

УДК 004.934.5

**МЕТОДЫ АВТОМАТИЧЕСКОГО АННОТИРОВАНИЯ РЕЧИ БЕЗ
ИСПОЛЬЗОВАНИЯ РАЗМЕЧЕННЫХ ДАННЫХ**

Зорькина А. А. (Университет ИТМО)

Научный руководитель – к.ф.-м.н. Рыбин С.В. (Университет ИТМО)

Введение. Просодическое аннотирование заключается в разметке речевых аудиозаписей для идентификации и анализа просодических характеристик и является одной из важнейших составляющих в разработке системы синтеза речи [1]. Однако большинство существующих методов аннотирования речи требуют использования больших объемов размеченных данных, сбор и обработка которых является ресурсоемкой задачей. В связи с этим, исследование методов автоматического аннотирования речи без использования размеченных данных, особенно в контексте просодической аннотации, представляет собой актуальную задачу.

Основная часть. На данный момент интонационная просодическая разметка является одним из важнейших этапов синтеза речи. Ее цель состоит в том, чтобы с помощью ограниченного набора единиц максимально адекватно отразить интонацию естественного речевого сигнала. В связи с этим, специалистами было разработано несколько видов интонационных моделей и различных алгоритмов для их получения. В рамках данной работы подробно рассматриваются алгоритм моделирования интонации MoMel и система аннотирования INTSINT [2], проводится их сопоставление с моделями, специфичными для русского языка. Результаты реализации алгоритма тестируются на ограниченном наборе фраз для русского языка.

Выводы. В ходе данной работы был подробно рассмотрен алгоритм аннотирования просодических характеристик. Данный подход представляет особый интерес для дальнейшего изучения, поскольку не требует большого объема размеченных данных, дефицит которых наблюдается в русскоязычном сегменте.

Список использованных источников:

1. Tan X. et al. A survey on neural speech synthesis //arXiv preprint. 2021. arXiv:2106.15561.
2. Daniel J Hirst Prosodic Theory and Practice. - MIT Press, 2022. 464 с.