

**ПРИМЕНЕНИЕ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА  
В ПРЕПОДАВАНИИ ФОНЕТИКИ В ЛИНГВИСТИЧЕСКОМ ВУЗЕ:  
ВОЗМОЖНОСТИ И ОГРАНИЧЕНИЯ**

Стремительное развитие технологий искусственного интеллекта (ИИ) обусловило их внедрение в педагогическую практику, которые сопровождается как очевидными преимуществами, так и трудностями, естественными при освоении всего нового. Общеизвестным является большой образовательный потенциал названных технологий в обучении иностранным языкам. Наиболее популярным является использование ИИ для создания учебных материалов по заданным параметрам (вид задания, тематическая направленность, уровень обучающихся), для автоматизации некоторых рутинных процессов, например, при проверке выполнения упражнений и получения развернутой обратной связи в виде комментирования ошибок, редактирования текста, для обеспечения доступа к информации и т.д. Тем не менее применение данных технологий имеет ограничения, которые определяются собственно учебной дисциплиной, ее целями, задачами и этапом обучения. В данной публикации мы проанализируем возможные способы использования ИИ для решения узконаправленной задачи, т.е. для преподавания фонетики в профессиональном вузе, где к результатам учебной деятельности предъявляются максимальные требования.

Преподавание фонетики как учебной дисциплины, как известно, сопровождается использованием технических средств, поскольку обучение произношению предполагает прослушивание и имитацию образца, в качестве которого выступает речь носителя того или иного языка, а также запись и прослушивание собственной речи. Таким образом, первой и главной технологией искусственного интеллекта, применяемой для целей получения образца речи носителя языка, является автоматический синтез речи. Тем более что при создании данной технологии с самых первых дней использовались достижения фонетики, без которых само существование автоматического синтеза было бы невозможным. Автоматический синтез уже применяется для записи дидактических материалов по иностранному языку для школ и нелингвистических вузов, а также при создании учебных пособий по английскому языку как второму иностранному в МГЛУ, поскольку качество синтезированной речи по своим сегментным характеристикам уже неотличимо от естественной речи носителя языка. Очевидно, что данная

технология обладает рядом достоинств, которые обуславливают ее применение в преподавании фонетики. Во-первых, синтез решает проблему технического качества аудиозаписи. Во-вторых, современные синтезаторы позволяют выбрать конкретный акцент: британский произносительный стандарт RP, американский или австралийский стандарт, ирландский английский, благодаря чему на примере одного и того же текста можно наглядно продемонстрировать фонетические сходства и различия национальных и региональных вариантов языка. Данное свойство открывает широкие возможности применения программ синтеза в преподавании теоретической фонетики, в учебно-исследовательской деятельности студентов и разработке курсов повышения квалификации для преподавателей. Во-вторых, пользователь может выбрать пол, возраст, тембр и индивидуальную манеру говорения: уверенную, решительную, ласковую, экспрессивную и др. Современный синтез, в отличие от более ранних программ, демонстрирует вариативность просодических характеристик, в том числе достаточную степень экспрессивности, что сближает его с естественной речью.

Казалось бы, при наличии такого количества преимуществ, программы автоматического синтеза речи можно использовать в практике преподавания без каких-либо ограничений. Мы поставили себе задачу провести предварительную оценку возможности использования программ автоматического синтеза в качестве замены записей учебных текстов и диалогов, многие из которых технически или морально устарели.

Мы синтезировали различные языковые примеры при помощи широко известной программы Elevenlabs и оценили их с точки зрения использования в качестве образцов для имитации, исходя из собственного опыта преподавания практической фонетики английского языка. Мы сделали запись фрагмента монологического описательного текста “Our Sitting-room”, используемого в практике преподавания уже более 40 лет в теме «Интонация нефинальных частей высказывания». Результат оказался вполне удовлетворительным: просодическая сегментация текста программой практически неотличима от естественной речи, просодические структуры, типичные для неконечных синтагм, также были реализованы и соответствовали контексту. Таким образом, аудиозаписи повествовательного или описательного текста могут быть использованы как звуковые иллюстрации, как минимум, явления просодической сегментации. Однако того же нельзя сказать о диалогах с участием двух или нескольких говорящих. Во-первых, чтение диалога синтезатором не позволяет отразить различные модально-прагматические значения, передаваемые в диалоге при обмене репликами с собеседником: одобрение, возражение, сомнение, предупреждение, ирония и т.д. То есть обучение модально-прагматическим значениям просодических структур на примере синтезированной речи не будет достаточно эффективным. Синтез реплик из учебного диалога, в оригинальной записи которого выражено значение сомнения при помощи нисходяще-восходящего тона, не показал

данное значение, заменив его простой констатацией. Во-вторых, слабым местом программ синтеза до сих пор остается фразовая акцентуация. Следует подчеркнуть, что лучшие современные программы сделали значительный шаг вперед и учитывают контекстные особенности при расстановке фразовых ударений. И тем не менее, на наш взгляд, в синтезированной речи отсутствует четкое и последовательное разграничение по степени просодической выделенности новой и уже известной информации: информация, потерявшая значимость в контексте, звучат так же, как если бы она упоминалась впервые. Например, в следующем диалоге из произведения «Алиса в Стране Чудес», где по контексту требовалось смещение ядерного акцента в ответе на вспомогательный глагол (*I don't know* – ‘Я не знаю’), синтезатор оставил его на лексическом глаголе:

*“Oh, I am not particular as to size,” Alice hastily replied. “Only one doesn’t like changing so often, you know.” “I don’t know,” said the Caterpillar.*

В-третьих, сохраняется проблема звучания омографов. Наконец, звучание синтезированного текста зависит от индивидуальной манеры чтения диктора, на чьем голосе обучалась программа. В случае необходимости специального обучения конкретным просодическим моделям (например, восходящему тону в общих вопросах) требуется пробовать синтез на основе нескольких вариантов дикторских голосов, пока нужная структура не будет реализована.

Мы отметили некоторые наиболее существенные достоинства и недостатки технологий автоматического синтеза речи с точки зрения их использования для создания звуковых образцов на занятиях по фонетике иностранного языка. С нашей точки зрения, такое применение оправдано как в учебном процессе, так и в учебно-исследовательской работе, для сравнения естественной и синтезированной речи с целью обнаружения и интерпретации сходств и различий, нахождения ошибок, несоответствий и т.д. Таким образом, в настоящее время современные технологии автоматического синтеза в преподавании фонетики иностранного языка могут быть дополнительным средством обучения, однако главным образцом для имитации должны оставаться записи естественной речи носителей иностранного языка.