

УДК 612.821.6

## фМРТ-КОРРЕЛЯТЫ КАТЕГОРИЗАЦИИ ЗВУКОВ РЕЧЕВОЙ И НЕРЕЧЕВОЙ ПРИРОДЫ В СЛУХОВОЙ КОРЕ

© 2019 г. В. М. Шкловский<sup>1</sup>, С. А. Варламов<sup>2, 3, \*</sup>,  
А. Г. Петрушевский<sup>1</sup>, Л. А. Майорова<sup>1, 3, \*\*</sup>

<sup>1</sup>ГБУЗ Центр патологии речи и нейрореабилитации Департамента здравоохранения Москвы, клиническая база ФГБУ Федеральный медицинский исследовательский центр психиатрии и наркологии им. В.П. Сербского Минздрава РФ, Москва, Россия

<sup>2</sup>Московский государственный университет им. М.В. Ломоносова, Москва, Россия

<sup>3</sup>ФГБУН Институт высшей нервной деятельности и нейрофизиологии РАН, Москва, Россия

\*E-mail: sv-kanino@mail.ru

\*\*E-mail: larimayor@gmail.com

Поступила в редакцию 04.03.2019 г.

После доработки 28.03.2019 г.

Принята к публикации 02.06.2019 г.

Статья посвящена изучению функциональной структуры слуховой коры путем выявления и сопоставления пространственной локализации зон активации (по данным фМРТ) в ответ на речевые и неречевые стимулы. Помимо этого, было проведено аналогичное сопоставление зон активации в ответ на мужской и женский голос. Было обнаружено, что имеются специфические речевая, неречевая и перекрывающаяся зоны, причем речевая имеет значительные размеры относительно других. Перекрывание зон активации при прослушивании мужского и женского голоса имеет место, но не значительно, при этом влияние женского голоса оказалось значительно сильнее. Данные результаты свидетельствуют о наличии особых зон слуховой коры, предназначенных для специализированной обработки звукового сигнала.

**Ключевые слова:** слухо-речевой гнозис, верхневисочная кора, фМРТ.

**DOI:** 10.1134/S0131164619060122

В настоящее время в мире ведется множество исследований устройства работы слуховой системы мозга. Слуховая кора, будучи высшим звеном анализа звука, является интегративным центром, отвечает за обработку сложных звуковых характеристик, в том числе и речи. На сегодняшний день до конца не ясно, по какому принципу устроена организация этой обработки. Существуют три возможных варианта: или все сигналы обрабатываются одними и теми же нейрональными группами, или существует четкая специализация этих групп, или имеются общие и специальные группы. Данное исследование дает новые сведения по этому вопросу.

Слуховой анализатор в корковом отделе расположен в височной области коры больших полушарий, в районе извилин Гешля и верхней височной, главным образом представлен полем 41 (по Бродману) и соседними полями 42, 22, которые отвечают за речевую функцию [1].

В настоящее время еще окончательно не известно, в какой форме происходит обработка речи на физиологическом уровне: или для этого

имеют место специальные нервные механизмы, или же речь обрабатывается общими механизмами, как и любые сложные звуки. В соответствии с этой теорией восприятия речи можно разделить на две категории [2]. Левая верхняя височная кора, играя существенную роль в обработке речи, более чувствительна к речевым звукам. Многие исследования показали, что при прослушивании речи можно наблюдать повышение активности нейронов в задних частях левых верхних височных извилин и борозды (*STS*, *STG*). Эти области используются для обработки быстроменяющихся характеристик звука, в частности, формантных переходов, в связи с этим существует мнение, что *STS* и *STG* скорее участвуют в обработке сложных звуков, а не речи как таковой [3, 4].

Верхняя височная извилина содержит несколько функционально неоднородных областей, которые реагируют на речь, играя важную роль в ее обработке, будучи чувствительной ко многим характеристикам речевого сигнала, например, к формантным частотам гласных, спектрально-временному составу согласных. При экспери-

ментальном исследовании в нескольких верхневисочных областях наблюдалась повышенная активность во время прослушивания речевых последовательностей. Так, оказалось, что передние верхневисочные области участвуют в декодировании речи независимо от крупномасштабной статистической структуры звуковых последовательностей, состоящих из речевых и неречевых фрагментов с различной степенью их стохастичности. В отличие от передних, несколько задних верхневисочных областей проявляли чувствительность к статистической структуре. Таким образом, можно говорить о наличии функционального разделения верхней височной плоскости на области, которые не могут быть активированы одновременно одним и тем же раздражителем [5].

Одной из таких специализированных областей является *planum temporale*, представляющая собой часть верхней височной плоскости непосредственно позади поперечной височной борозды, ограниченная медиально Сильвиевой бороздой, а сзади — надкраевой извилиной. *Planum temporale* представляет собой треугольную область, которая расположена каудально относительно извилины Гешля в верхней височной плоскости. Первоначальные исследования этой области показали, что она является одной из наиболее асимметричных областей в мозге, будучи существенно больше в левом полушарии головного мозга, чем в правом, причем данная асимметрия наблюдается не только у человека, но и у обезьян [6]. *Planum temporale* имеет множество различных связей с другими областями мозга, она анатомически неоднородна, определить ее точные границы достаточно трудно. По сравнению с первичной слуховой корой, в *planum temporale* представлены этапы обработки звука более высокого порядка. Известно, что *planum temporale* участвует в опосредованных сенсомоторных процессах [7, 8]. Предполагается, что она является частью регуляторной системы, имеющей важное значение для беглости речи. Как было выяснено, в *planum temporale* имеются отдельные зоны, которые преимущественно реагируют или на восприятие речи, или на ее выработку, что свидетельствует о наличии внутреннего функционального распределения. Это может означать, что имеют место функционально отличные популяции нейронов, которые преимущественно реагируют на сенсорные или моторные стимулы [9, 10].

Существует гипотеза “вычислительного центра” [11], согласно которой левая *planum temporale* устраняет неоднозначность сложных звуков, сопоставляя их временные и спектральные характеристики с хранимыми шаблонами. Можно предположить, что левые передняя и средняя *planum temporale* участвуют в построении предсказания поступающей слуховой информации. Такие механизмы могут использоваться для выявления не-

соответствий в речевом потоке, особенно если последовательности хорошо предсказуемы, последующей корректировки внутренних представлений и сигнализации об ошибке [5]. Существует предположение, что тонкая структура акустического сигнала важна для сохранения каждого конкретного шаблона в памяти [12, 13].

Согласно экспериментальным данным, у испытуемых при прослушивании “синусоидальной” речи (*SWS* — *sine wave speech*) наблюдалась значительно более сильная активация после обучения в левой задней верхней височной борозде (*STSp*), однако это имело место только у тех испытуемых, кто научился выявлять речевой сигнал из *SWS*. Таким образом, активация нейронных речевых представлений в левой *STSp* может быть необходимым условием для восприятия звуков как речи. Это подтверждает предположение, что в основе восприятия речи лежит специализированная нейронная обработка в левой задней верхней височной коре (в зоне Вернике) [14].

Помимо областей височной коры, наблюдалась чувствительность к человеческому голосу в районе левой нижней префронтальной коры человека (*pars orbitalis*). Повышенная реакция при прослушивании голоса имела место как для речи, так и для неречевых звуков, и была сильнее при наличии эмоциональной окраски. Ранее было обнаружено существование слухового домена в вентrolатеральной префронтальной коре мозга макаки, вне гомологичной человеческой зоны Брока, нейроны которого устойчиво реагировали на сложные звуки, сильнее активируясь на видоспецифичные и человеческие вокализации [15–17]. Таким образом, *pars orbitalis* может являться аналогичным префронтальным слуховым доменом в мозге человека. В эксперименте при сравнении активирующего воздействия на мозг человека звуков речевой и неречевой природы при помощи фМРТ было показано, что только одна область левой нижней лобной извилины, *pars orbitalis* (47 зона Бродмана), проявила более высокую активность для звуков человеческого голоса, включая нелингвистические (смех, кашель), по сравнению со звуками животных. При этом гомологичная область в правом полушарии, реагировала только на нелингвистические вокализации, активируясь меньше. Можно предположить, что в левой префронтальной коре человека существует слуховой домен, отличный от области Брока [18].

Таким образом, можно отметить, что в верхней височной области существует сложная функциональная организация. Обработка речевой информации проводится путем спектрального анализа, построения предсказания, сравнения с шаблоном. Помимо этого существует слуховой домен в левой префронтальной коре, способный

специфично активироваться на звуки, производимые человеком.

Вопрос определения пола говорящего на физиологическом уровне является тоже весьма актуальным, но менее исследованным. В частности, было показано, что задача классификации ассоциации лица/голоса в зависимости от пола приводит, во-первых, к унимодальным активациям как зрительной, так и слуховой области (билатеральная верхняя височная извилина), во-вторых, к специфическим супрамодалным активациям левой верхней теменной и правой нижней лобной извилин. При этом оба унимодальных региона оказались связаны между собой и с префронтальной извилиной и скорлупой, а левая теменная извилина имела усиленную связь с теменно-премоторной сетью, которая участвует в перекрестном контроле внимания [19].

В данном исследовании в качестве стимульного материала использовались короткие фрагменты звуков с целью выявления специфической активации коры головного мозга в зависимости от их природы: речевой и неречевой, в том числе в зависимости от пола говорящего. Ключевым принципом для выбора основного речевого звука было наличие у него неречевых аналогов. Таким звуком стал [р] (записанный как в мужском, так и в женском исполнении), к которому были подобраны подобные ему рык животных и шум моторов. Во время прослушивания этих звуков проводилось фМРТ-сканирование с дальнейшей обработкой полученных данных и построением карт активации мозга.

## МЕТОДИКА

В исследовании принимали участие 20 чел., 3 из них были исключены из-за неоднозначности полученных активационных карт. Средний возраст оставшихся 17 чел. (8 женщин, 9 мужчин) составил  $45.5 \pm 8.6$  лет. У всех испытуемых отсутствовал неврологический и психиатрический дефицит в анамнезе.

У каждого испытуемого перед сканированием оценивали профиль функциональной сенсомоторной асимметрии (использовали опросник Аннет). Индекс латерализации по группе в среднем составил  $0.65 \pm 0.21$  (преимущественно правши).

Перед функциональным сканированием давали инструкцию лежать спокойно, расслабленно, не фокусировать внимание на слуховых стимулах.

**Стимульный материал.** В качестве слуховой стимуляции в наушники 1 раз в 15 с подавали звуки речевой природы ([р], произнесенные 6 дикторами мужского и женского пола, а также искусственно синтезированные) и [р]-подобные звуки неречевой природы (шум моторов, рык животных). Все звуки были выровнены по длитель-

ности (380 мс) основной форманте и громкости. Всего каждому испытуемому предъявляли 126 звуков (63 речевой и 63 неречевой природы). Слуховую стимуляцию производили бинаурально через наушники, поставляемые в комплектации с презентационным оборудованием (*In vivo*, США).

**Сканирование и обработка данных.** МРТ проводили на томографе *MAGNETOM AVANTO 1.5 T* (*Siemens*, Германия). Для получения анатомического изображения в сагиттальной плоскости использовали последовательность *T1 MPRAGE* ( $TR - 1900$  мс,  $TE - 3.4$  мс, 174 среза, толщина среза – 1 мм,  $FoV - 250$  мм, матрица реконструкции –  $256 \times 256$ , размер вокселя –  $1 \times 1 \times 1$  мм).

Для сбора функциональных данных использовали последовательность со следующими характеристиками:  $TR - 3000$  мс, задержка – 0 мс,  $TE - 50$  мс, 32 среза, толщина среза – 3 мм,  $FoV - 192$  мм, матрица –  $64 \times 64$ , размер вокселя –  $3 \times 3 \times 3$  мм. Исследование включало сбор 630 измерений (объемов) для каждого участника.

Анализ фМРТ-данных проводили с помощью пакета статистической обработки *SPM12* [18] на платформе *MATLAB R2012b* (*MathWorks, Natick, MA, USA*) на индивидуальном и групповом уровнях. На индивидуальном уровне два первых функциональных объема исключали из анализа, оставшиеся изображения выравнивали по времени (процедура *slice timing*), а также относительно первого объема для коррекции движений. Далее проводили корегистрацию среднего функционального изображения со структурным. Процедуру пространственной нормализации в обеих группах проводили с помощью инструмента *DARTEL* (*Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra*) и *New Segment*. Сглаживание изображения с помощью фильтра Гаусса проводили с ядром фильтра размером  $6 \times 6 \times 6$  мм.

Статистические параметрические карты строили с помощью общей линейной модели (*GLM*) [21]. На групповом уровне результаты *t*-статистики (*SPM{t}*) (применяли одновыборочный *t-test*, а также попарный *t-test* для сравнения активации на различные виды стимулов) трансформировали в нормальное стандартное распределение (*SPM{z}*). Воксельный порог активации соответствовал  $p_{FWECorr} < 0.05$  с поправкой на множественные сравнения.

Для построения индивидуальных карт активации использовали контраст, направленный на выявление активации коры полушарий головного мозга на звуки речевой и неречевой природы, также отдельно выделялся контраст на диктора, произносившего звук [р], мужского и женского пола.

Степень латерализации активаций рассчитывали на групповом и индивидуальном уровне по

формуле  $LI = (Left - Right) / (Left + Right)$  с применением приложения  $LI$  для  $SPM$  [22, 23].

## РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

*Активация слуховой коры на звуки речевой и неречевой природы.* В ходе проведения эксперимента и дальнейшей обработки были получены усредненные по группе карты активации мозга. В первом случае (рис. 1) рассматривается влияние природы [р]-звуков: речевой (рис. 1, А) и неречевой (рис. 1, Б). Из рис. 1 видно, что в целом активизируются схожие области, однако точная локализация и интенсивность активации оказались различными. Речевой сигнал активизирует специфические зоны мозга, расположенные на периферии относительно “неречевой” зоны активации. У испытуемых в ответ на предъявление звуков речевой природы активация наблюдалась в слуховой коре с обеих сторон (69 и 98 активированных вокселей слева и справа соответственно), коэффициент латерализации составил  $-0.17$ . Активация наблюдается в 13, 22, 38, 40, 41 полях Бродмана (рис. 1, А, табл. 1). В случае неречевых звуков активация была меньше, но тоже с обеих сторон (17 и 15 активированных вокселей слева и справа соответственно) при коэффициенте латерализации  $0.06$ . В данном случае активация имела место в следующих полях Бродмана: 22, 40, 41, а также в 1-м (рис. 1, Б, табл. 1). При попарном сравнении активационных карт фМРТ (контраст на групповом уровне “звуки речевой природы > звуки неречевой природы”) у испытуемых при прослушивании звуков разной природы получены статистически значимые присты активации в левых 41 и 22 зоне Бродмана, суммарно 17 вокселей, а также в правых 41 и 22 зоне Бродмана, суммарно 5 вокселей; коэффициент латерализации составил  $0.55$  (рис. 1, В, табл. 1).

Особый интерес вызывает пространственное сопоставление зон активации мозга различной природы, показанное на рис. 2. Из рис. 2 видно, что общая зона активации существенно меньше специфической речевой зоны, в то время как специфическая неречевая зона меньше общей зоны. Видно, что в правом полушарии расположено более половины чисто речевой зоны, включающей 38, 40, а также частично 22 и 41 зоны Бродмана, в то время как чисто неречевая в нем представлена 1 и частично 41 зонами Бродмана, а общая — частично 22 и 41 зонами Бродмана. В левом полушарии чисто речевая зона представлена 13, 22; чисто неречевая — частично 40; общая — частично 40 и 41 зонами Бродмана соответственно (табл. 1).

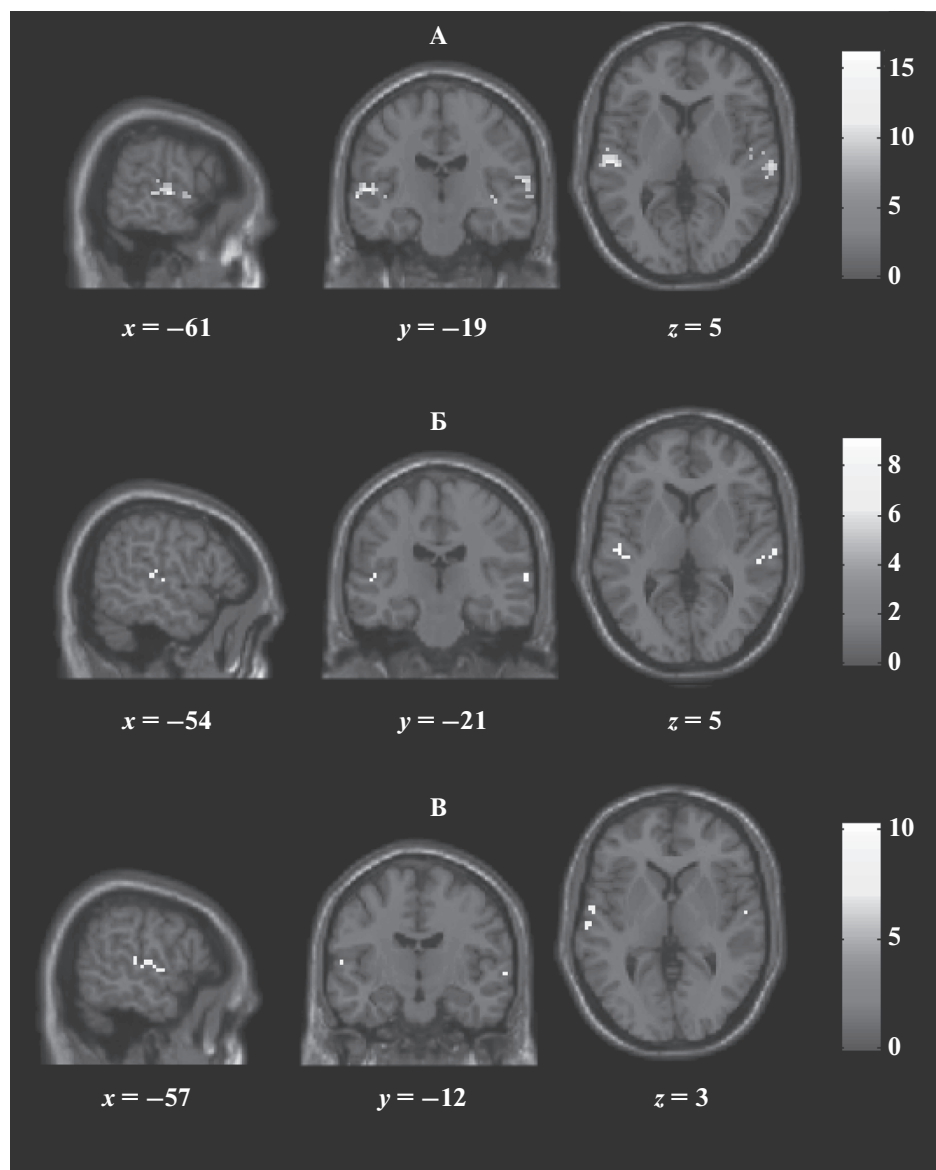
*Влияние пола диктора на активацию зон мозга при прослушивании звуков речевой природы.* Во втором случае (рис. 3) изучалось влияние пола диктора на активацию зон мозга при прослушивании. Оказалось, что в случае женского голоса

объем активированных областей несколько больше, чем при прослушивании мужского. У испытуемых в ответ на предъявление звуков женского голоса активация наблюдалась в слуховой коре с обеих сторон (38 и 41 активированных вокселей слева и справа соответственно), коэффициент латерализации составил  $-0.04$ . Активация наблюдается в 22, 40, 41, а также в 1-м полях Бродмана (рис. 3, А, табл. 2). В случае звуков мужского голоса активация была меньше, но тоже с обеих сторон (20 и 16 активированных вокселей слева и справа соответственно) при коэффициенте латерализации  $0.1$ . В данном случае активация имела место в следующих полях Бродмана: 13, 22, 41 (рис. 3, Б, табл. 2).

При попарном сравнении активаций на [р]-подобные звуки речевой природы, произносимые дикторами мужского и женского пола, значимых отличий на уровне  $pFWE < 0.05$  выявлено не было. Однако при пространственном сопоставлении активационных карт (рис. 4) можно видеть, что были активированы различные области, т.е. перекрывание, и как следствие, функциональная общность зон имеет место, но незначительно.

## ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

*Активация слуховой коры на звуки речевой и неречевой природы.* Таким образом, полученные результаты свидетельствуют о принципиальном существовании специфических зон активации слуховой коры в ответ на неречевые звуки, схожие с речевыми. Данный факт может говорить о наличии физиологического механизма первичной обработки и разделения звуков по видам, распределяющего дальнейшую их обработку в специализированные зоны мозга. При сопоставлении пространственных карт активации коры полушарий головного мозга было выявлено относительное увеличение объема активации на речевые стимулы, что свидетельствует о ведущей функции обработки речевого сигнала. Концептуально, это согласуется с результатами [5], где тоже наблюдалось наличие как общей, так и специализированных зон при прослушивании речевых и неречевых звуков, а также латерализация (большая речевая зона в левом полушарии). Наблюдалась, в частности, речевая специфическая активация в районе левой верхней височной борозды, что согласуется с результатами [14]. Однако существенной активации левой нижней лобной извилины как маркера речевых звуков, в отличие от [18] замечено не было, что, возможно, связано с различием в методиках экспериментов. Специфическая активация речевыми стимулами участков в левой передней верхней височной извилине, которую обнаружили [24], подтверждает наши наблюдения. Показанное в эксперименте [25] отсутствие существенного различия в ответе слухо-



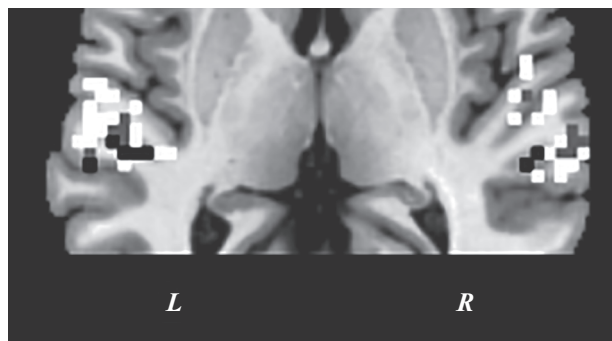
**Рис. 1.** Усредненные по группе карты активации мозга при предъявлении [r]-подобных звуков речевой природы (А) *one sample t-test* ( $T = 7.2, p_{FWE} < 0.05$ ), неречевой природы естественного (рев животных) и неестественного происхождения (рев мотора) (Б) *one sample t-test* ( $T = 7.2, p_{FWE} < 0.05$ ). (В) – сравнительный анализ активаций (А) и (Б), *paired t-test* ( $T = 7.2, p_{FWE} < 0.05$ ).

вой коры на звук валторны при наличии такового на звуки речи у носителей немецкого и китайского языков также подкрепляет идею специализированной обработки речевой и неречевой информации. Полученная при сравнении контрастов левосторонняя латерализация активации на речевые стимулы относительно неречевых, учитывая специфику стимулов, еще раз подтверждает предположение С. *Wernicke* и А.Р. *Luria* [26, 27] о том, что дефицит восприятия речи при одностороннем поражении левого полушария, вызван ухудшением способности воспринимать простые речевые звуки.

*Влияние пола диктора на активацию зон мозга при прослушивании звуков речевой природы.* Небольшое увеличение объема активации на предъявление женского голоса относительно мужского, полученное при сопоставлении пространственных карт активации коры полушарий головного мозга, должно способствовать лучшему восприятию женского голоса. Значительное пространственное разделение обработки голоса по половому признаку проявляет глубину и фундаментальность данной характеристики для человека. Увеличение активации на предъявление [r]-подобных звуков речевой природы (а также на жен-

**Таблица 1.** Области активации коры больших полушарий головного мозга в ответ на слуховую стимуляцию

Полушарие	Область	Поле по Бродману	<i>x</i>	<i>y</i>	<i>z</i>	<i>T</i>	<i>p</i> ( <i>FWE-corr</i> )	Размер кластера
Звуки речевой природы								
<i>R</i>	Височная доля	41	66	-22	11	14.85	0.00	84.00
		22	69	-31	2	11.34	0.00	
		22	66	-13	-1	10.85	0.00	
<i>L</i>	Височная доля	40	-39	-34	14	12.05	0.00	67.00
		41	-54	-19	5	11.94	0.00	
		22	-63	-19	-1	10.52	0.00	
<i>R</i>	Височная доля	40	54	-25	17	8.53	0.01	4.00
<i>R</i>	Височная доля	22	42	-19	-4	8.11	0.01	4.00
<i>R</i>	Височная доля	38	60	5	-10	7.97	0.01	3.00
<i>R</i>	Височная доля	41	42	-31	11	7.91	0.02	1.00
<i>L</i>	Височная доля	22	-42	-19	-1	7.58	0.03	1.00
<i>L</i>	Височная доля	13	-39	-13	-1	7.45	0.03	1.00
<i>R</i>	Височная доля	41	51	-16	5	7.25	0.05	1.00
<i>R</i>	Височная доля	22	60	-40	14	7.21	0.05	1.00
Звуки неречевой природы								
<i>L</i>	Височная доля	41	-51	-19	5	8.97	0.00	15.00
<i>L</i>	Височная доля	40	-54	-25	11	8.12	0.01	
<i>R</i>	Височная доля	41	57	-25	5	8.75	0.00	2.00
<i>R</i>	Височная доля	41	66	-22	8	8.71	0.00	9.00
<i>R</i>	Височная доля	41	57	-4	-1	8.67	0.00	3.00
<i>R</i>	Височная доля	1	66	-13	14	8.05	0.01	1.00
<i>L</i>	Височная доля	40	-39	-31	14	7.68	0.02	2.00
<i>R</i>	Височная доля	22	60	-37	11	7.37	0.04	1.00
<i>R</i>	Височная доля	1	60	-10	11	7.31	0.04	1.00
Контраст “звуки речевой природы > звуки неречевой природы”								
<i>L</i>	Височная доля	41	-60	-25	11	10.26	0.00	4.00
<i>R</i>	Височная доля	22	66	-10	-1	9.80	0.00	2.00
<i>L</i>	Височная доля	н/д (в просв. бор.)	-54	-4	-1	8.62	0.00	9.00
<i>L</i>	Височная доля	41	-57	-13	8	7.83	0.02	
<i>R</i>	Височная доля	22	54	-1	-1	8.44	0.01	1.00
<i>R</i>	Височная доля	41	57	-7	2	7.98	0.01	1.00
<i>L</i>	Височная доля	22	-60	-16	2	7.95	0.01	3.00
<i>R</i>	Височная доля	22	60	5	-7	7.30	0.04	1.00
<i>L</i>	Височная доля	22	-45	-19	-4	7.21	0.05	1.00

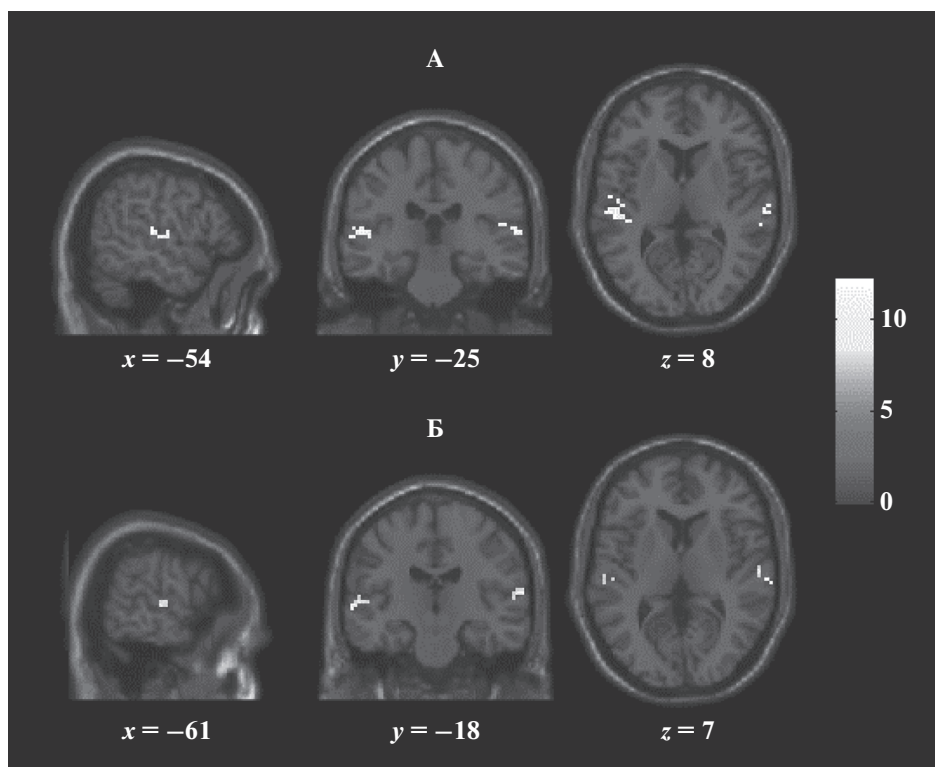


**Рис. 2.** Сопоставление пространственной локализации активности слуховой коры в ответ на два типа стимулов (речевой/неречевой природы). Слуховая кора (аксиальный срез на уровне  $z = 5$ ). Белым представлена активация на речевые стимулы, черным – на неречевые, серым – общие области активации.

ский голос) возможно объясняется эволюционной значимостью стимулов такого рода.

Согласно другим исследованиям, механизм реакции на голос диктора основан на пространственном функциональном разделении обработки сигнала. В частности, было показано [28], что нормализация гласных звуков (для их унификации) начинается на раннем этапе обработки речи,

при этом основной тон обрабатывается в извилине Гешля, а формантно-частотная и типологическая обработка гласных локализуется сразу за извилиной Гешля, в *planum temporale*. В другом исследовании [29] было продемонстрировано, что во время прослушивания различных дикторов (мальчик, девочка, мужчина) при выполнении задачи категоризации диктора имеет место увеличение активации в правых верхних височных извилинах и борозде, в то время как категоризация гласного осуществляется в правой задней височной коре. При этом данные области хотя и являются различными, имеют зону перекрытия. Также известно [30], что гласные звуки кодируются первыми двумя формантами, а вид диктора – частотой основного тона, являясь промежуточными вычислительными объектами, которые формируются не только в специализированных, но и в ранних слуховых областях мозга, причем за классификацию диктора отвечают зоны в борозде Гешля, рядом с зоной, отвечающей за классификацию гласных, а также три зоны в правой верхней височной борозде с вкраплениями областей обработки гласных. Помимо этого, как стало известно из исследования [31], посвященного восприятию голоса матери детьми, имеется усиление активации в задней слуховой коре, верхней височной и нижней лобной извилинах при прослу-



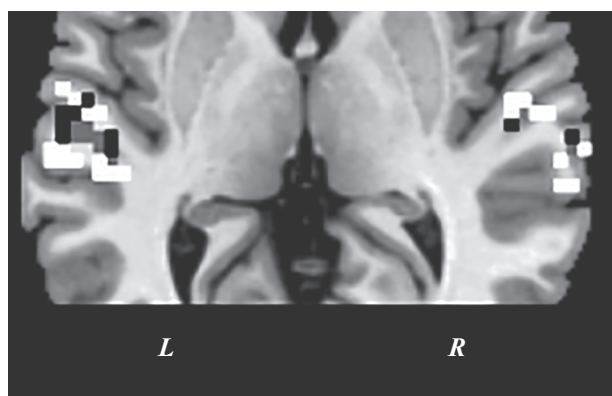
**Рис. 3.** Карты активации мозга на предъявление женского (А) и мужского (Б) голоса *one sample t-test* ( $T = 7.2$ ,  $p_{FWE} < 0.05$ ).

**Таблица 2.** Влияние пола диктора на активацию зон мозга при прослушивании звуков речевой природы

Полушарие	Область	Поле по Бродману	$x$	$y$	$z$	$T$	$p(FWE-corr)$	Размер кластера
Диктор женского пола								
$R$	Височная доля	41	63	-10	-1	10.09	0.00	20.00
		22	60	2	-1	9.70	0.00	
		41	54	-7	-1	9.01	0.00	
$R$	Височная доля	41	66	-22	11	10.01	0.00	12.00
$L$	Височная доля	41	-51	-28	8	9.62	0.00	34.00
		41	-42	-31	11	9.57	0.00	
		41	-45	-19	11	9.33	0.00	
$R$	Височная доля	22	60	-34	8	9.25	0.00	1.00
$R$	Височная доля	22	63	-31	5	8.58	0.01	2.00
$R$	Височная доля	22	42	-19	-4	8.48	0.01	2.00
$L$	Височная доля	22	-60	-1	-1	8.02	0.01	2.00
$R$	Височная доля	40	54	-25	14	7.82	0.02	3.00
$L$	Височная доля	1	-57	-16	11	7.65	0.02	1.00
$R$	Височная доля	41	48	-28	11	7.54	0.03	1.00
$L$	Височная доля	22	-63	-19	-1	7.34	0.04	2.00
Диктор мужского пола								
$R$	Височная доля	41	66	-22	8	12.25	0.00	16.00
		41	60	-16	8	8.38	0.01	
$L$	Височная доля	41	-51	-22	5	9.88	0.00	19.00
		22	-63	-19	2	8.70	0.00	
$L$	Височная доля	13	-39	-16	-1	7.91	0.02	1.00

шивании голоса своей матери относительно незнакомой.

Поскольку различие мужского и женского голосов в первую очередь зависит от частоты звука,



**Рис. 4.** Сопоставление пространственной локализации активности слуховой коры в ответ на диктора мужского и женского пола.

Слуховая кора (аксиальный срез на уровне  $z = 5$ ). Белым представлена активация на женский голос, черным — на мужской, серым — общие области активации.

карты активации в ответ на эти голоса должны примерно совпадать с томотопической картой слуховой коры. В исследовании [20] было показано, что предпочтительная зона активации низкими частотами расположена в центральной части извилин Гешля, а высокими частотами — медиально и сзади, окружая первую. Это соответствует нашим результатам, в частности, на рис. 4 видно, что зона активации на женский голос окружает зону активации на мужской.

Полученные данные в целом не противоречат данным исследованиям, подтверждая гипотезу о наличии специфического пространственного разнесения обработки сложных (речевых) звуков, в частности, для случая их формантного (гендерного) различия. При этом они вносят дополнения, показывая, что только женский голос приводит к активации в 1 (постцентральная извилина) и 40 (краевая извилина) зонах Бродмана, и только мужской активирует 13 (островковая область) зону.

*Топографическая сегрегация нейронных ансамблей слуховой коры при выделении подклассов слуховых стимулов.* В целом можно отметить, исходя из полученных данных, что в обоих случаях наблюдается наличие трех групп нейронов: две специфические и одна общая. При этом при ви-



зуальном сравнении рис. 2 и рис. 4 можно отметить, что в целом специфическая неречевая группа в случае сравнения пола диктора не активировалась (не наблюдалось кластера левой 40 зоны и двух кластеров правой 1 зоны Бродмана). При этом по всей видимости, кластер чисто речевой левой 13 зоны Бродмана преимущественно реагирует на мужской голос, а чисто речевой правой 40 зоны, и по меньшей мере 2 кластера правой 22 зоны Бродмана преимущественно реагируют на женский голос.

Таким образом, основываясь на полученных данных, можно говорить о наличии единого механизма обработки и распределения звуковой информации, который выстроен в иерархическую цепь: исходный звук проходит предобработку в общей зоне, далее попадая в зависимости от его типа в неречевую или в речевую зону, где в последней сигнал, пройдя через вторую общую зону, попадет в зависимости от пола диктора в чисто мужскую или чисто женскую зону. При этом вполне возможно существование наряду с процессом “*bottom-up*” объединения показателей низких уровней на последующих этапах, динамического влияния “*top-down*” единиц более высокого порядка, которые ограничивают и помогают идентифицировать единицы меньшего порядка. Т.е., распознавание слова или фразы помогает распознать сложные фонемы или слоги.

Концептуально это согласуется с исследованием [33], где было показано наличие развития корковой речевой сети, уточнения нейронных механизмов обработки речи в зависимости от возраста и личных фонологических навыков, что в нашем случае можно по аналогии интерпретировать на филогенетическом уровне. Как было отмечено в [34], активация зоны в *planum temporale* модулируется избирательным вниманием и не зависит от громкости одновременно звучащих голосов: это означает отделимость посредством слухового внимания объектных нейронных представлений от звукового фона. В то же время было показано [35], что имеется связь между речевыми звуками и кортикальными паттернами: в частности, предсказательная информация гласных и слогов поступает из разных нейронных кластеров (левая нижняя лобная борозда и левая вентральная премоторная кора, центральная борозда, левая средняя верхняя височная борозда соответственно). В исследовании [36] было обнаружено, что в областях средней и задней верхней височной извилины отсутствует усиление активации нейронов в ответ на речевые стимулы относительно контрольных, причем имеет место в данном случае возрастание этого усиления от средней и задней верхней височной извилины до передней средней височной извилины. И хотя точных совпадений контраста “речевые звуки > звуки животных + музыкальные инструменты”

с нашими данными найти не удалось, описанные зоны лежат рядом. Помимо этого, в данном исследовании было обнаружено существенное превышение активации нейронов во время прослушивания речи относительно других звуков в верхней средней височной извилине, чего у нас не наблюдалось, что возможно связано с различием в постановке эксперимента. В целом, описанные выше эксперименты подкрепляют гипотезу о наличии функционального пространственного разделения (специализации) групп нейронов при обработке звуковой информации, что и было продемонстрировано нашим исследованием.

## ЗАКЛЮЧЕНИЕ

Данное исследование показало, что обработка звуковой информации в слуховой коре происходит пространственно неоднородно. По всей видимости, первичный звуковой анализ происходит в тех областях, которые одинаково активируются всеми видами звуков, а вторичный анализ происходит в специфических зонах, развитие которых, скорее всего, является эволюционно обусловленным и соразмерным, с одной стороны, со сложностью решаемой задачи (такой как обработка речи), так и, с другой стороны, жизненной необходимостью (женский голос более важен для выживания вида в целом). Следует также отметить наличие специализированных зон для обработки неречевых звуков, что может свидетельствовать о равном иерархическом уровне речевой и неречевой информации с точки зрения ее обработки, т.е. о вероятном наличии разделения информационных потоков (речевого и неречевого) и их отдельной обработки (с дальнейшим разделением речевого потока), в противовес концепции последовательной обработки от простого к сложному.

**Этические нормы.** Все исследования проведены в соответствии с принципами биомедицинской этики, сформулированными в Хельсинкской декларации 1964 г. и ее последующих обновлениях, и одобрены локальными биоэтическими комитетами Центра патологии речи и нейрореабилитации и Института высшей нервной деятельности и нейрофизиологии РАН (Москва).

**Информированное согласие.** Каждый участник исследования представил добровольное письменное информированное согласие, подписанное им после разъяснения ему потенциальных рисков и преимуществ, а также характера предстоящего исследования.

**Благодарности.** Авторы выражают благодарность сотрудникам ЦПРиН за помощь в наборе стимульного материала.

**Конфликт интересов.** Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией данной статьи.

### СПИСОК ЛИТЕРАТУРЫ

1. *Лурия А.Р.* Высшие корковые функции человека и их нарушения при локальных поражениях мозга. М.: Издательство МГУ, 1962. 431 с.
2. *Diehl R.L., Lotto A.J., Holt L.L.* Speech perception // *Annual Review of Psychology*. 2004. № 55. P. 149.
3. *Zatorre R.J., Belin P., Penhune V.B.* Structure and function of auditory cortex: music and speech // *Trends Cogn. Sci.* 2002. № 6. P. 37.
4. *Joanisse M.F., Gati J.S.* Overlapping neural regions for processing rapid temporal cues in speech and non-speech signals // *NeuroImage*. 2003. № 19. P. 64.
5. *Tremblay P., Baroni M., Hasson U.* Processing of speech and non-speech sounds in the supratemporal plane: Auditory input preference does not predict sensitivity to statistical structure // *NeuroImage*. 2013. V. 66. P. 318.
6. *Marie D., Roth M., Lacoste R. et al.* Left Brain Asymmetry of the Planum Temporale in a Nonhominid Primate: Redefining the Origin of Brain Specialization for Language // *Cereb. Cortex*. 2018. V. 28. № 5. P. 1808.
7. *Zheng Z.Z., Munhall K.G., Johnsrude I.S.* Functional overlap between regions involved in speech perception and in monitoring one's own voice during speech production // *J. Cogn. Neurosci.* 2010. V. 22. № 8. P. 1770.
8. *Christoffels I.K., Formisano E., Schiller N.O.* Neural correlates of verbal feedback processing: An fMRI study employing overt speech // *Hum. Brain Mapp.* 2007. V. 28. № 9. P. 868.
9. *Hickok G., Okada K., Serences J.T.* Area Spt in the Human Planum Temporale Supports Sensory-Motor Integration for Speech Processing // *J. Neurophysiol.* 2008. V. 101. № 5. P. 2725.
10. *Zheng Z.Z.* The Functional Specialization of the Planum Temporale // *J. Neurophysiol.* 2009. V. 102. № 6. P. 3079.
11. *Griffiths T.D., Warren J.D.* The planum temporale as a computational hub // *Trends Neurosci.* 2002. V. 25. № 7. P. 348.
12. *Hawkins S.* Roles and representations of systematic fine phonetic detail in speech understanding // *J. Phon.* 2003. V. 31. P. 373.
13. *McMurray B., Tanenhaus M.K., Aslin R.N.* Gradient effects of within-category phonetic variation on lexical access // *Cognition*. 2002. V. 86. P. B33.
14. *Mottonen R., Calvert G., Jaaskelainen I. et al.* Perceiving identical sounds as speech or non-speech modulates activity in the left posterior superior temporal sulcus // *NeuroImage*. 2006. № 30. P. 563.
15. *Petrides M., Pandya D.N.* Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey // *Eur. J. Neurosci.* 2002. V. 16. № 2. P. 291.
16. *Romanski L.M., Averbach B.B.* Neural representation of vocalizations in the primate ventrolateral prefrontal cortex // *J. Neurophysiol.* 2005. V. 93. P. 734.
17. *Romanski L.M., Goldman-Rakic P.S.* An auditory domain in primate prefrontal cortex // *Nat. Neurosci.* 2002. V. 5. № 1. P. 15.
18. *Fecteau S.* Sensitivity to Voice in Human Prefrontal Cortex // *J. Neurophysiol.* 2005. V. 94. № 3. P. 2251.
19. *Joassin F., Maurage P., Campanella S.* The neural network sustaining the crossmodal processing of human gender from faces and voices: An fMRI study // *NeuroImage*. 2011. V. 54. № 2. P. 1654.
20. Welcome. Trust. Centre. for Neuroimaging: <http://www.fil.ion.ucl.ac.uk/spm>.
21. *Friston K.J., Holmes A.P., Worsley K.J. et al.* Statistical parametric maps in functional imaging: A general linear approach // *Hum. Brain Mapp.* 1994. V. 2. № 4. P. 189.
22. *Wilke M., Schmithorst V.J.* A combined bootstrap/histogram analysis approach for computing a lateralization index from neuroimaging data // *NeuroImage*. 2006. V. 33. № 2. P. 522.
23. *Wilke M., Lidzba K.* LI-tool: A new toolbox to assess lateralization in functional MR-data // *J. Neurosci. Methods*. 2007. V. 163. № 1. P. 128.
24. *Chan A., Dykstra A., Jayaram V. et al.* Speech-specific tuning of neurons in human superior temporal gyrus // *Cereb. Cortex*. 2014. V. 24. № 10. P. 2679.
25. *Fan C.S.D., Zhu X., Dosch H.G. et al.* Language related differences of the sustained response evoked by natural speech sounds // *PLoS ONE*. 2017. V. 12. № 7. P. e0180441.
26. *Wernicke C.* The Symptom Complex of Aphasia // *Proceedings of the Boston Colloquium for the Philosophy of Science 1966/1968*. 1969. V. 4. P. 34.
27. *Luria A.R.* Traumatic aphasia (Translated from the Russian). The Hague, Paris: Mouton, 1970. 479 p.
28. *Andermann M., Patterson R.D., Vogt C. et al.* Neuro-magnetic correlates of voice pitch, vowel type, and speaker size in auditory cortex // *NeuroImage*. 2017. V. 158. P. 79.
29. *Bonte M., Hausfeld L., Scharke W. et al.* Task-Dependent Decoding of Speaker and Vowel Identity from Auditory Cortical Response Patterns // *J. Neurosci.* 2014. V. 34. № 13. P. 4548.
30. *Formisano E., De Martino F., Bonte M., Goebel R.* "Who" is saying "what"? Brain-based decoding of human voice and speech // *Science*. 2008. V. 322. № 5903. P. 970.
31. *Liu P., Cole P., Gilmore R. et al.* Young children's neural processing of their mother's voice: An fMRI study // *Neuropsychologia*. 2019. V. 122. P. 11.
32. *Gardumi A., Ivanov D., Havlicek M. et al.* Tonotopic maps in human auditory cortex using arterial spin labeling // *Hum. Brain Mapp.* 2017. V. 38. № 3. P. 1140.
33. *Bonte M., Ley A., Scharke W., Formisano E.* Developmental refinement of cortical systems for speech and voice processing // *NeuroImage*. 2016. V. 128. P. 373.
34. *Simon J.Z.* The encoding of auditory objects in auditory cortex: Insights from magnetoencephalography // *Int. J. Psychophysiol.* 2015. V. 95. № 2. P. 184.
35. *Markiewicz C.J., Bohland J.W.* Mapping the cortical representation of speech sounds in a syllable repetition task // *NeuroImage*. 2016. V. 141. P. 174.
36. *Bethmann A., Brechmann A.* On the definition and interpretation of voice selective activation in the temporal cortex // *Front. Hum. Neurosci.* 2014. V. 8. P. 499.

## Speech and Non-Speech Sound Categorization in Auditory Cortex: fMRI Correlates

V. M. Shklovsky<sup>a</sup>, S. A. Varlamov<sup>b, c, \*</sup>, A. G. Petrushevsky<sup>a</sup>, L. A. Mayorova<sup>a, c, \*\*</sup>

<sup>a</sup>*Center for Speech Pathology and Neurorehabilitation of the Moscow Department of Health, clinical base of the Federal Medical Research Center for Psychiatry and Addictology named after Serbsky Ministry of Health of Russia, Moscow, Russia*

<sup>b</sup>*Lomonosov Moscow state university, Moscow, Russia*

<sup>c</sup>*Institute of Higher Nervous Activity and Neurophysiology of the Russian Academy of Sciences, Moscow, Russia*

\**E-mail: sv-kanino@mail.ru*

\*\**E-mail: larimayor@gmail.com*

We studied the functional structure of the auditory cortex by identifying and comparing the spatial localization of activation zones in response to speech and non-speech stimuli using functional magnetic resonance imaging (fMRI). We also performed a similar comparison of activation zones in response to male and female voices. We found that there are specific zones for speech and non-speech auditory stimuli and overlapping zones; the speech zone is significantly larger as compared with others. The activation zones responding to male and female voices are overlapping, although not significantly; the influence of female voice was stronger. These results suggest that there are special areas in the auditory cortex for auditory signal processing.

*Keywords:* speech perception, supratemporal cortex, fMRI, planum temporale.