

ОТЗЫВ
официального оппонента
о диссертации Соломенник Анны Ивановны
«Оценка качества селективного синтеза речи: методы и результаты»,
представленной на соискание ученой степени
кандидата филологических наук
по специальности 10.02.21 Прикладная и математическая лингвистика

Диссертация Соломенник А.И. посвящена проблеме объективной оценки результатов работы селективных синтезаторов речи для русского языка. Актуальность работы не вызывает сомнений, т.к. посвящена наиболее распространенным в настоящее время типам синтезаторов, и для русского языка к настоящему времени недостаточно разработана.

Предложенные автором методы охватывают все этапы синтеза, достаточно полно проанализированы в работе, что свидетельствует о хорошей информированности автора в рассматриваемой области и его высоком уровне квалификации в компьютерной лингвистике в целом.

Диссертационная работа насчитывает 195 страниц, содержит 21 рисунок и 22 таблицы. В 19 приложениях содержатся тестовые тексты для предложенных в диссертации способов оценки и результаты экспериментов.

Во введении к диссертации сформулирована основная цель исследования – разработать комплекс методов оценки качества селективного синтеза речи на материале русского языка, и задачи, которые потребовалось решить в процессе ее достижения. Материалом для исследования явилась синтезированная речь, полученная с помощью нескольких современных

русскоязычных селективных синтезаторов, список которых приведен автором. Также здесь обосновывается актуальность проделанной работы и отмечается ее практическая значимость. Материалы работы прошли апробацию на ряде международных конференций, проходивших с 2009 по 2015 гг. По теме диссертации опубликовано 16 научных работ, из них 14 статей, 7 из которых опубликовано в журналах из перечня ВАК.

Основной текст диссертации состоит из четырех глав, разбитых на разделы и содержащих краткие заключения.

В первой главе очень подробно описана история синтеза речи, применяемые решения в технологии синтеза в разные периоды развития, рассмотрено изменение целей задач, а также требований к качеству в процессе совершенствования данной технологии. Отмечены факторы, влияющие на развитие технологии синтеза: развитие техники, а также соответствующих разделов прикладной и теоретической лингвистики. Отдельное внимание уделено перспективам развития технологии.

Во второй главе дается описание технологии исследуемого в диссертации селективного синтеза речи. Селективный синтез речи является разновидностью конкатенативного синтеза, то есть при генерации речевого сигнала используются заранее полученные звукозаписи естественной речи. Автором отмечаются особенности этого синтеза по сравнению с другими применяемыми способами получения синтезированной речи. Сделанное описание этапов селективного синтеза позволяет выявить проблемы, которые отличают данный тип синтеза от других, чем обосновывается необходимость проведения специальных видов тестов, разработанных с учетом этих особенностей, а наличие общих этапов в разных технологиях позволяет использовать разработанные ранее тесты.

Также во второй главе приведен список синтезаторов речи для русского языка, на основе которых были проведены эксперименты по тестированию.

Список снабжен описанием особенностей конкретных синтезаторов и особенностей получения материалов для тестирования, который может оказаться полезным в подобных исследованиях.

Третья, основная, глава посвящена описанию структуры предлагаемой системы оценки качества селективного синтеза речи, дается обоснование возможности использования существующих методов оценки и необходимости разработки специфических методов. В первом разделе приводится список задач, которые позволяет решить разработанная система оценки качества синтезированной речи. К ним относятся: тестирование системы синтеза в процессе ее разработки, оценка собственной системы синтеза речи в сравнении с конкурентами, участие в конкурсах, проводимых независимыми компаниями. Во втором разделе данной главы рассматриваются методы оценки разборчивости речи. Отдельное внимание в главе посвящено задаче, которая становится наиболее важной на современном этапе развития синтезаторов речи – методам оценки естественности синтезированной речи. В работе подчеркивается тот факт, что наряду с субъективными методами оценки может использоваться и инструментальная оценка, позволяющая оценить качество звучания речи. К сожалению, в работе нет подробного описания и анализа способов получения таких оценок. В четвертом разделе главы «Методы оценки качества селективного синтеза речи» обсуждаются факторы, влияющие на восприятие синтезированной речи человеком. К ним относятся: условия окружения, ограничения, присущие системе обработки информации, которой обладает человек, опыт и подготовленность слушателя, лингвистическая структура сообщения, качество звука. В пятом разделе главы обосновывается необходимость адаптации общих методов оценки качества к селективному синтезу речи. Речь, порождаемая селективным синтезатором, имеет специфические особенности. Это, прежде всего, неравномерность распределения мест с неудачным звучанием: нередко отдельная фраза или ее часть звучит гораздо естественнее остальных, а при стыковке «гладких»

участков появляются помехи. Указанные особенности связаны с базовым алгоритмом выбора звуковых единиц. В данном разделе подчеркивается, что следует разграничивать причины возникновения ошибок в лингвистической обработке, связанные с работой лингвистического процессора, и ошибки, появившиеся вследствие неудачно подобранных звуковых элементов для конкатенации. Для селективного синтеза невозможно, как, например, для аллофонного или дифонного конкатенативного синтезатора, составить тест, содержащий все или большинство элементов его речевой базы для тестирования их звучания, так как сегментные единицы языка (фонемы и их аллофоны) в базе будут представлены не одним, а, возможно, сотнями или даже тысячами вариантов. В шестом разделе главы рассматривается общая структура системы комплексной оценки селективного синтеза речи. В соответствующих подразделах рассматриваются тесты для отдельных модулей синтезаторов, а также интегральной оценки качества синтезированной речи. Даются конкретные рекомендации по составлению и проведению сравнительного и диагностического тестирования синтезаторов.

В четвертой, экспериментальной, главе описаны эксперименты и тесты, проведенные в диссертационном исследовании по оценке качества нескольких современных систем селективного синтеза речи на русском языке, дается анализ полученных результатов. Эксперименты проводились на материалах полученных с использованием современных русскоязычных синтезаторов речи.

Первый эксперимент посвящен оценке влияния различных типов ошибок на общее качество синтезированной речи. В нем делается попытка оценить, какие ошибки наиболее распространены в современных селективных синтезаторах высокого качества и какие из них вызывают наибольшие проблемы при восприятии синтезированной речи, заставляя слушающих оценивать ее как менее естественную. Из проведенного теста автор делает вывод, что ошибки в интонационном оформлении синтезированной речи являются главной проблемой современных русскоязычных селективных

синтезаторов и тестируемые синтезаторы отличаются по качеству лингвистической и акустической обработки, причем большее влияние на ухудшение естественности имеют ошибки, связанные с неправильной постановкой словесного ударения и неверной транскрипцией.

В следующем разделе обсуждается серия из нескольких тестов модуля лингвистической обработки. Анализируется точность выделения предложений, оценка чтения аббревиатур, цифровых обозначений, специальных символов, иностранных слов на латинице и правильности определения места ударения. Общие результаты исследований, описанных в данном разделе, как и отмечает автор, являются безусловно ожидаемыми. Отдельный раздел посвящен оценке модуля фонетической обработки, в нем одновременно проверяются результаты работы просодической и фонетической обработки, при этом автор отмечает, что без доступа к результатам работы соответствующих модулей невозможно протестировать правильность интонационной транскрипции независимо от результирующих физических характеристик речи. В данном разделе приводятся результаты тестов правильности сегментной транскрипции, паузирования и интонации. В четвертом разделе «Оценка акустической обработки» даются рекомендации по оценке модуля выбора звуковых элементов из речевой базы и модуля модификации. Автор посчитал невозможным провести их при использовании только демо-версий синтезаторов.

Заключительный пятый раздел посвящен обсуждению интегральной оценки качества синтезированной речи, в нем описывается эксперимент по сравнительной оценке качества речи по ГОСТ Р 50840-95 «Передача речи по трактам связи. Методы оценки качества, разборчивости и узнаваемости», а также приводятся данные по зависимости оценок от «знакомства» аудиторов с синтезом речи.

В заключении автор дает краткую характеристику основных разделов работы, приводит основные результаты, полученные в рамках

диссертационного исследования и рассматривает перспективы дальнейшей разработки темы.

Автореферат и публикации соответствуют содержанию диссертации.

Из недостатков работы можно отметить:

1. Чрезмерную увлеченность автора изложением исторического аспекта, и как следствие, увеличенный объем соответствующего раздела работы.
2. Некоторую неаккуратность при подготовке иллюстраций. Рисунки 3, 5, 11 на страницах 14, 17, 27 несколько перекошены.
3. Названия двух разделов работы, «3.1 Задачи и критерии оценки качества синтезированной речи» и «3.6 Структура и задачи системы оценки качества синтезированной речи» по смыслу практически идентичны, при этом в последнем пункте речь идет о системе оценки качества именно селективного синтеза речи.
4. В разделе 4.3.3 автор изначально говорит о показателе «адекватности» интонации, однако затем (на стр. 118) использует термин «точность» для того же явления.
5. Незначительные технические ошибки в автореферате (стр. 20 «в ходе рамках диссертационного исследования»)

Сделанные замечания не снижают общую высокую оценку диссертационной работы А. И. Соломенник. Диссертация соответствует критериям, сформулированным в пп. 9, 10 «Положения о порядке присуждения ученых степеней» ВАК Министерства образования РФ и является научно-квалификационной работой, в которой содержится решение задачи, имеющей существенное значение для области прикладной лингвистики и речевых технологий, а ее автор, Анна Ивановна Соломенник, заслуживает присуждения

ей искомой степени кандидата филологических наук по специальности 10.02.21
– Прикладная и математическая лингвистика.

Кузнецов В. Б., *Кузнецов* 12.05.2016

канд. филол. наук, профессор кафедры прикладной и
экспериментальной лингвистики

федерального государственного бюджетного

образовательного учреждения высшего образования

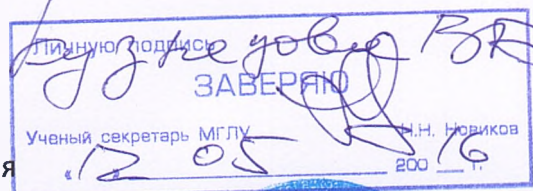
«Московский государственный лингвистический
университет»

119034, Москва, ул. Остоженка, д. 38.

тел: (499) 245 06 12, (495) 637 56 97

сайт: <http://www.linguanet.ru>

e-mail: kuvlad2007@yandex.ru



К теме диссертации относятся следующие публикации оппонента:

Автоматический синтез речи. Алгоритмы преобразования "Буква-звук" и
управление длительностью речевых сегментов. (монография) // Таллинн.
Валгус. - 1989. Соавтор - Отт А.

Лингвистическое обеспечение систем синтеза речи по правилам: достижения,
проблемы и перспективы (монография) // ИНИОН РАН, М., - 1992

Почему русский дрожащий трудно синтезировать? (на англ. языке) // Сборник
трудов Международного семинара "Речь и компьютер" (SPECOM'2003),
Москва, 27-29 октября 2003 г.