

## ИССЛЕДОВАНИЕ АРТИКУЛЯТОРНОЙ БАЗЫ РУССКОГО ЯЗЫКА МЕТОДАМИ МАГНИТНО-РЕЗОНАНСНОЙ ТОМОГРАФИИ

МГУ им. М. В. Ломоносова

Россия, 119992 Москва, ГСП-2, Ленинские горы, МГУ, корп. 1, филологический ф-т, комн. 983  
Тел.: (095) 939-1478; Факс: (095) 939-5596

Эл. Почта: [kedr@philol.msu.ru](mailto:kedr@philol.msu.ru)

МГУ им. М. В. Ломоносова

Россия, 119992 Москва, ГСП-2, Ленинские горы, МГУ, ул. Лебедева, 5, ЦМТС МГУ  
Тел.: (095) 939 1669;

Эл. Почта: [nisimovnv@mail.ru](mailto:nisimovnv@mail.ru)

*В докладе представлены промежуточные результаты работ, выполняемых в МГУ им. М. В. Ломоносова в рамках инициативного междисциплинарного проекта, основной целью которого является разработка методик использования новейших технических средств (магнитно-резонансной томографии) в исследовании механизмов речепорождения. На основе разработанных методик в дальнейшем планируется сформировать компьютерную базу данных, в которой будут представлены графические образы для всех типов артикуляторных комплексов русского языка, включая переходные (внутрислововые и межслововые) участки. На данном этапе исследования решалась задача выявления с помощью ФМРТ-визуализации особенностей работы артикуляторных органов при произнесении гласных звуков (как в изолированной позиции, так и в составе слогов и слов). При этом особое внимание было уделено визуализации движений губ, языка, конфигурации ротовой полости, нёбной занавески, гортани в процессе порождения звука.*

Наше исследование посвящено изучению возможностей использования функциональной магнитно-резонансной томографии для получения верифицированных экспериментальных данных, наиболее подробно и полно отражающих артикуляторные процессы порождения речи. В докладе будут изложены промежуточные результаты работ, выполняемых в МГУ им. М. В. Ломоносова в рамках междисциплинарного проекта, основной целью которого является разработка методик использования новейших технических средств (магнитно-резонансной томографии) в исследовании механизмов речепорождения и формирование на основе разработанных методик электронных баз данных, в которых будут представлены графические образы для всех типов артикуляторных комплексов русского языка. На первом этапе исследования (конец 2002 – начало 2003 г.) решалась задача выявления с помощью МРТ-визуализации особенностей работы артикуляторных органов при произнесении гласных звуков (как в изолированной позиции, так и в составе слогов и слов). При этом особое внимание было уделено визуализации движений губ, языка, конфигурации ротовой полости, нёбной занавески, гортани в процессе порождения звука.

Актуальность проекта обусловлена отсутствием достоверных знаний о динамических механизмах речепорождения в разных языках, и в частности, применительно к русскому языку, поскольку до настоящего времени единственным средством экспериментального исследования речевой артикуляции здесь была кинофоторентгенография. До сегодняшнего дня практически единственным источником сведений об артикуляторной базе русского языка (в части стационарных участков звучания) является справочное издание “Атлас звуков русской речи”, подготовленное по кинофоторентгенограммам в Венгрии [1]. Основные недостатки этого метода (снятие, расшифровка и прорисовка кинофоторентгенограмм) связаны с очень большими техническими сложностями в его практическом применении (собственно технологические и санитарно-гигиенические ограничения), которые обусловили принципиальную невозможность изучения речи в процессе ее развертывания (в реальном времени). Разрабатываемый нами метод (ФМРТ-визуализация) позволяет решить эту задачу, что будет иметь самые существенные

последствия для успешного решения как ряда фундаментальных задач (прежде всего совершенствования теории речепорождения), так и для множества прикладных областей речеведения (лингводидактики, автоматического распознавания и синтеза речи, компьютерных речевых баз данных и т.д.).

Основной метод нашего исследования – функциональная магнитно-резонансная томография процесса речепроизводства. Задача ФМРТ – получать МР-изображения в режиме реального времени. МР-изображения используются для визуализации органов в движении, анализа физиологических процессов, исследования реакции структур мозга на внешний раздражитель. Простейшим вариантом ФМРТ является последовательная съемка одной и той же зоны сканирования через определенные интервалы времени для анализа эволюции этой зоны.

ФМРТ имеет свои технологические особенности. Для наших задач наиболее существенными являются ограничения на скорость считывания информации при сканировании. Разрешающая способность современных томографов в большинстве случаев превышает 0.5 сек., что затрудняет съемку быстротекущих процессов. Исследователь должен найти компромисс между высокой скоростью получения изображений и их информативностью. Для повышения скорости считывания МР-сигнала: а) уменьшают количество считываемых точек, что приводит к снижению разрешающей способности; б) используют импульсные последовательности с короткими задержками, из-за чего падает взвешенность изображения; в) минимизируют количество накоплений сигнала, что снижает отношение сигнала к шуму; г) отказываются от вспомогательных задержек и импульсов, используемых для компенсации аппаратных несовершенств.

Задача данного исследования – с помощью МР-визуализации выявить особенности работы органов, ответственных за голосовое воспроизведение речи. Особое внимание уделяется визуализации губ, языка, ротовой полости, глотки, гортани при произнесении гласных звуков. Время сканирования одного среза не должно превышать времени произнесения гласного звука – примерно 0.5 - 1 с. (последнее – для тянутого произнесения).

Кроме того, представляет интерес МР отображение процесса произнесения последовательности звуков – слов или фраз. Для визуализации этого процесса необходимо иметь кадры, снятые с интервалом не более 0.1с (10 кадров/сек). Наша МРТ-аппаратура позволяет сканировать срез со скоростью более 0.5 сек. Поэтому используется стробоскопический способ сбора данных: быстрое МР сканирование ведется при многократном повторении одной и той же фразы. Кадры для анимационного представления создаются из МР изображений, которые расставляются с учетом временной привязки каждого кадра относительно начала звуковой осциллограммы.

Материалом эксперимента послужили русские гласные звуки, произнесенные в изолированной позиции и в составе фразы (некоторые гласные). В качестве дикторов выступали 2 человека (мужской и женский голоса). Томографическая съемка проводилась в несколько сеансов, для диктора-мужчины в целях оценки устойчивости моделей произнесения были осуществлены две серии записей. Параллельно в ходе эксперимента велась контрольная аудиозапись (на два канала с установление сигнальных меток в момент начала съемки), а в ходе имитации эксперимента вне магнита – видеосъемка. Полученный массив экспериментальных данных был вручную обработан: осуществлена расшифровка аудиотреков, их разметка, проведена атрибуция изображений, для оценки степени информативности полученных изображений был также проведен специальный перцептивный эксперимент.

Одна из целей перцептивного эксперимента заключалась в исследовании степени информативности полученных Т-изображений и возможностей их кластеризации на основе воспринимаемых глазом различий в особенностях конфигураций артикуляторных органов, а также степень совпадения полученных от разных испытуемых классификаций изображений. В эксперименте приняли участие 10 человек, время работы с экспериментальным материалом было неограничено, размер кластеров и их число также никак не ограничивались.

В результате нам удалось показать высокую информативность полученных изображений, с высокой степенью достоверности представить все основные особенности артикуляторных движений при произнесении русских гласных звуков. Было проведено сравнение полученных Т-изображений с материалами К. Боллы [1]. При значительном совпадении основных параметров обоих типов изображений отмечена большая детализация действий органов артикуляции в случае томографической

съемки, особенно в отношении мягких тканей, которые могут проявлять себя как наиболее подвижные (напр., кончик языка).

В ходе доклада будут подробно представлены полученные результаты – результаты перцептивного эксперимента, изображения основных конфигураций артикуляторных органов при производстве русских гласных звуков, полученные в режиме реального времени (по данным томографического исследования) для изолированного произнесения звуков и для звуков в составе фразы.

Рис. 1. Структура эксперимента и типы полученных данных.

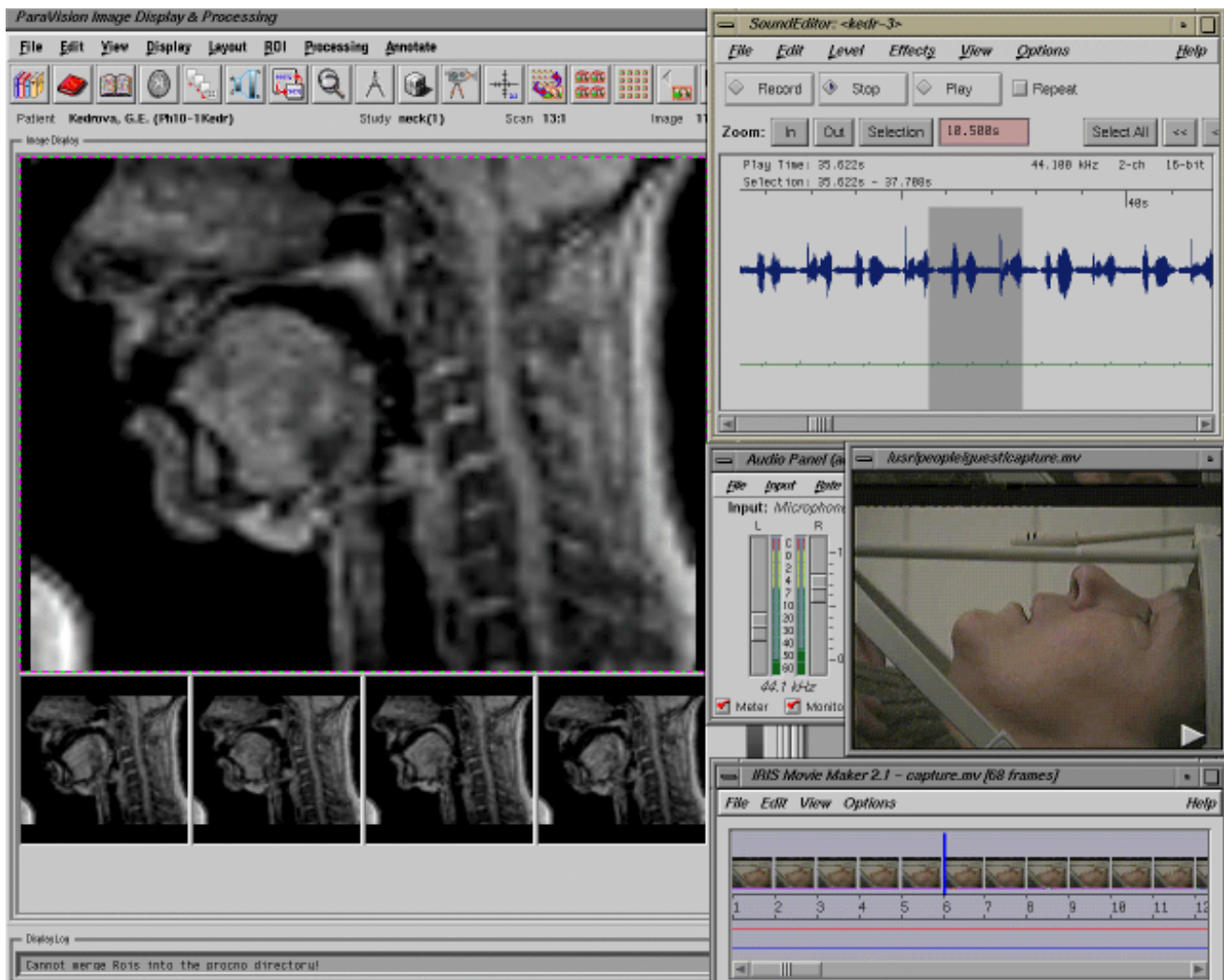
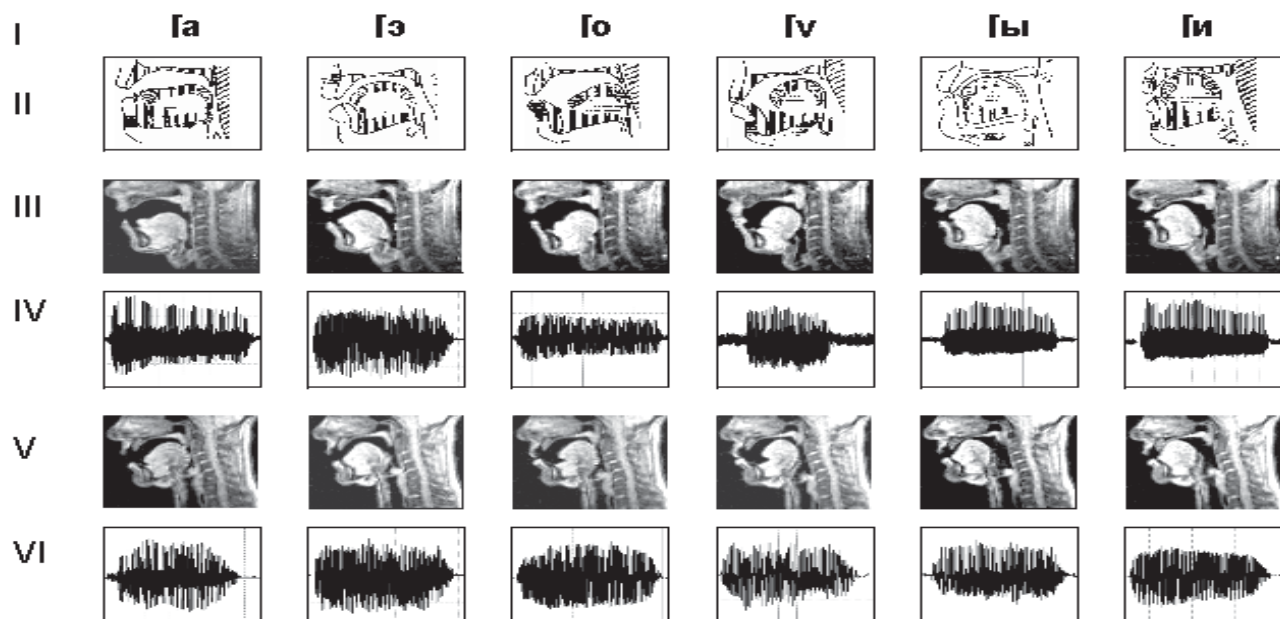


Рис. 2. МР визуализация процесса произнесения гласных звуков русского языка в сопоставлении с рентгеновским изображением и звуковой осциллограммой.



*Примечание* к рис. 2. **I** – основные гласные звуки русского языка, **II** – изображения ротовой полости при произнесении соответствующих звуков. Изображения получены с помощью рентгеновской съемки с последующим ручным ретушированием [1], **III**- МР изображения ротовой полости и шейного отдела испытуемого Л. (муж. 48 лет), при произнесении тех же звуков. **IV** – осциллограммы тех же звуков. Длительность осциллограммы ~2 с. **V** и **VI** – аналоги III и IV, для испытуемой Г. (жен. 45 лет). МР изображение с разрешением 0.35x0.35 мм получается за 0.55 с.

## ЛИТЕРАТУРА

1. К. Болла “Атлас звуков русской речи”, Будапешт, 1981 г.