

Новые подходы в исследовании базовых артикуляций русского вокализма

Фонетика и фонология на стыке веков могут быть охарактеризованы как науки, далекие от процесса затухания и ухода в небытие. ... Напротив, они обрели как бы «второе дыхание», черпая силу из источника идей новых теоретических и прикладных задач, новых речевых и ин-формационных технологий, новых потребностей в изучении вечно живой и вечно влекущей к себе звучащей речи.

Р.К. Потапова, В.В. Потапов.
Язык, речь, личность.

Речевая артикуляция, как основная форма материализации речевой деятельности, естественным образом аккумулирует в себе все базовые свойства и особенности языка – социальную и индивидуальную природу языковой деятельности, ее зависимость от общих закономерностей функционирования физиологии человека и его психологии, а также нейрофизиологические и анатомические особенности производства речи, роль которых нельзя недооценивать. Как пишет Л.Р. Зиндер: «Звуки человеческой речи – продукт действия произносительных органов и формы резонаторных полостей произносительного аппарата человека; следовательно, они определенным образом связаны с анатомией этого аппарата» [Зиндер 1984].

Такая многоаспектность природы речевой артикуляции делает ее, с одной стороны, достаточно сложным объектом исследования, однако с другой – способствует привлечению к изучению этого аспекта языковой деятельности всего широкого спектра приборов, методов и методик, разработанных в сфере естественнонаучного знания, т.е. превращает речевую артикуляцию в один из самых естественных и многообещающих объектов междисциплинарного научного исследования. Такого рода междисциплинарность традиционно рассматривается в фонетике как необходимое условие истинно научного подхода, который характерен для российской лингвистической науки. В частности, Л.В. Щерба «всегда ратовал за содружество языковедов с представителями других наук» [цит. по: Зиндер 1984], а Л.В. Зиндер с удовлетворением отмечал, что «такое содружество в наши дни уже не является редкостью» [цит. соч.]. Действительно, именно междисциплинарные, в первую очередь, акустические и физиолого-артикулографические, исследования процесса речепроизводства дали в России очень значимые результаты, которые в конце XX в. позволили блистательно реализовать целый ряд прикладных языковых разработок: таких как

автоматический синтез и распознавание речевого сигнала, вокодерные технологии и методы идентификации личности и ее состояний по речи и т.п.

В этом контексте нам представляется весьма целесообразным и продуктивным использовать при изучении речевых артикуляций уникальные возможности, которые предоставляет современная томографическая техника. Магнитно-резонансная томография (МРТ) в настоящее время является общепризнанным методом медицинской диагностики – в первую очередь, из-за высокой информативности и точности получаемых данных, а также (что немаловажно) благодаря безопасности исследования как для пациента, так и для самого экспериментатора. Вместе с тем, уникальные технические возможности современных томографов уже с конца 1980-х гг. обратили на себя внимание исследователей и из других, не связанных с медициной напрямую, областей. Примером такого нестандартного, не связанного с медицинской диагностикой использования МРТ стали эксперименты по визуализации положения артикуляторных органов при произнесении как отдельных звуков речи, так и в процессе порождения слитной речи. В первую очередь, лингвистов привлекает в этой технологии то, что с помощью томографа фонетист-исследователь может получить изображения артикуляционных процессов как бы «напрямую», без акустического (или какого-либо иного) «посредника».

Актуальность такого рода исследований обусловлена также и тем, что до настоящего времени практически нет исчерпывающих детальных достоверных описаний динамических механизмов речепорождения в разных языках, которые были бы основаны на результатах фиксации в режиме реального времени процесса развертывания речевых артикуляций. Большинство наших знаний в этой области являются в значительной степени реконструкцией, которая сформирована по косвенным, в сущности, данным (в первую очередь, по результатам акустических исследований и/или исследований с применением разного рода артикулографических методов). Однако, несмотря на то, что связь между акустикой и артикуляцией безусловно существует, в интерпретации акустической картины в артикуляционных терминах остается все же слишком много неясного. Так, С.В. Кодзасов, А.М. Красовицкий и Е.В. Щигель, изучая динамические спектрограммы гласных в русской речи, приходят к выводу, что многие вопросы, которые встают в связи с интерпретацией наблюдаемых в спектре речевого сигнала различий, до сих пор остаются без ответа. Посвященная результатам этого исследования статья заканчивается следующим замечанием: «Мы приводим здесь далеко не полный перечень тех “загадок”, который задает обследованный материал... Однако... результаты ясно показывают, что существующие представления о спектрах русских гласных нуждаются в дополнениях» [Кодзасов, Красовицкий, Щигель 2002].

Мы считаем, что для русского языка эта проблема стоит особенно остро, так как в настоящий момент работы в области томографического исследования речевых артикуляций в России в силу ряда объективных и субъективных обстоятельств существенно отстают от мирового уровня. До настоящего времени наиболее разработанным и общепризнанным средством экспериментального исследования артикуляции русской речи считалась кино(фото)рентгенография. Разработке этого метода в СССР было посвящено значительное количество работ – в этой связи нельзя не упомянуть в первую очередь фундаментальные исследования Л.Г. Скалозуб, М.И. Матусевич, Е.Э. Павловской, Н.И. Тоцкой, И.А. Стариковой и др., а также многоаспектное и разноплановое исследование артикуляторной базы русского языка венгерского лингвиста К. Боллы, которые достигли значительных успехов в применении кинофоторентгенографии в экспериментах по изучению речевой артикуляции. С помощью разработанных этими исследователями и их учениками методов кино(фото)рентгеносъемки процесса речепроизводства была подготовлена широко известная база артикуляторных моделей (артикуляторных поз) русского языка, наиболее полно представленная, в частности, в классическом справочном издании К. Боллы «Атлас звуков русской речи» [Болла 1981]. Однако, несмотря на достигнутый в этой области фонетической науки существенный прогресс, в ней все еще оставалось много вопросов и многие проблемы, в первую очередь, связанные с лингвистической интерпретацией получаемых результатов, не могли быть решены на том уровне программного и аппаратного обеспечения, который был доступен исследователю-фонетисту.

Нам представляется, что именно магнитно-резонансная томография (МРТ) позволяет преодолеть многие базовые ограничения рентгенографического метода и, благодаря ее чувствительности к сигналам от мягких тканей, получить высококонтрастные онлайн-овые изображения практически всех наиболее значимых артикуляционных органов – губ, языка, ротовой полости, нёбной занавески, гортани и др. мягких тканей – непосредственно в процессе порождения речевого сигнала. Поэтому именно этот метод – онлайн-овое МРТ-сканирование артикуляторного тракта в момент продуцирования звучащей речи – был выбран нами для достижения основной долгосрочной цели проводимого в Центре новых информационных технологий филологического факультета МГУ им. М.В. Ломоносова исследования, а именно: формирование электронных баз данных по реальным онлайн-овым изображениям основных конфигураций речевого тракта в момент продуцирования звучащей речи, в которых исчерпывающим образом могут быть представлены все типы артикуляторных жестов русского языка как в статике, так и в динамике, формирующие его артикуляционную базу в целом. Как отмечают Р.К. Потапова и В.В. Потапов, именно создание «машинных фондов и фонетических баз данных, ... более широкое использование в при-

кладных фонетических исследованиях новых информационных технологий ... являются доминирующим направлением в фонетике и фонологии конца XX в.» [Потапова, Потапов 2006]. Актуальность таких баз данных обусловлена недостаточностью достоверных подробных описаний динамических механизмов речепорождения на русском языке, что представляется тем более недопустимым на фоне активной разработки аналогичной проблематики, в первую очередь, для английского и других европейских языков, а также для ряда неевропейских (японского, корейского и китайского) языков.

На начальной стадии работ в результате экспериментального и теоретического исследования, осуществленного по оригинальной методике¹, разработанной в ЦНИТ ГО филологического факультета МГУ совместно с Центром магнитной томографии и спектроскопии МГУ (ЦМТС МГУ), нам удалось получить отсканированные в режиме реального времени контрастные, высокоинформативные и достоверно интерпретируемые МРТ-изображения, отображающие реальную конфигурацию артикуляторных органов и их динамику при произнесении основных гласных звуков русской речи.

Все эксперименты были проведены в ЦМТС МГУ на 0.5 Тл томографе Tomikon S50 («Bruker»). В МРТ-исследовании артикуляторных моделей русских гласных звуков сканирование проводилось по сагиттальному срезу шириной 9 мм и зоне сканирования – 20x12 см. Быстрое сканирование – примерно 2-2,7 кадра в секунду – получено с использованием импульсной последовательности градиентного эха. Сканирование проводилось многократными (128-256) запусками без перерывов внутри каждой серии МР-импульсов.

Известно, что использование методов МРТ в исследованиях речевой артикуляции, для которой характерны очень быстро меняющиеся во времени состояния артикуляторного тракта, связано с определенными трудностями, так как технологическая база современных томографов в большинстве случаев не позволяет вести съемку в реальном масштабе времени с требуемой скоростью. Для получения динамических изображений быстро протекающих процессов (каковым является и процесс речепроизводства) наиболее перспективным представляется применение методов магнитно-резонансной томографии с использованием специальных методик быстрого сканирования, позволяющих неинвазивно получать посрезовые изображения со скоростью порядка 10 кадров в секунду. В идеале это требует специального аппаратного и программного обеспечения, позволяющего реализовать наиболее быстрый на сегодняшний день метод МР-сканирования – эхо-планарный метод. Однако из-за невозможности реализовать данный метод на имевшемся в нашем распоряжении томографе Tomikon S50 основные наши усилия были направлены на разра-

¹ Интеллектуальный приоритет авторов в 2003 г. защищен Патентом РФ № 34756 на полезную модель «Устройство для исследования артикуляторных процессов порождения речи» от 10 декабря 2003 г.

ботку и реализацию новых технических методов оптимизации собственно процесса получения томографического изображения и совершенствования методик «псевродинамического» отображения артикуляционных движений, т.е. методик стробоскопического сбора данных. Напомним, что стробоскопический метод состоит в подборе изображений, соответствующих реальной последовательности артикуляторных движений при произнесении слова, из большого массива последовательно ведущейся с заданным временным сдвигом томографической съемки, который получен при многократном повторе заданного слова в процессе сканирования. В соответствии с этим принципом в процессе эксперимента нам удалось сформировать достаточно представительный массив томографических изображений, полученных после многократных запусков быстро сканирующей импульсной последовательности.

Стробоскопический метод подбора экспериментальных данных потребовал разработки специального сценария эксперимента и соответствующих инструкций как для дикторов, так и для самого экспериментатора. Общий сценарий МРТ-экспериментов требовал, чтобы каждый диктор, находясь внутри томографа, повторял многократно определенным образом подобранные речевые стимулы (долгий, «тянутый» звук, слог, слово/псевдослово, предложение, фраза, произнесенные в естественном темпе), при этом одновременно с фонацией речи осуществлялась последовательная томографическая съемка артикуляторного тракта с заданным временным сдвигом и фиксацией начала каждой съемки в виде специальных импульсов на двухканальной фонограмме/осциллограмме. На основе таких осциллограмм после расшифровки, временной привязки и последующей атрибуции полученных изображений проводилась реконструкция динамики изменения конфигурации артикуляторных органов в момент звучания речевых стимулов и анализ их взаимного расположения. Впоследствии точно такие же условия эксперимента и соответствующие инструкции дикторам и экспериментатору были воспроизведены при повторной съемке тех же испытуемых (дикторов) с временным промежутком в один месяц и один год.

Для двухканальной аудиозаписи процесса фонации и импульсов начала МР-сканирования использовались аппаратные и программные средства основного, управляющего томографом, компьютера INDY (фирмы SiliconGraphics). Особенностью процесса аудиозаписи было параллельное использование двух аудиоканалов (стерео-запись). На первый канал поступал сигнал от микрофона, на второй – импульс начала МР-сканирования кадра. Микрофон (фирмы LifeVideo™) крепился к приемной катушке (катушка Гельмгольца), используемой для МР-съемки шейного отдела и был расположен в непосредственной близости от рта диктора (рис. 1). Такое расположение позволило мини-

минимизировать влияние акустического шума, обычно возникающего при МР сканировании из-за вибрации катушек при пропускании через них импульсных токов. Влияние микрофона на качество съемки (из-за наличия в его конструкции кабеля и металлических компонентов, а также протекания тока) оказалось пренебрежимо малым. Для формирования импульса начала МР-сканирования каждого кадра были внесены изменения в управляющую программу сканирования и задействован резервный интерфейс. В итоге нам удалось достаточно точно соотнести каждый участок осциллограммы с конкретным кадром МРТ-изображения артикуляторного тракта, что позволило получить однозначную интерпретацию для каждого снимка.

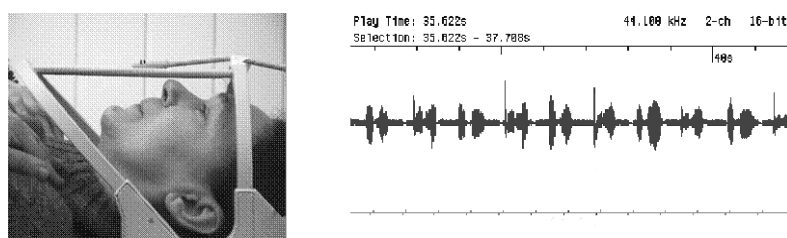


Рисунок 1. Положение испытуемого (диктора) внутри камеры томографа (микрофон закреплен на приемной катушке, используемой для стандартной МР-съемки шейного отдела). Справа: двухканальная осциллограмма речевого сигнала (верхняя линия) и начала МР-сканирования (нижняя линия).

Для дополнительного контроля нами была проведена также видео- и аудиозапись диктора (профильная проекция) в процессе артикулирования тех же самых экспериментальных речевых стимулов вне магнитной камеры томографа, при этом мы стремились к идентичности всех остальных условий эксперимента, поэтому диктор продуцировал речь в положении лежа и с так же прикрепленным микрофоном. Все виды полученных нами в ходе исследования экспериментальных данных представлены на рисунке 2.

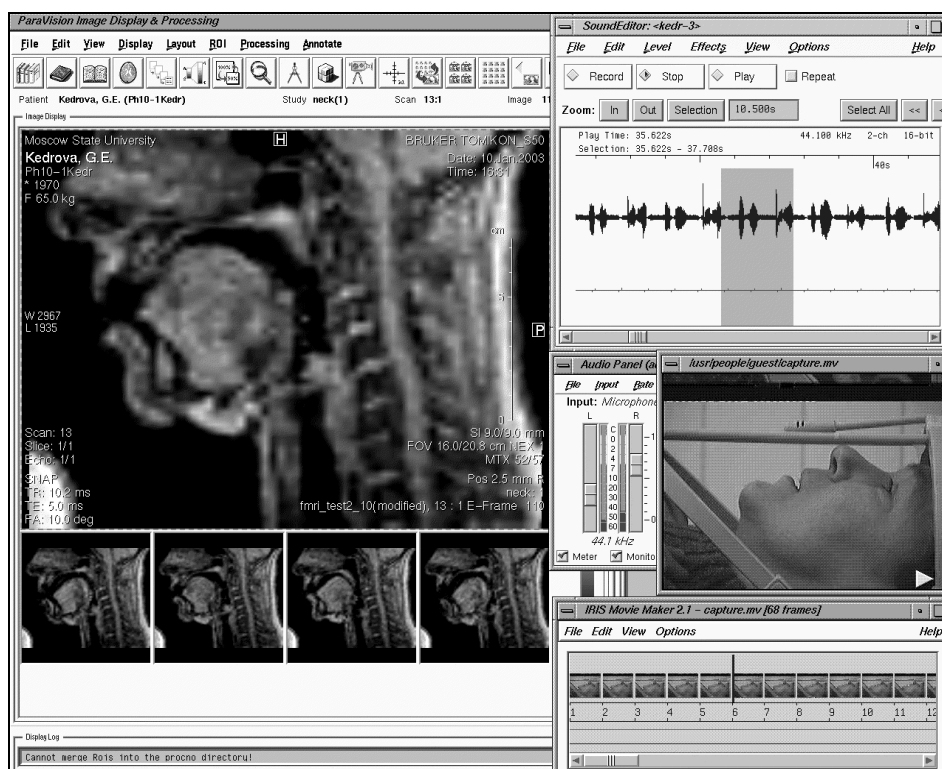


Рисунок 2. Сводный интерфейс МРТ-исследования по визуализации артикуляторных органов. Параллельно работают: программа МРТ-сканирования и программа аудио-записи. Для наглядного отображения процесса речепроизводства приведено обычное видеоизображение. Показано, что полученные МРТ-изображения представляют собой набор кадров, разделенных по времени на 0,5–0,4 сек. и отображающих одну и ту же зону сканирования.

Обратимся непосредственно к основным результатам проведенного нами совместно с ЦМТС МГУ исследования артикуляционных моделей русского вокализма. В качестве дикторов в наших экспериментах выступали 2 человека (мужской и женский голоса). В каждом из сеансов МРТ-сканирования дикторы-испытуемые получали идентичное задание, а именно: многократно (произвольное число раз) повторять с небольшой паузой и, по возможности, с ровной интонацией долгие, «тянутые»², гласные звуки: [a:], [o:], [y:], [ы:], [и:], [е:]; а также произнести без паузы их последовательность [аоуые]. Каждая серия повторов изолированных гласных звуков включала от 16 до 33 стимулов для каждого звука, что в итоге позволило получить от каждого из дикторов и проанализировать 384 МРТ-снимка процесса артикуляции гласных звуков в каждом из сеансов записи. Поскольку томографическая съемка проводилась в несколько сеансов, разделенных временным промежутком приблизительно в один месяц, а кроме того, для оценки воспроизводимости (устойчивости) моделей произнесения по каждому из дикторов аналогичные сеансы МРТ-

визуализации проводились спустя один год, в итоге мы получили для анализа более 2000 МР-изображений артикуляции гласных звуков.

В процессе подбора экспериментального речевого материала мы опирались, в первую очередь, на широко известные замечания Л.В. Щербы, высказанные им в защиту метода исследования звучащей речи через изучение изолированного произнесения речевых стимулов. Так, Л.В. Щерба пишет: «Обыкновенно мы изолируем наиболее типичную часть того гласного варианта, который наиболее часто встречается и находится в наиболее независимом положении, ... т.е. тот вариант, с которым ассоциируется фонема, ибо фонема – это не отвлеченные общие признаки группы частных представлений, ... это совершенно конкретное звуковое представление. Если так, то первая задача исследователя состоит в том, чтобы описать именно этот вариант, который, как, например, в русском языке, может представлять и отдельное слово, правомерность изолированного произнесения которого не может вызывать сомнений. Наряду с этим должны быть описаны и прочие варианты, не встречающиеся в независимом положении, а для этого исследователь должен уметь изолировать и их» [Щерба 1983]. В дальнейшем именно эти идеи Л.В. Щербы послужили для нас не только основой для формирования всех элементов экспериментального звукового массива, но и своеобразной «инструкцией» по определению последовательности этапов исследования.

Прежде чем перейти непосредственно к результатам анализа полученных МР-изображений, отметим, что для дополнительной верификации подобранных в результате вышеописанной процедуры атрибуции и идентификации изображений мы посчитали необходимым сопоставить полученные нами МРТ-снимки артикуляторных поз, соответствующих квазистационарной фазе производства гласного звука, с широко известными аналогичными изображениями, построенными с использованием метода кинофторентгенографии [Болла 1981].

Как мы и предполагали ранее, сравнение полученных нами МРТ-изображений с результатами кинофторентгенографического исследования показало, что при определенном (довольно значительном) совпадении основных параметров изображений – в первую очередь, это взаиморасположение основных артикуляторных органов – разработанный нами метод позволяет получить большую детализацию изображения, особенно в отношении мягких и/или наиболее подвижных тканей (в частности, кончика языка, его передней части и др.), что особенно значимо для описания артикуляции переднеязычных гласных и

² т.е. такие гласные звуки, которые дикторы-непрофессионалы способны без затруднений произносить в изолированной позиции на протяжении некоторого периода времени – в нашем эксперименте его длительность составляла от 1,0 сек. до 3 сек.

согласных, а также их сочетаний. Более подробно результаты сопоставления обсуждаются в [Кедрова, Захаров, Пирогов, Анисимов 2003].

Хотя сегодня тот факт, что в физиологии движений «одни и те же действия могут входить в различные деятельности и наоборот – один и тот же результат может быть достигнут путем разных действий»³, является уже общепризнанным, полученные нами результаты обнаруживают достаточно высокую степень постоянства артикуляторных конфигураций для каждого из экспериментальных речевых стимулов (гласных звуков) у каждого из дикторов. Эта стабильность артикуляционных поз сохранялась практически во всех экспериментальных сериях, т.е. как при многократности повторения каждого из речевых стимулов в однородных последовательностях реализации гласного звука и при его произнесении в окружении других гласных, так и в разные моменты времени в ходе одного и того же сеанса записи и в рамках разных сеансов, проведенных с временным сдвигом в один месяц и один год. На рисунках 3 и 4 представлены МРТ-изображения артикуляторных поз для разных реализаций одного и того же гласного звука (квазистационарных сегментов его звучания) в произнесении диктора-мужчины и диктора-женщины, полученные в ходе разных экспериментальных сессий.

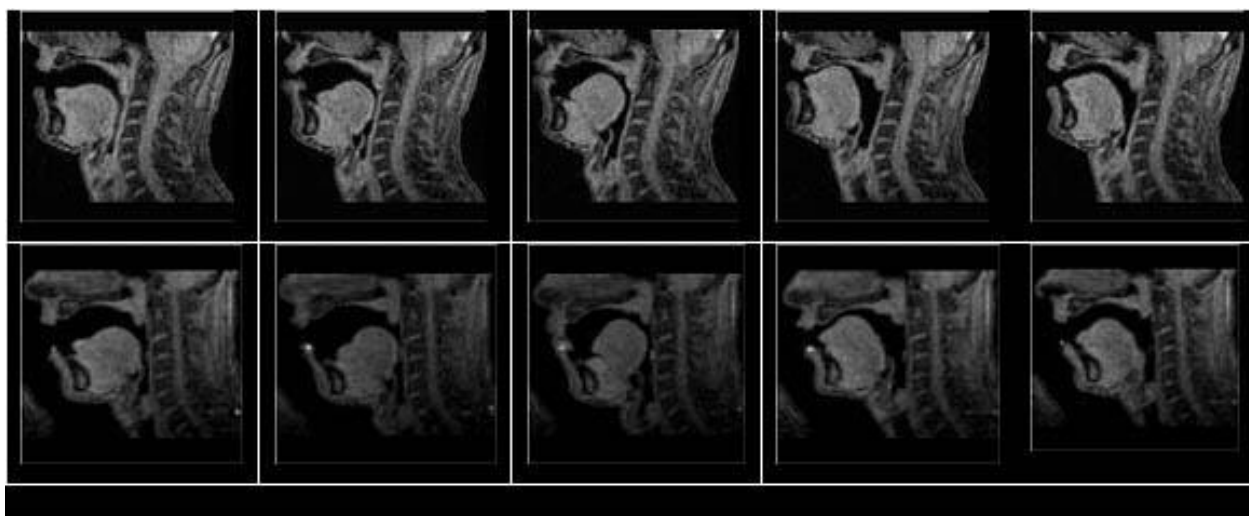


Рисунок 3. МРТ-изображения артикуляторных поз для разных реализаций одного и того же гласного звука (квазистационарных сегментов его звучания) в произнесении диктора-мужчины, полученные в ходе разных экспериментальных сессий. Верхний и нижний ряды отображают результаты МРТ-экспериментов, проведенных с промежутком в 1 год. Изображения в ряду соответствуют звукам [а:], [о:], [у:], [и:], [е:] (слева направо).

³ Как пишет Б. Серебренников, ссылаясь на Н.А. Бернштейна, «в этом сказывается ... “метрический” характер человеческой деятельности, позволяющий использовать при фиксированной цели различные способы ее достижения и по ходу выполнения намеченного плана изменять эти способы соответственно изменившейся обстановке». (Цит. по: [Серебренников 1970]).

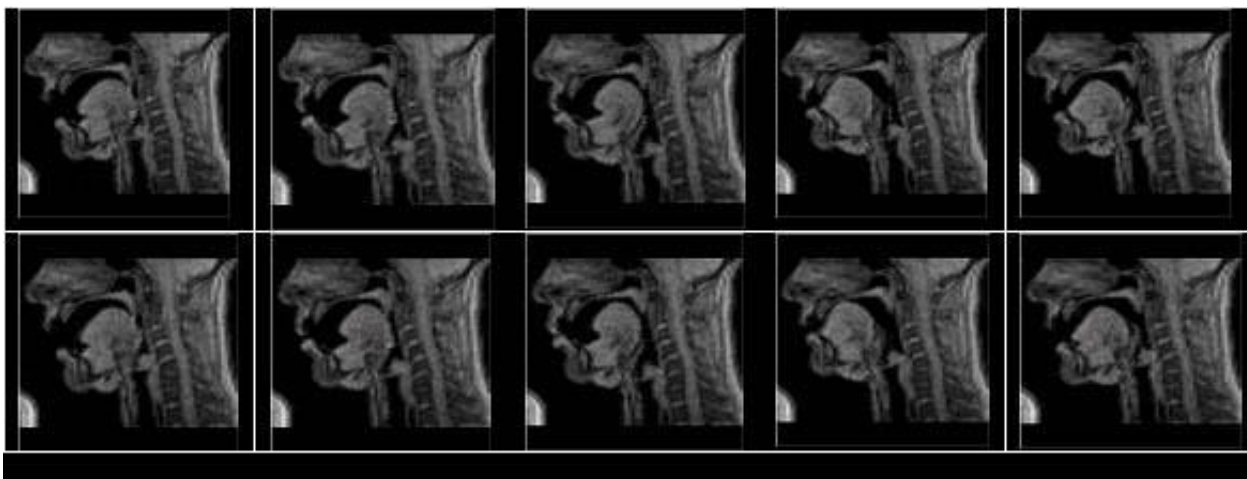


Рисунок 4. МРТ-изображения артикуляторных поз для разных реализаций одного и того же гласного звука (квазистационарных сегментов его звучания) в произнесении диктора-женщины, полученные в ходе разных экспериментальных сессий. Верхний и нижний ряды отображают результаты экспериментов, проведенных с промежутком в 1 год. Изображения в строке соответствуют звукам [a:], [o:], [y:], [и:], [е:] (слева направо).

Таким образом, полученный нами в результате проведенного МРТ-исследования экспериментальный материал, который в первую очередь отличается значительной внутридикторской стабильностью речевых артикуляций, как нам кажется, полностью подтверждает гипотезу, выдвинутую еще И.А. Бодуэном де Куртенэ и В.А. Богородицким, о том, что звук речи (Богородицкий предлагает называть его «фонемой») представляет собой сложный психомоторный (произносительно-слуховой) комплекс, формирующийся на основании ассоциации смежности у каждого носителя языка в раннем детском возрасте через упражнения. В.А. Богородицкий описывает этот комплекс как состоящий из «волевых импульсов, идущих из определенных центров головного мозга к соответствующим мышцам говорильного аппарата... [эти] импульсы вызывают сокращение этих мышц, вызывающее в свою очередь в сенсорных центрах головного мозга мышечное ощущение. Артикуляционные работы частей речевого аппарата вызывают во внешней воздушной среде определенное колебательное движение, передающееся барабанной перепонке уха, которое возбуждает в мозгу слуховое ощущение соответствующего звука, причем у говорящего оно является в сочетании с звукопроизводством, а у слушающего – только оно одно» [Богородицкий 2004].

Нам представляется, что полученные нами МРТ-изображения артикуляций русских гласных звуков с высокой степенью достоверности могут быть признаны кандидатами на отображение таких стабильных моторно-сенсорных комплексов, которые характеризуют базовую артикуляцию гласных звуков русского языка, как минимум, у данных дикторов. И здесь понятие базовых артикуляций, или артикуляторной базы языка представляет для

нас особый интерес. Более того, можно утверждать, что именно понятие артикуляторной базы языка является тем ключевым лингвистическим понятием, которое и обеспечивает единство такого многоаспектного явления, как речевая артикуляция, в том числе и в ее индивидуальных проявлениях. Ведь, как отмечает Л.Р. Зиндер, артикуляционная база, с одной стороны, «отражает механизм артикуляции звуков данного языка его носителями», т.е. она самым тесным образом «связана с биологическим аспектом языка»; но с другой стороны, артикуляционная база «вырабатывается постепенно в результате усвоения ребенком фонетической системы того языка, как котором говорят окружающие его люди», и может быть поэтому определена как «совокупность артикуляционных и фонотактических привычек», характеризующих то или иное языковое сообщество [Зиндер 1984]. Сопоставление базовых артикуляций по всем звукам (с учетом особенностей их фонотактики) у разных носителей языка, как нам представляется, позволит сформировать адекватные и объективные представления артикуляторной базы языка в целом, о фундаментальном значении такого рода представлений пишут практически все лингвисты-фонетисты – как теоретики, так и прикладники-практики.

Литература

- Богородицкий 2004 – Богородицкий В.А. Очерки по языковедению и русскому языку. М., Едиториал УРСС, 2004.
- Болла 1981 – Болла К. Атлас звуков русской речи. Будапешт, 1981.
- Зиндер 1984 – Зиндер Л.Р. К вопросу об артикуляторной базе // Экспериментально-фонетический анализ речи. Вып. 1. Изд-во Ленинградского университета. 1984.
- Кедрова, Захаров, Пирогов, Анисимов 2003 – Кедрова Г.Е., Захаров Л.М., Пирогов Ю.А., Анисимов Н.В. Исследование артикуляторной базы русского языка методами магнитно-резонансной томографии // XIII сессия Российского акустического общества (25-29 августа, 2003). Сб. трудов. Т. 3. Акустика речи. Медицинская и биологическая акустика. М., ГУОС, 2003.
- Кодзасов, Красовицкий, Щигель 2002 – Кодзасов С.В., Красовицкий А.М., Щигель Е.В. Проблемы описания спектров русских гласных // Проблемы фонетики. Вып. IV. Сб. статей / Отв. ред. Р.Ф. Касаткина. М.: Наука, 2002.
- Потапова, Потапов 2006 – Потапова Р.К., Потапов В.В. Язык, речь, личность. М.: «Языки славянской культуры», 2006.
- Серебренников 1970 – Серебренников Б. Общее языкознание. М., 1970.
- Щерба 1983 – Щерба, Л.В. Русские гласные в качественном и количественном отношении. Л.: Наука, 1983.