольного опьянения возможно применение искусственного интеллекта для выявления уровня трезвости водителя и последующего принятия решения об обоснованности управления транспортным средством.

Автор

Магдиева З.Х.

Научный руководитель

Балакшин П.В.