

Отзыв о диссертации
на соискание ученой степени кандидата филологических наук
Елены Владимировны Моисеевой
"Реализация гласных после мягких согласных
на стыках слов в современном русском языке",
специальность 10.02.01 - русский язык.

Диссертация Е.В. Моисеевой посвящена одному из ключевых участков русской фонетической системы. Литературный вокализм находится под пристальным наблюдением лингвистов уже более века, но до сих пор предлагает задачи, требующие решения специалистов. Один из таких ответов на вопрос о функционировании фрагмента безударного вокализма дает рецензируемая диссертация.

Работа имеет установку на практическое применение результатов исследования, необходимых для создания и совершенствования систем автоматического распознавания и синтеза речи. Нужно заметить, что современные системы синтеза речи, такие как телекоммуникационный продукт ЦРТ (Центр речевых технологий, <http://www.speechpro.ru/>) или Nuance (<http://www.nuance.com/>), которым пользуются крупнейшие банки и операторы сотовой связи для генерации звучащих текстов из орфографических ("текст-в-речь"), имеют существенные недостатки в том числе в части реализации качества безударных гласных в неканонических случаях. Даже сложные артикуляционные модели американского исследователя Бреда Стори имеют существенные сложности, касающиеся реализации такого «композита» как линейное сочетание звуков в речи.

Практическая часть диссертации Е.В. Моисеевой предваряется замечательной «историей вопроса» во введении и первой главе. Эта часть претендует на крайне содержательный и даже справочный по своему характеру раздел в учебнике фонетики. Автор дает все необходимые современные сведения не только о литературе, которая поможет читателю понять место диссертационного исследования в кругу прочих фонетических трудов, но эти сведения также позволят расширить кругозор в области использования программ акустического анализа речи. История вопроса, опубликованная отдельно, может быть крайне полезна студентам-лингвистам и филологам.

После широкого вступления у читателя возникает вопрос о непосредственной сути самого диссертационного исследования и на с. 39 автор ставит яркую точку в определении значимости и места работы в ряду прочих фонетических проблем: «С одной стороны, позиция в начале слова способствует сохранению тембра гласного, с другой стороны, предшествующий мягкий согласный обуславливает изменение формантной структуры гласного. Изучение влияния взаимодействий этих

двух факторов на формантную структуру начального гласного слова является целью настоящей работы».

Детальное и весьма объёмное описание множества факторов, влияющих на качество гласного, формирует у читателя внутреннее переживание, так как хорошо известна сложность многофакторного описания гласных. Однако Елена Владимировна на с. 40 предлагает разграничение факторов и объясняет экспериментальную базу исследования: «Некоторые из них были учтены при подборе и анализе материала, а именно: положение по отношению к ударению, положение в сильной или слабой фразовой позиции, интонационное оформление высказывания, особенности идиолекта и пол диктора. Влияние остальных факторов – правого сегментного контекста, места образования предшествующего согласного, эмоционального состояния диктора – мы постарались максимально уменьшить».

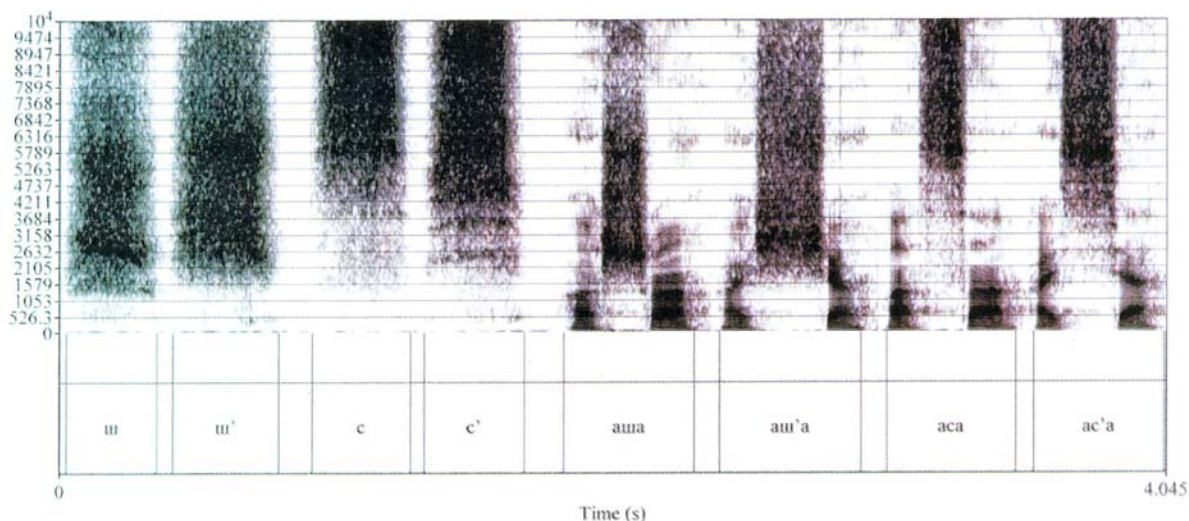
Спектрографический анализ и расчет формантных данных автор проводил силами программы акустического анализа звука PRAAT. Так на с. 45 размещен рисунок 6, осциллограмма и спектрограмма фрагмента слова *обелиск* в произношении диктора-женщины, где формантные переходы происходят скачкообразно. Эта особенность формантной структуры женских и детских голосов делает анализ формант затруднительным и заставляет отводить часть материала. Однако надо полагать, узкополосная спектрограмма с настройками шага анализа 0,0125 могла бы снять часть графических накладок. В этом отношении не хватает данных о настройках параметра *spectrum*, а также *formant*, которые использованы автором. Настройки существенно влияют на выходные данные и знать их необходимо для понимания того, какой сигнал после обработки программой получил автор исследования. Стандартные настройки программы могут скрывать расположенные рядом форманты (или возникающие «паразитные» форманты), объединяя их в одну в окне спектрограммы, что приведёт к «фантомному» браку материала.

Не лучшее качество иллюстраций из программы Praat может быть отмечено также на рисунках 7-8. Иллюстрация 7, судя по спектру, представляет мужской скрипучий голос, форманты разбиты на равные краткие периоды. Такой материал не мешает глазу видеть формантные треки, но предъявляет особые требования к выбору места вычисления формант на треке. Рисунок 8 показывает речь женщины с низкой собственной энергией артикуляции. Это тихий голос, потому, как совершенно верно замечает Е.В. Моисеева, формантные максимумы реализованы слабо или смазано. Высокий внешний шум и низкая энергия голоса делают иллюстрацию неубедительной с точки зрения качества картинки. Кроме того, в работе «мужские» картинки всегда громче, чем шумные «женские», что говорит о неудовлетворительных технических

условиях записи, которая, вероятно, выполнялась прямо на компьютер в программе Praat.

Есть и другие вопросы по работе автора с программой акустического анализа звука. Эти вопросы имеют целью уточнить выбор способа подачи графического материала, но не должны приравниваться к фактическому замечанию.

На с. 64, например, следует признать не совсем удачными иллюстрации из Praat, сделанные в режиме фотографии основного рабочего экрана программы. На такой картинке отсутствует частотная разметка, чтение иллюстрации и определение формантных значений, которое выполняет читатель глазом, превращается в гадание. Отсутствие качественных подписей также мешает читать текст. Преодолеть это несовершенство диалогового окна программы можно, используя вшитую функцию "textgrid" и "praat picture". При таком оформлении иллюстративный материал может стать более «дружелюбным» к читателю.



На с. 80 описывается методика подсчета результатов, то же и во второй главе. Выводы о формантной структуре гласного строятся на средних арифметических значениях в произношении информантов. Несмотря на подобранный контекст и единообразное просодическое оформление звуковых фрагментов (такова установка эксперимента) форманты могут иметь значительные расхождения от примера к примеру. К этому должны быть добавлены возможные ошибки расчета формант в программы Praat. Насколько необходимо знать среднее арифметическое значение, игнорирующее частные отклонения?

Продолжая разговор о некоторых технических недостатках работы, стоит сказать, что на с. 56, например, смущают подписи к спектрам в орфографии. Хотя можно догадаться, что это сделано для удобства восприятия, но, тем не менее, этим стоит пренебречь, так как диссертация рассчитана на чтение специалистами. То же на с 96-97. "В СРЛЯ в первом

предударном слоге в абсолютном начале слова смысловозначительной является оппозиция трех гласных [а], [и], [у]: унёс – инёс – анёс [ун'о́с] – [ин'о́с] – [ан'о́с], а в позиции после мягкого согласного внутри слова – только двух: *тюки – теки, люблю – леплю*". Из-за отсутствия транскрипции в двух последних парах примеров может возникнуть ощущение, что орфографическая подача материала акцентирует различие гласных [у]-[э].

Есть и другие моменты, которые не совсем понятны и требуют разъяснения автором.

С.76. «Абсолютная частота формант гласного непервого предударного слога *приблизительно* вычисляется по формуле $F(\#_k) - (F(\#_k) - F(\#t'_k))/2$ в положении перед твердым заднеязычным согласным, но для положения между мягкими согласными эта формула нуждается в корректировке». Приблизительное вычисление по формуле, кажется, затруднит её применение.

То же на с.87, таблицы 12-13. Описание таблиц говорит о том, что «Первая же форманта испытывает влияние не столь единообразное, **вероятно**, существуют и иные факторы, влияющие на ее частоту, кроме соседних согласных». В свете того, что актуальность работы определяется необходимостью исчерпывающего описания акустических свойств гласных звуков с целью создания систем автоматического распознавания и синтеза речи, слово «вероятно» кажется в этом контексте чересчур осторожным и даже лишним.

Следует признать неудачной формулировку одного из вводов к 4 главе (с. 128). "В идиолектах дикторов спектральные характеристики гласных на месте фонем <a>, <o> и гиперфонемы <a/o> сильно различаются, причем это различие сохраняется как в ударном, так и в безударном положении". Следуя этому выводу можно подумать, что существует различие в реализации фонем <o>, <a> и гиперфонемы <o/a>, то есть некое "оканье" в широком смысле, хотя ясно, что автор подразумевал совершенно иное явление.

Интерпретация некоторых данных вызывает вопрос. На с. 88: «Мягкий согласный предшествующего слова повышает значение F2 на 35% в идиолекте диктора 1 и 15–18% в идиолекте диктора 2». Каким образом можно употребить столь широкие расхождения данных для автоматического распознавания и синтеза?

С чем связаны пробелы в табличных данных в приложении, что значит отсутствие релевантных примеров и не влияют ли они на статистическую картину в вычислении среднего значения? Эти частные вопросы и замечания требуют ответа, но не влияют на общее качество работы.

В своих выводах автор следует внутренней логике развития исследования. Гипотеза о связи длительности гласного и качественной редукции рассматривается и отводится как нерабочая. Обычно же в диссертационных исследованиях даются лишь положительные выводы. Честная позиция отрицающего автора заслуживает всяческого одобрения: «Объяснить изменение спектральной картины гласного недостижением целевой артикуляции вследствие сокращения длительности в положении после согласного предшествующего слова на исследованном материале не удалось».

Один из самых сложных пунктов экспериментально-фонетического исследования заключается в выборе точки снятия формантных значений. Решение такой задачи уже само по себе является значительным исследованием и приходит только после работы с большим количеством спектров. В работе Е.В. дается два ответа для разных условий. Так на с. 84-86 делается убедительный выбор точек измерения и сопоставления формант. «<...>при анализе гласного первого предударного слога (как и гласного непервого предударного слога) было принято решение сравнивать частоты формант на начальном отрезке гласного после паузы (а не в его середине) с частотами на центральном участке гласных между двумя согласными или в точке экстремума (если таковая имеется)».

Важным для понимания достигнутого результата является дважды подтвержденное формантное развитие гласных в положении перед заднеязычными и палатализованными согласными (с. 90-93). "<...>F1 и F2 [a]-образного гласного в положении перед твердым заднеязычным (*сломать акацию*) занимают по частоте положение практически ровно посередине между соответствующими значениями формант гласного абсолютного начала слова и гласного первого предударного слога после мягкого согласного, как и в непервом предударном слоге. Что касается позиции перед мягким согласным (*заслать агента*), то первые две форманты находятся по частоте ближе к формантам гласного после мягкого согласного внутри слова, что тоже коррелирует с результатами, полученными для непервого предударного слога". А следующая далее таблица 15 дает представление о *соотношении* частот F1 и F2. Эти относительные значения, может быть, функциональнее и содержательнее, чем изменчивые частотные характеристики первых двух формант.

Неполное совпадение формант в однородном контексте хорошо известно фонетистам. Думается, что важнейшим ответом на существенный вопрос о природе несоответствий будет тот, который позволит понять, почему столь существенная разница позволяет нашему мозгу отождествлять различные звуки в разных идиолектах в один. Этот ответ важен не только для понимания фонологической и психофизиологической сущности явления, но, прежде всего, для безошибочной машинной

обработки входящего сигнала и формировании исходящего, то есть для автоматического распознавания и генерации звучащей речи.

Один из самых замечательных выводов главы 3 (с. 99) является предположение (хотя и выполненное на малом материале) об отсутствии связи длительности, тембра и реализацией целевой артикуляции. Этот соблазнительный механизм объяснения редукции артикуляционным "недострелом" теперь требует отдельного осмысления. Очень важно, что практическая по своей сути работа ставит теоретические вопросы перед русской фонетикой.

В разделе 4.3.7. "Интерпретация коартикуляции на стыках слов в рамках интегральной модели звукового поведения (лексической фонологии)" (с. 124) Е.В. Моисеева применяет правила "горизонтальной" и "вертикальной" фонетики для объяснения последовательных изменений (реализации) звуков. Этот раздел снимает часть вопросов, возникших ранее. Вариативность поведения начального гласного слова, которая устанавливается экспериментально, получает теоретическое объяснение.

Между тем хотелось бы понять, где в стройной модели порождения слова автор видит зазор, возможность колебания для осуществления фонетической эволюции, или, если угодно, для смены моделей.

Нужно сказать, что Е.В. Моисеевой удается сочетать два теоретических подхода к оценке фактов языка. Иерархические отношения единиц, очевидно, оцениваются в рамках Московской фонологической школы, а последовательность реализации вариантов и вариаций речи оценивается в рамках лексической фонологии.

Завершая оценку работы Е.В. Моисеевой необходимо сказать, что *актуальность* работы не вызывает сомнений, так как определяется необходимостью описания акустических свойств гласных звуков с целью создания систем автоматического распознавания и синтеза речи. *Научная новизна* диссертации также очевидна, она состоит в выявлении закономерностей функционирования гласных фонем, а также теоретического объяснения наблюдаемых фактов, что раньше не находило системного описания. *Теоретическая значимость* определяется применением современной объяснительной системы к фактам языка, а *практическая значимость* диссертации заключается в том, что данные, полученные в исследовании, могут быть использованы для создания фактической базы акустических образов русских безударных гласных, которые могут использоваться в преподавании, машинном переводе и т.д.

Автореферат диссертации Елены Владимировны Моисеевой "Реализация гласных после мягких согласных на стыках слов в современном русском языке" и публикации полно отражают текст диссертации.

Диссертация Елены Владимировны Моисеевой «Реализация гласных после мягких согласных на стыках слов в современном русском языке», представленная на соискание ученой степени кандидата филологических наук, соответствует пп. 9, 10 Положения о присуждении ученых степеней и является научно-квалификационной работой, в которой содержится решение задачи, имеющей существенное значение для специальности 10.02.01 – русский язык.

Считаю, что диссертация Елены Владимировны Моисеевой выполнена на высоком профессиональном уровне, соответствует требованиям, предъявляемым ВАК РФ к диссертациям на соискание ученой степени кандидата наук, а ее автор заслуживает присвоения искомой ученой степени.

Кандидат филологических наук,
старший научный сотрудник отдела
диалектологии и лингвистической
географии Федерального
государственного бюджетного
учреждения науки Институт
русского языка им.
В. В. Виноградова РАН.

И.И. Исаев



Сведения об оппоненте. Игорь Игоревич Исаев. 143980, Московская область, гор. Железнодорожный, ул. Рождественская, д. 8, кв. 316. Телефон: +7 915-435-76-12. E-mail: ignatis@mail.ru

Место работы. Федеральное государственное бюджетное учреждение науки Институт русского языка им. В. В. Виноградова РАН. 119019, Москва, ул. Волхонка, д. 18/2. Телефон: (+7 495) 695-26-60. Факс: (+7 495) 695-26-03. E-mail: ruslang@ruslang.ru. www.ruslang.ru. Старший научный сотрудник.

Публикации по специальности рецензируемой диссертации.

Исаев И.И. Экспериментально-фонетическое исследование гласных звуков как метод диалектной компаративистики // Экспериментальный метод исследования русского языка в вузе и школе. Владимир: ВГПУ, 2011. С. 19-26.

Исаев И.И. Артикуляционное пространство и формантная характеристика гласных в русских говорах // Исследования по славянской диалектологии, выпуск 15. М., 2012. С. 215-223.

Исаев И.И. Проблемы фонологического анализа диалектной системы согласных с многоступенчатой фонетической реализацией признака «твёрдость-мягкость» // Діалекти в синхронії та діахронії: загальнослов'янський контекст. Київ: Інститут української мови НАН України, 2014. С.194-197