

ИССЛЕДОВАНИЕ МЕТОДОВ РАСПОЗНАВАНИЯ ФОНЕМ РУССКОГО ЯЗЫКА И ИХ ИНТЕГРАЦИИ В ПРОЦЕСС ДИАГНОСТИКИ ЗВУКОПРОИЗНОШЕНИЯ У ДЕТЕЙ ДОШКОЛЬНОГО ВОЗРАСТА

Пономарева С.В.¹, Рахманов С.А.¹

Научный руководитель – канд. техн. наук, доцент Добренко Н.В.¹

¹Университет ИТМО

Введение

Около 60% российских школьников поступают в первый класс с нарушениями устной речи [1]. Это свидетельствует о низкой доступности своевременной диагностики и коррекции логопедических нарушений у детей дошкольного возраста. Сокращение числа обязательных диспансеризаций, острый дефицит профильных специалистов и высокая стоимость очной первичной диагностики существенно ограничивают доступ семей к своевременной логопедической помощи [2]. Существующие зарубежные логопедические программные решения не адаптированы для русского языка из-за различий в фонемной системе, а российские аналоги не предоставляют автоматизированной оценки навыков звукопроизношения в соответствии с возрастными логопедическими нормами. В связи с этим целью данного исследования является изучение методов автоматического распознавания фонем русского языка, их интеграции в процесс автоматизированной диагностики звукопроизношения у детей дошкольного возраста с использованием программных решений.

Основная часть

В рамках исследования были проанализированы научные работы в области онтогенеза речи дошкольников. На основе классических (А.Н. Гвоздев, М.Е. Хватцев) и современных (М.Б. Елисеева, Н.В. Нищева, Т.Б. Филичева) трудов по логопедии была сформирована методология диагностики и таблица возрастных нормативов формирования звукопроизношения. Алгоритм диагностики предполагает проверку произношения звуков в различных позициях (изолированно, в начале, середине, конце слова и в стечении согласных).

Для реализации автоматической оценки анализировались методы распознавания фонем русского языка в аудиоматериалах. Традиционные методы, такие как скрытые марковские модели (HMM) и гауссовские смеси (GMM), эффективны в условиях ограниченных вычислительных ресурсов. Однако для достижения более высокой точности целесообразно использование архитектур на базе нейронных сетей: рекуррентных (RNN), сверточных (CNN) сетей и Трансформеров. Одной из наиболее перспективных является архитектура self-supervised моделей (например, wav2vec2-Large-Ru-Golos), которая показывает низкий уровень ошибки (WER 10.144%) на русскоязычных датасетах (Golos, SOVA Dataset) [3-4].

Определено, что основной проблемой интеграции данных моделей в детские приложения является снижение их точности при распознавании специфичной детской речи. Детские акустические корпуса, такие как CHILDRU, имеют недостаточный объем для обучения сложных моделей с нуля, что требует применения стратегий дообучения [5]. Кроме того, базовые модели лишь транскрибируют речь, не делая проверки на предмет соответствия логопедическим нормам.

Выводы

Результатом исследования стало формирование комплексных требований к системе автоматизированного скрининга речи. Доказана необходимость адаптации

систем распознавания речи не только акустически, но и алгоритмически — путем интеграции с верифицированными логопедическими нормами и методиками. Дальнейшие работы будут направлены на реализацию прототипа системы в соответствии с требованиями и проведении тестирования разработанных алгоритмов распознавания фонем на целевой аудитории.

Литература

1. Около 60% детей в РФ приходят в первый класс с нарушениями устной речи - эксперт // Независимая газета [Электронный ресурс]. — URL: <https://www.ng.ru/news/721700.html> (дата обращения: 15.01.2026).
2. Единая межведомственная информационно-статистическая система (ЕМИСС) [Электронный ресурс]. — URL: <https://www.fedstat.ru/indicator/62414> (дата обращения: 15.01.2026).
3. Sber Golos Dataset [Электронный ресурс] / SberDevices. — 2023. — URL: <https://sberdevices.ru> (дата обращения: 15.01.2026).
4. SOVA Dataset [Электронный ресурс] / SOVA AI. — 2023. — URL: <https://sova.ai> (дата обращения: 15.01.2026).
5. Тхал, Л.А. Корпуса детской речи «INFANT.RU», «INFANT.MAVS», «CHILD.RU», «ЕмоChildRu» на материале русского языка и их использование в исследованиях речевого онтогенеза [Текст] / Л.А. Тхал // Вестник Амурского государственного университета. Серия: Филология. — 2017. — № 1. — С. 28–58. — URL: https://lingua.amursu.ru/upload/lingua/2017/1/thal_2017_1_03_028_058.pdf (дата обращения: 15.01.2026).