

«Московский государственный университет имени М.В. Ломоносова»

Факультет психологии

На правах рукописи

ШЕВЧЕНКО АНДРЕЙ ОЛЕГОВИЧ

**МОЗГОВЫЕ МЕХАНИЗМЫ ВНУТРЕННЕГО
ПРОГОВАРИВАНИЯ ФОНЕМ И СЛОГОВ В НОРМЕ**

5.3.2. Психофизиология (психологические науки)

Диссертация

на соискание ученой степени

кандидата психологических наук

Научный руководитель:

кандидат психологических наук,

старший научный сотрудник

А.В. Вартанов

Москва– 2024

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	5
ГЛАВА 1. ТЕОРЕТИЧЕСКИЙ АНАЛИЗ ПРОЦЕССОВ ВНУТРЕННЕГО ПРОГОВАРИВАНИЯ	16
1.1. Внутреннее проговаривание в структуре речевой деятельности.....	16
1.2. Модели речепроизводства	20
1.2.1. Иерархическая модель В. Левелта	22
1.2.2. Сетевая модель Г.С. Делла.....	26
1.3. Фонематическая система речи и артикуляция	28
1.4. Анализ мозговых механизмов речевой деятельности	35
1.4.1. Конвенциональные речевые зоны	35
1.4.2. Таламус и мозжечок как вторичные посредники речевой деятельности.....	39
1.4.3. Модели мозгового взаимодействия при восприятии и порождении речи.....	42
1.5. Подходы к построению классификаторов внутреннего проговаривания	47
1.5.1. Классификация фонематических стимулов методом ЭЭГ ...	50
1.5.2. Миографические подходы к классификации внутреннего проговаривания	54
1.6. Постановка исследовательской проблемы, гипотезы и схема исследования.....	56
ГЛАВА 2. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ	64
2.1. Методика исследования внутреннего проговаривания методом электромиографии	64
2.2. Результаты исследования внутреннего проговаривания методом электромиографии	65

2.3. Методика исследования внутреннего проговаривания методом электроэнцефалографии.....	67
2.4. Результаты исследования внутреннего проговаривания методом электроэнцефалографии.....	73
2.4.1. Результаты исследования внутреннего проговаривания методом электроэнцефалографии на визуальную и аудиальную инициализацию	73
2.4.2. Результаты исследования внутреннего проговаривания на инициализацию условным стимулом методом электроэнцефалографии	81
2.4.3. Построение классификатора на основе метода опорных векторов (SVM-RE).....	81
2.4.4. Построение классификатора на основе сверточных нейронных сетей	86
2.5. Исследование внутреннего проговаривания методом функциональной магнитно-резонансной томографии.....	91
2.6. Результаты исследования внутреннего проговаривания методом функциональной магнитно-резонансной томографии	93
2.7. Построение фМРТ-связности с помощью анализа независимых компонент	99
ГЛАВА 3. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ.....	103
3.1. Выделение связанного с событием потенциала внутреннего проговаривания и условия различных инициализаций	103
3.2. Сравнение внутреннего проговаривания фонем и слогов с их восприятием	106
3.3. Сравнение внешнего проговаривания и внутреннего проговаривания	110
3.4. Мозжечок в процессах внутреннего проговаривания.....	114
3.5. Интегративная модель внутреннего проговаривания и восприятия фонем и слогов.....	117

3.6. Построение классификатора как область практико-ориентированных исследований внутреннего проговаривания.....	121
ЗАКЛЮЧЕНИЕ.....	124
ВЫВОДЫ.....	129
СПИСОК СОКРАЩЕНИЙ.....	131
СПИСОК ЛИТЕРАТУРЫ.....	132
ПРИЛОЖЕНИЕ 1. СХЕМА РАСПОЛОЖЕНИЯ ЭЛЕКТРОДОВ ЭЛЕКТРОМИОГРАФИИ.....	146
ПРИЛОЖЕНИЕ 2. ПРИМЕРЫ КЛАССИФИКАЦИИ.....	147
ПРИЛОЖЕНИЕ 3. ТАБЛИЦЫ СРАВНЕНИЯ ФМРТ-СВЯЗНОСТИ СЕРИЙ.....	151
ПРИЛОЖЕНИЕ 4. ФМРТ ИЗОБРАЖЕНИЯ АКТИВНОСТИ ГОЛОВНОГО МОЗГА В СЕРИИ ФМРТ-СВЯЗНОСТИ.....	158

ВВЕДЕНИЕ

Актуальность исследования. В отечественной и зарубежной психологии актуальным является изучение процессов внутреннего проговаривания (П.Я. Гальперин, К. Парнин, А.Н. Соколов и др.), как одного из фундаментальных механизмов переработки и усвоения информации с помощью повторения во внутреннем плане внешних элементов. Так, П.Я. Гальперин выделяет процессы внутреннего проговаривания и определяет внутреннее проговаривание как скрытую внешнюю речь или внешнюю речь про себя (Гальперин, 2000). Внутреннее проговаривание – «форма скрытой речевой активности, близкая к внешней речи» (Азимов, Щукин, 2009, с.41). Существующие определения сопоставляют внешнюю речь и внутреннее проговаривание, в результате чего можно выделить ряд свойств внутреннего проговаривания для изучения. В отличие от внешнего проговаривания, в процессе внутреннего проговаривания происходит подавление конечной (полной) артикуляции. Многие исследователи отмечают наличие скрытых, неполных двигательных речевых актов и во внутреннем проговаривании. А.Н. Соколов в своих исследованиях внутреннего проговаривания и внутренней речи обращал внимание на речедвигательный компонент – зачаточную артикуляцию (Соколов, 1984). В исследованиях Г.М. Оппенгейма и Г.С. Делла (Oppenheim, Dell, 2010) говорится о влиянии артикуляции на фонологический уровень внутреннего проговаривания. Известно, что артикуляция, кроме конвенциональных речевых зон, включает и подкорковые структуры, и мозжечок (Indefrey, Levelt, 2004; Ackermann et al., 2007; Wildgruber et al., 2001 и др.), однако в настоящее время нет исследования, в котором бы проводилось сравнение внешнего и внутреннего проговаривания минимальных речевых единиц с оценкой вовлеченности мозговых структур. Кроме этого, можно выделить ряд уникальных свойств внутреннего проговаривания: мотивация к внутреннему проговариванию, внутренняя репрезентация проговариваемой фонемы или слога (внутренний план), степень артикуляции – все эти свойства должны иметь свои мозговые

механизмы. Все эти этапы релевантны и для любой другой речевой деятельности: механизм обеспечения ощущений для речевой функции «должен быть внутримозговым и иметь своей основой единую интегрированную систему связей проекционных и ассоциативных зон коры с речевыми» (Иваницкий, 1997, с. 224).

И фонематическое восприятие, и внутреннее проговаривание становятся процессами, которые определяются тесной связью с удержанием речевого образа. Сравнивая процессы восприятия и внутреннего проговаривания, мы можем полагать, что при внутреннем проговаривании происходит произвольное формирование внутреннего образа проговариваемой единицы, что является одним из свойств внутреннего проговаривания. В процессе восприятия также создается образ речевой единицы, однако воссоздание образа будет непроизвольным. Таким образом, для оценки произвольности, как свойства внутреннего проговаривания, релевантно сравнить намеренное воссоздание образа проговариваемой единицы при внутреннем проговаривании с непроизвольным внутренним отражением в процессе восприятия этих единиц. Кроме этого, произвольность должна быть детерминирована мотивом на произнесение.

В процессе формирования образа произносимой единицы возникает такое свойство внутреннего проговаривания, как полимодальность – способность совмещать в образе несколько модусов познания. Это свойство должно характеризоваться активностью различных зон мозга, связанных с разными модальностями, например, при внутреннем проговаривании возможно включение зон, связанных со зрительным восприятием, в связи с ассоциациями и навыками письма и чтения у взрослых респондентов. Известно, что порождение речевого высказывания включает несколько этапов: этап идеи или мотива речевого высказывания, замысел, лексический подбор, грамматическое и синтаксическое оформление фразы, артикуляция необходимых слов (Седов, Горелов, 2001; Levelt et al., 1999; Dell, 1986; Oppenheim, Dell, 2008, 2010). В настоящее время не существует моделей,

которые описывают процесс внутреннего проговаривания. В соответствии с этим полимодальность, как свойство внутреннего проговаривания, и этапы кодирования ставят перед исследователями два важных вопроса: вопрос о влиянии задающего стимула на процесс программирования и реализации внутреннего проговаривания, а также вопрос о возможных временных интервалах связанных с событием потенциалов, на которых инициализация – команда, задающая фонему или слог – проявляется сильнее. Выявление мозговых механизмов внутреннего проговаривания на основе электрической активности мозга является важной проблемой для усовершенствования построений интерфейсов «мозг – компьютер». Таким образом, изучение и сравнение электрофизиологических показателей внутреннего проговаривания на разных его этапах является не только актуальной научной, но и практико-обусловленной задачей.

В настоящем психофизиологическом исследовании предполагается изучить внутреннее проговаривание на психологическом уровне с учетом данных о закономерностях восприятия речи и ее артикуляторных механизмах, исследовать мозговые механизмы процесса внутреннего проговаривания по данным электроэнцефалографии и функциональной магнитно-резонансной томографии. В работе для оценки мозговых механизмов внутреннего проговаривания используются задачи, исключаящие синтаксическое лексическое кодирование. Актуальным становится изучение внутреннего проговаривания фонем и слогов как минимальных речевых единиц. Применение ряда психофизиологических методов позволяет комплексно оценить процессы внутреннего проговаривания. Метод электроэнцефалографии обладает свойствами хорошего временного разрешения, что важно для дифференциации различных этапов внешнего и внутреннего проговаривания. Метод функциональной магнитно-резонансной томографии позволяет оценить вовлеченность различных зон коры больших полушарий и подкорковых структур в процессы внутреннего проговаривания и восприятия с хорошим пространственным разрешением. С помощью

электромиографии возможно зафиксировать наличие и свойства внутреннего проговаривания, а также временные интервалы его реализации.

Парадигма построения интерфейсов «мозг – компьютер» способствует усовершенствованию методов выделения паттернов, дифференцирующих по мозговой активности проговариваемые про себя речевые единицы. Известно, что восстановление внутреннего проговаривания на основе данных биоэлектрической активности мозга является возможным (Yamaguchi et al., 2015; Sereshkeh, 2017; Porbadnigk, Wester, Calliess, 2009; Brigham, Kumar, 2010; DaSalla et al., 2009; Matsumoto, Hori, 2014), поскольку в процессе внутреннего проговаривания генерируется активность артикуляторов, а также головного мозга.

Степень научной разработанности проблемы.

В исследованиях не раз изучалась внутренняя речь как психологический феномен (Б.Г. Ананьев, Т.В. Ахутина, П.П. Блонский, Л.С. Выготский, П.Я. Гальперин, А.А. Леонтьев, А.Н. Соколов, R. Dodge, C. Parnin). Существуют различные подходы к пониманию внутреннего проговаривания. Так, исследованиями в рамках выделения внутреннего проговаривания, как одной из форм внутренней речи, занимались такие ученые, как П.Я. Гальперин, А.Н. Соколов, C. Parnin и др. В большинстве моделей речепроизводства не выделяют отдельно внутреннюю речь и внутреннее проговаривание (И.Н. Горелов, К.Ф. Седов, G.S. Dell, W.J.M. Levelt, G.M. Oppenheim и др.). В области психофизиологии вопросом локализации речевых функций занимались многие ученые (Н.А. Бернштейн, П. Брока, К. Бродман, К. Вернике, А.Р. Лурия, Z. Breznitz, M. Fujii, A.J. Zhong и др.), однако отдельных нейровизуализационных исследований процессов внутреннего проговаривания намного меньше (B. Alderson-Day, G.S. Dell, A. Miyake, G.M. Oppenheim, M. Scott и др.), при этом они чаще всего изучают семантическое кодирование, что не является отражением процессов только внутреннего проговаривания, так как вовлекаются зоны, ассоциированные с семантическими процессами. При этом существует множество подходов,

которые рассматривают вовлеченность различных «неречевых» зон и структур в речевую деятельность во внешней речи, в том числе на уровне фонематических единиц (Н. Ackermann, А.А. Bohsali, А.А. Ford, Р. Indefrey, W.J.M. Levelt, А. Starowicz-Filip, D. Wildgruber и др.).

Большим достижением в области реконструкции внутреннего проговаривания стала разработка инвазивных методов классификации фонем и слогов (А.Т. Петросян, М. Angrick, S. Martin, N. Mesgarani, B.N. Pasley и др.). В области реконструкции также используются неинвазивные методы, например, метод электроэнцефалограммы. Существует много исследований по выделению и классификации проговариваемых фонем и слогов с разным процентом точности распознавания (точность распознавания – 65% Ямагучи с соавт.; точность распознавания – 69–75% Серешке и соавт.; точность распознавания – 44% Бригама и Кумара; точность распознавания – 68–78% Да Салла и соавт., др.). Данное диссертационное исследование детализирует информацию о локализации внутреннего проговаривания минимальных языковых единиц и соединяет эти данные с электроэнцефалографическими исследованиями в единую модель.

Цель: выявление мозговых механизмов внутреннего проговаривания фонем и слогов в норме.

Объект исследования: внутреннее проговаривание фонем и слогов в норме.

Предмет исследования: мозговые механизмы внутреннего проговаривания фонем и слогов в норме.

Общая гипотеза. Реализация и мониторинг процесса внутреннего проговаривания фонем и слогов обеспечивается специальной мозговой системой, которая отличается от системы реализации и мониторинга при внешнем проговаривании или восприятии фонем и слогов.

Частные гипотезы:

1. Процессы внутреннего проговаривания, в отличие от внешнего проговаривания, в большей степени связаны с формированием зрительного образа проговариваемой единицы.
2. Подкорковые структуры более активны во внешнем проговаривании, чем во внутреннем проговаривании, в связи с большей включенностью артикуляционных процессов.
3. Условия инициализации внутреннего проговаривания визуально, аудиально или посредством условной ассоциации влияют на ранние характеристики связанного с событием потенциала, тогда как его поздние компоненты сходны вне зависимости от условий инициализации.
4. По амплитудно-временным характеристикам связанного с событием потенциала можно классифицировать проговаривание отдельной фонемы или слога русского языка с помощью алгоритмов машинного обучения.

Задачи исследования:

1. Провести теоретический анализ существующих исследований внутреннего проговаривания.
2. Выявить артикуляционную активность, как индикатор внутреннего проговаривания, с помощью электромиографического исследования.
3. Выявить ассоциированные с внутренним проговариванием временные интервалы при разных типах условий, задающих проговаривание, с помощью электроэнцефалографического исследования.
4. Классифицировать отдельные фонематические единицы внутреннего проговаривания с помощью метода машинного обучения по данным электроэнцефалографии.
5. Выявить функциональные характеристики процессов внутреннего и внешнего проговаривания, а также восприятия фонем и слогов с помощью функциональной магнитно-резонансной томографии.

- б. На основе теоретического анализа и проведенных эмпирических исследований создать модель внутреннего проговаривания фонем и слогов.

Методологическая основа исследования. Исследование опирается на следующие теоретические представления: принцип психофизиологического исследования «Человек-Нейрон-Модель» Е.Н. Соколова, моторную концепцию внутренней речи А.Н. Соколова, иерархическую модель построения речи В. Левелта.

Методы и методики исследования:

1. Метод электроэнцефалографии с регистрацией на 19-канальном электроэнцефалографе «Нейро-КМ» (компания «Статокин», Россия).
2. Метод электромиографии на базе системы «Нейро-КМ» по двум отведениям.
3. Метод функциональной магнитно-резонансной томографии с помощью томографа Siemens Magnetom Skyra 3 TMRI Machine (Siemens Medical Systems, Эрланген, Германия). Исследование проводилось на базе ФГБУ «Национальный медико-хирургический Центр им. Н.И. Пирогова».

Характеристика выборки. Всего в исследовании приняло участие 62 человека. В электроэнцефалографическом исследовании участниками стали 25 человек: 15 участников женского пола и 10 участников мужского пола в возрасте от 18 до 28 лет; средний возраст – 20,8 лет. В апробационном исследовании методом электромиографии приняли участие 7 человек: 4 участника женского пола и 3 участника мужского пола в возрасте от 18 до 23 лет; средний возраст – 19,6 лет. Исследование методом функциональной магнитно-резонансной томографии прошли 30 участников: 9 мужского пола и 21 женского пола в возрасте от 20 до 30 лет; средний возраст – 24,0 года. Все участники исследования не имели в анамнезе черепно-мозговых травм и психоневрологических заболеваний. Все участники были праворукими. Все участники подписали информированное согласие после ознакомления с

регламентом исследования. Исследование проходило с соблюдением этических принципов Российского Психологического Общества.

Научная новизна результатов исследования.

В диссертационном исследовании разработана интегративная модель внутреннего проговаривания фонем и слогов, которая является развитием иерархической модели В. Левелта и теории моторного кодирования внутренней речи А.Н. Соколова. Представленная модель включает этап ориентировочной артикуляционной предобработки и детализирует иерархическое формирование единого фонематического образа.

Получены новые электроэнцефалографические данные, показывающие психофизиологические различия в процессах внутреннего проговаривания фонем и слогов в зависимости от типа инициализации. Выявлено, что наиболее дифференцирующими являются ранние и средние компоненты связанного с событием потенциала. Так, обнаружены отличия в реакции на аудиальный стимул по сравнению с визуальным, при этом реакция на «условный стимул» (подаваемый аудиально) похожа на реакцию на визуальный. Схожесть «условного стимула» и визуального объясняется тем, что респонденту приходится самостоятельно восстанавливать образ звука в процессе концептуальной обработки, так как при аудиальной инициализации многие характеристики звукового образа (например, ритмика, частотные характеристики) уже имеются, в отличие от визуальной инициализации, где необходимо их воссоздать самостоятельно.

С помощью метода функциональной магнитно-резонансной томографии впервые получены данные, показывающие различия в процессах внешнего и внутреннего проговаривания фонем и слогов в ряде структур головного мозга. Для воссоздания полного образа проговариваемой единицы (аудиальных, кинестетических, визуальных характеристик) в процессе внутреннего проговаривания требуется участие верхней теменной доли, угловой извилины, нижней затылочной извилины, веретенообразной извилины и мозжечка. Только при внешней речи активна зона передней поясной

извилины, что позволяет выделить эту структуру как значимую зону в сети обратной слуховой связи.

На основе данных электроэнцефалографии с помощью методов машинного обучения созданы классификаторы фонем и слогов русского языка при различных условиях инициализации. Впервые проведено сравнение качества классификаторов фонем русского языка на основе метода опорных векторов и сверточных нейронных сетей, что является значимым основанием для разработки и усовершенствования интерфейсов «мозг – компьютер».

Теоретическая значимость результатов исследования. В работе представлена интегративная модель внутреннего проговаривания фонем и слогов, расширяющая модель В. Левелта и моторную теорию внутренней речи А.Н. Соколова. Представленная модель включает такие блоки, как мотивационный блок, блоки ориентировочной артикуляционной предобработки, концептуальной детекции/концептуальной обработки, фонематического кодирования и артикуляции. В работе представлены данные о различиях условий инициализации, выраженные в ранних и средних компонентах связанных с событием потенциалов.

Кроме того, полученные в ходе исследования данные помогают расширить существующие знания в области изучения речи об участии различных зон мозга в процессе внутреннего проговаривания. В работе описана роль мозжечка как вторичного речевого посредника. Данные, полученные в исследовании, показали, что передняя поясная кора наиболее активна во внешней речи в сравнении с внутренним проговариванием. Это может говорить о том, что передняя поясная кора является значимой областью в сети слухового контроля за произносимой речью. Отмечено участие верхней теменной доли, угловой извилины, нижней затылочной извилины и веретенообразной извилины в процессах внутреннего проговаривания. Активность этих зон связана с тем, что во внутреннем плане происходит естественная полимодальная связь визуальных и аудиальных отделов, которые совместно формируют единый образ.

Практическая значимость результатов исследования. Результаты диссертационного исследования могут использоваться при диагностике различных функциональных нарушений. Полученные данные о роли зрительных и конвенциональных речевых зон, а также подкорковых структур (включая мозжечок) при внутреннем проговаривании минимальных речевых единиц подтверждают положение об опоре на различные анализаторы при коррекции речевых расстройств. Результаты исследования показывают, что использование полимодального подхода в коррекции может улучшать коррекционный процесс.

Кроме этого, информация о локализации процессов внутреннего проговаривания может использоваться в практике построения интерфейсов «мозг – компьютер». Одним из направлений при работе над нейроинтерфейсами является пространственная фильтрация данных электроэнцефалограммы с целью улучшения распознавания проговариваемых про себя речевых единиц. Активно развиваются технологии классификации элементов внутреннего проговаривания, и одним из подходов является использование методов сетевой локализации процессов проговаривания (Santos et al., 2021; Frikha et al., 2021). При этом использование интерфейсов возможно не только в дешифровке и декодировании речи, но и при столкновении с речевыми нарушениями. Также нейроинтерфейсы распознавания внутреннего проговаривания могут быть применены в индустрии развлечений, например, в устройствах виртуальной реальности (Гавриленко и соавт. 2019). Созданный классификатор на основе полученных данных электроэнцефалографии может быть использован для распознавания речевых стимулов и является актуальной технологией в области неинвазивного речевого ввода.

Надежность и достоверность полученных результатов обеспечивается применением методов регистрации, обработки и анализа данных, адекватных предмету и задачам исследования; использованием при

обработке и анализе данных современного программного обеспечения и статистических методов, отвечающих специфике эмпирических данных.

Положения, выносимые на защиту:

1. Условия инициализации, которые определяют содержание проговариваемого элемента, отражаются в ранних характеристиках связанного с событием потенциала, что связано с механизмами артикуляционной предобработки в процессе начала проговаривания.
2. Мозжечок больше активируется при реализации процессов внутреннего проговаривания в сравнении с внешним проговариванием, при этом произнесение во внутреннем плане речевых элементов разного уровня – фонем или слогов – активирует различные области мозжечка.
3. Процессы внутреннего проговаривания, в том числе при слуховой инициализации, более связаны с активацией зрительных зон мозга в сравнении с внешним проговариванием фонем и слогов.

Апробация работы. Результаты работы обсуждались на заседаниях кафедры психофизиологии факультета психологии МГУ имени М.В.Ломоносова. Основные положения и результаты исследования докладывались на: XXVI и XXVII Международной конференции студентов, аспирантов и молодых ученых «Ломоносов-2019» и «Ломоносов-2020» (Севастополь, 2019; Москва, 2020); Annual International Conference on Brain-Inspired Cognitive Architectures for Artificial Intelligence (Натал, Бразилия, 2020; Гвадалахара, Мексика, 2022); I Национальном Конгрессе по когнитивным исследованиям, искусственному интеллекту и нейроинформатике (Москва, 2020); XII Международной научной конференции «Интеллектуальные системы и компьютерные науки» (Москва, 2021).

Структура и объем диссертации. Работа состоит из введения, трех глав, заключения, выводов, списка сокращений, списка литературы и приложений. Текст диссертации изложен на 165 страницах и включает в себя 41 рисунок, 11 таблиц. Список литературы включает 158 публикаций, из них 54 – на русском языке, 104 – на английском языке.

ГЛАВА 1. ТЕОРЕТИЧЕСКИЙ АНАЛИЗ ПРОЦЕССОВ ВНУТРЕННЕГО ПРОГОВАРИВАНИЯ

1.1. Внутреннее проговаривание в структуре речевой деятельности

Традиционно речь разделяется на внешнюю и внутреннюю. Под внешней речью понимается озвученная речь или же речь, облаченная в звуки (Розенталь, Теленкова, 1976). К пониманию внутренней речи существуют различные подходы: она может описываться как неозвученная внешняя речь (с теми же механизмами) или же как отдельные механизмы, реализующиеся с помощью иной мозговой цепи. При рассмотрении внутренней речи важно отметить работы Л.С. Выготского и П.П. Блонского, которые понимали внутреннюю речь как механизм словесного мышления и словесно-логической памяти. Л.С. Выготский определял внутреннюю речь как своеобразное психологическое явление, живой процесс рождения мысли в слове, возникающий из эгоцентрической речи и характеризующийся фрагментарностью, упрощением синтаксиса и полной предикативностью (Выготский, 1956). Л.С. Выготский характеризовал внутреннюю речь через ее идиоматичность и предикативность, при этом отвергая бихевиористическое понимание (обеззвучивание произносимой речи), а также критиковал определение внутренней речи исключительно как вербальной памяти (Выготский, 1966). П.П. Блонский, критикуя гипотезы Л.С. Выготского, считал, что внутренняя речь возникает из взаимного общения с людьми одновременно с внешней речью, но также отмечал ее чрезвычайную фрагментарность и незаконченность (Блонский, 1964). Ученый настаивал на том, что внешняя и внутренняя речь возникают одновременно: в его исследованиях отмечались такие характеристики речи, как фрагментарность, незаконченность, недоразвитость и необходимость к репродукции, обоснованная тем, что при внимательном прослушивании речи наличие сторонних мыслей становится невозможным (Блонский, 1935). Б.М. Теплов связывал проблемы внутренней речи с необходимостью изучения

«внутреннего пения» и «внутреннего слуха» и утверждал, что внешне выраженные или зачаточные движения голосовых связок или пальцев имеют большое влияние на работу «внутреннего слуха» и «внутреннего пения» (Теплов, 1947; Соколов, 1967).

Элементы беззвучных речевых процессов проявляются и в произвольных, и в сознательных действиях в виде самоинструкций, вербальной интерпретации восприятий или речевых установках (Соколов, 1984). А.Н. Соколов под внутренней речью понимал базовый механизм, характерный для процессов мышления. В качестве функций внутренней речи исследователь выделял следующий ряд операций: обобщение, отбор и хранение сенсорной информации. Важным является тот факт, что в рамках данной теории мышление не тождественно речи, поскольку, представляя собой постоянное взаимодействие сенсорной и вербальной информации, процесс мышления включает в себя не только речевой компонент (Соколов, 1984, Верани, 2010). Элементы беззвучных речевых процессов проявляются и в произвольных, и в сознательных действиях в виде самоинструкций, вербальной интерпретации восприятий или речевых установках (Соколов, 1984, Верани, 2010).

Внутреннее проговаривание могут выделять как один из типов внутренней речи, определяя внутреннее проговаривание как «речь про себя», сохраняющая структуру внешней речи, но без фонации, типичная для решения мыслительных задач в затрудненных условиях (Азимов, Щукин, 2009, с.42). При этом ряд авторов связывают внутреннюю речь и внутреннее проговаривания, как параллельные процессы. В работе К. Парнина (Parnin, 2011) внутренняя речь представляет собой беззвучную мысленную речь, которая сопровождается мыслительным процессом. Автор связывает процесс внутренней речи с процессом внутреннего проговаривания через процесс чтения про себя. Во время этого и других подобных процессов движения артикуляционных органов происходят незаметно для невооруженного глаза (Parnin, 2011). Р. Додж (Dodge, 1896) в своем исследовании эмпирически

показал, что при обезболивании и достижении онемения своих губ и языка не обнаруживается какого-либо влияния на внутреннюю речь (Dodge, 1896). Движения, осуществляемые органами речевого аппарата, не являются необходимыми для процесса мышления, однако они становятся конечным получателем сигналов двигательной и премоторной коры головного мозга. В настоящее время в науке существует спор о роли моторного кодирования в акте внутреннего проговаривания, который до сих пор является не разрешенным.

Во время внутреннего проговаривания органы речевого аппарата воспроизводят процессы артикуляции как во время внешней речи, однако осуществляются они в скрытой форме. П.Я. Гальперин определял внутреннее проговаривание как скрытую внешнюю речь или внешнюю речь про себя (Гальперин, 2000). Внутреннее проговаривание связано с умственными действиями, которые еще не автоматизировались, чаще всего оно возникает при решении математических задач, при переводе иностранных текстов и так далее (Глухов, 2005). Внутреннее проговаривание – форма скрытой речевой активности, близкая к внешней речи (Азимов, Щукин, 2009, с.41). Еще одним термином, описывающим процесс внутреннего проговаривания, является субвокализация – только мысленное проговаривание, которое позволяет вообразить звучание слов. Субвокализацию, в первую очередь, относят к механизмам чтения. Говоря об отношениях между внутренним проговариванием и субвокализацией, следует отметить, что внутреннее проговаривание может включать в себя субвокализацию: в процессе внутреннего проговаривания органы периферического речевого аппарата воспроизводят процессы, характерные для внешней речи, в скрытой форме. Данный процесс связан с умственными задачами, реализующимися в непроизвольной форме.

В свою очередь А.Н. Соколов в своих исследованиях внутренней речи и внутреннего проговаривания обращал внимание на речедвигательный компонент – зачаточную артикуляцию, микродвижения речевых органов как

константу повышения активности и тонуса их мускулатуры. По данным электромиографических исследований, при мыслительной деятельности выделяется два вида речедвигательных реакций: тонические и фазические (Соколов, 1984).

В соответствии с трудами А.Н. Соколова, внутренняя речь является психологической трансформацией внешней, с которой она связана функционально. Если для озвучивания сообщения необходима отсрочка, то посредством внутренних речевых процессов информация фиксируется в форме мыслительного конспекта, плана высказывания. Данная особенность наблюдается и при письменной передаче информации: перед тем, как написать слово или фразу, происходит их мысленное (внутреннее) проговаривание. А.Н. Соколов изучал процессы внутренней речи и субвокализации посредством исследовательской программы, включающей в себя запись ЭМГ (электромиографии) во время перевода текста испытуемыми. Во время исследования он отметил некоторые особенности эффекта подавления артикуляции у испытуемых. Подавление артикуляции привело к уменьшению точности выполнения перевода, однако результаты ЭМГ отличались не только у разных испытуемых, но и у одного и того же испытуемого, выполняющего повторно то же задание. Таким образом, можно прийти к выводу, что люди могут варьировать степень артикуляции в зависимости от внимания и степени обучения (Соколов, 1967). Например, известно, что в процессе овладения навыком письма дети долгое время проговаривают записываемое слово сначала вслух, а потом сворачивают его в скрытое внутреннее проговаривание, о чем свидетельствуют регистрируемые тонкие движения гортани в процессе письма (Лурия, 2002).

Субвокализация тесно связана с репетиционными (повторение фраз перед тем, как сказать) процессами для понимания, сохранения в памяти и концентрации внимания на целях (Smith, 1995; Miyake, 2004). Во время субвокализации активируются дополнительные мозговые пути, эти активации стимулируют и ускоряют запоминание при чтении (McGettigan et al., 2011).

Так, при переводе слова необходимо не только извлечь значение слова, но и восстановить контекст, связанный с ним.

Г.М. Оппенгейм и Г.С. Делл (Oppenheim, Dell, 2010) провели исследование, чтобы понять, как артикуляционное сходство влияет на фонологические ошибки в процессе внутреннего проговаривания, а также исследовать влияние лексического предубеждения. Лексическое предубеждение – это тенденция фонологических речевых ошибок чаще образовывать слова, чем неслова (Oppenheim, Dell, 2010). В ходе эксперимента участникам предлагалось произносить фразы, в которых ошибки могли проявить эффект фонемного сходства и эффект лексической предвзятости. Произнесение фраз осуществлялось с внутренним проговариванием и с внутренним называнием. Авторы пришли к выводу, что артикуляция изменяет внутреннее проговаривание, которое, в свою очередь, не может быть независимым от артикуляции, используемой человеком для выражения (Oppenheim, Dell, 2010).

Так, становится заметной связь между артикуляционным актом и внутренним проговариванием. При этом внутреннее проговаривание возможно и без включения явной артикуляции, ведь в структуре внутреннего проговаривания существуют различные уровни артикуляторной активации. Кроме того, во внутреннем плане должен сохраняться темп и интенсивность речи – эта особенность представлена внутренним образом проговариваемых единиц.

1.2. Модели речепроизводства

Описание процессов внутреннего проговаривания и внутренней речи может быть интерпретировано средствами концептуального моделирования. Многие ученые считают, что порождение речевого высказывания – это процесс с определенной иерархией. Ряд авторов полагает, что порождение речевого высказывания включает следующие этапы: этап идеи или мотива речевого высказывания, лексический подбор, грамматическая и

синтаксическая организация фразы, артикуляция необходимых слов (Седов, Горелов, 2001; Levelt et al., 1999; Dell, 1986; Oppenheim, Dell, 2008, 2010).

Л.С. Выготский (1956) писал о том, что речь происходит от мотива к мысли, потом фаза внутренней речи, после которой следует построение программы речевого высказывания во внутреннем плане, а затем происходит её реализация во внешней речи (произношение). В дальнейшем в отечественной науке развивалась эта идея. Одна из последних отечественных схем, основанных на теории Л.С. Выготского, включает в себя мотив, мысль, смысловое, семантическое и поверхностное синтаксирование и кинетическое программирование (Ахутина, 2022). В схеме Т.В. Ахутиной и А.А. Леонтьева существует два этапа выбора слова во внутреннем плане: по звучанию и по значению. На этапе грамматического структурирования происходит нахождение полных грамматически оформленных слов, которые в последующем перекодируются в кинетическую слоговую программу высказывания. При её реализации происходит артикуляция, а контроль за высказыванием осуществляется системой кинестетического и слухового контроля (Ахутина, 2014). Чистович Л.А. в своих исследованиях отмечала, что «для описания и моделирования процесса управления речеобразованием необходимо использовать такое понятие, как цель движения» (Чистович и соавт., 1976, стр.33), то есть создание определенного акустического эффекта (Чистович и соавт., 1976). Так, стоит упомянуть моторную теорию восприятия речи человеком, согласно которой фонемная классификация происходит благодаря артикуляторным признакам речевых единиц, а акустические компоненты находятся на втором месте (Глухов, 2005)

Также существует горизонтальный уровень, представленный парадигматической и синтагматической категориями. Под парадигматической системой понимается система отношений однородных элементов языка одного уровня и порядка. Эти элементы образуют набор однородных языковых единиц, противопоставленных по определенному ряду признаков – каждая единица способна входить в несколько подобных наборов. Например,

рассматривая уровень звуков, можно отметить, что внутри каждой парадигмы гласных и согласных звуков существуют собственные подпарадигмы. Синтагматическая система, в свою очередь, является системой правил и норм сочетаемости элементов, благодаря которым происходят процессы формирования речевых высказываний. В качестве единицы синтагматической системы выделяют синтагму – сложный языковой знак, состоящий из слов или морфем, находящихся по отношению друг к другу как определяемое к определяющему (Глухов, 2005).

1.2.1. Иерархическая модель В. Левелта

Согласно этой модели, производство слов представляет собой поэтапный процесс (рис.1). На каждом этапе существует своя рабочая единица: лексическое понятие, лемма, морфема, фонологические слова и фонетические оценки – программы артикуляции.

Первый этап *концептуальной подготовки* – выбор лексической структуры слова – исходя из контекста, связан с процессом восприятия. Посредством сбора множества элементов можно сказать, что концептуальная обработка является частью концептуального восприятия. На этом этапе проходит выбор необходимого слова, включая все его возможные проявления в речи. Кроме того, данная система активируется и при слуховом и визуальном входе слова. После концептуальной подготовки происходит лексический выбор – извлечение леммы из ментального лексикона с выбранным выше лексическим значением (Indefrey, Levelt, 2004).

Согласно схеме, на каждую лексическую структуру приходится один лемматический узел. Активная лексическая структура распространяет часть своей активации на "свой" узел леммы, а *выбор леммы (лексический отбор)* – это механизм, который определяет наиболее подходящую лемму. После выбора леммы происходит процесс ее грамматического кодирования и создания соответствующей синтаксической структуры – *доступ к морфо-фонологическому коду*.

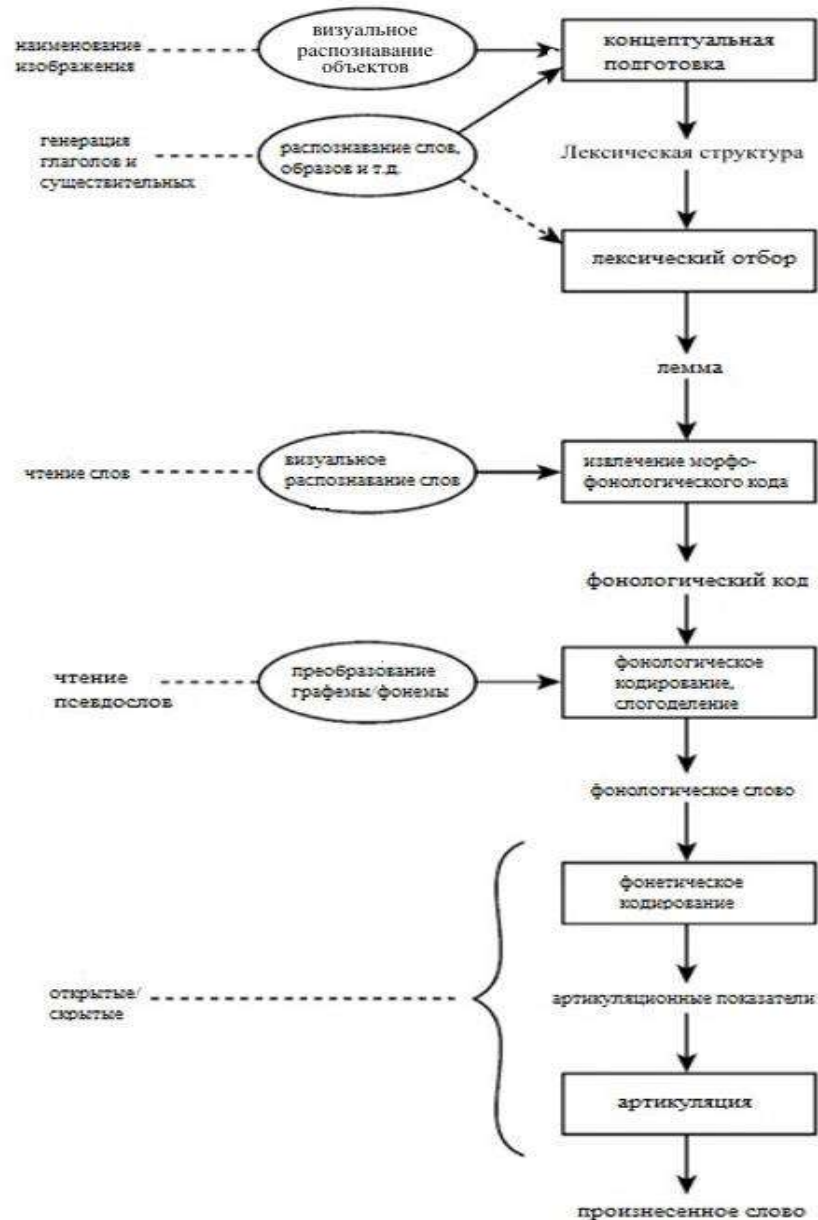


Рис. 1. Модель В. Левелта (пер. Шевченко А.О.) (Indefrey, Levelt, 2004)

Далее следует этап морфологического кодирования. Открывается морфологический состав слова и его форма. При обращении к форме слова говорящий извлекает структурную информацию. В дальнейшем сегменты вставляются в структурный каркас или присоединяются к нему, что обеспечивает их правильное упорядочивание и структуру. После того, как были определены морфологические компоненты слова, начинается этап **фонологического кодирования**, происходит **слоговоеобразование (слоговоеделение)** – процесс, зависящий от фонологического окружения слова. Слоговоеделение выходит за рамки лексического слова и относится к фонологическим рамкам

слова. Область слога деления называется "фонологическим" или "просодическим словом". Фонологическое слово шире рамок лексического слова, так как небольшие функциональные слова без ударения присоединяются к соседним словам содержания, если это позволяет синтаксис, вследствие чего создаются фонологические рамки слова, которые превосходят сохраненные лексические рамки (Levelt, 1999). Этот процесс происходит в режиме реального времени и с учетом положения фонем в слове и взаимодействия слогов между собой (Indefrey, Levelt, 2004).

Эволюционная особенность системы, кодирующей слово в моторную программу, заключается в ее возможности генерировать бесконечное разнообразие взаимно контрастирующих паттернов в артикуляционном и слуховом восприятии. Эта система участвует в фонологическом кодировании. По мере составления фонологических слогов последовательно извлекаются соответствующие артикуляционные характеристики, т.е. происходит *фонетическое кодирование*. В. Левелт (1999) предполагает, что у человека есть доступ к слоговому словарю, в котором хранятся наиболее частотные слоги, а именно их артикуляционные партитуры. Для говорящего функционально выгоднее иметь прямой доступ к часто используемым и, вероятно, внутренне согласованным слоговым партитурам, то есть к хорошо усвоенным артикуляторным движениям. Фонетические формы слов, состоящих из низкочастотных слогов, собираются с использованием сегментной и метрической информации (Levelt, 1999). После фонетического кодирования артикуляционный порядок действий фонологического слова исполняется речевым аппаратом, т.е. происходит *артикуляция* слова.

Также в модели В. Левелта в производстве речи присутствует самоконтроль. Человек осуществляет контроль за своей внешней и внутренней речью, то есть мы можем контролировать некоторые внутренние представления, которые создаются во время кодирования речи. Тем самым В. Левелт (1989) говорит о наличии двух систем самоконтроля: внешней, подразумевающей систему акустического контроля (услышал – понял –

исправил), и внутренней – артикуляционной, когда в начале артикуляции внутренняя система, заметив ошибку, начинает проверять скрытые процессы (Levelt, 1999).

В. Левелт (Levelt, 1999) придерживается точки зрения, что процесс речеобразования происходит последовательно, что было представлено им в модели. Согласно схеме, процесс производства речи составляет намерение, отбор информации, которую необходимо выразить, концептуализацию, а также упорядочивание информации. Данные процессы возможно реализовать при помощи концептуализатора, продуктом которого является доречевое сообщение. В. Левелт описывает концептуальную подготовку как процесс, в котором говорящий генерирует сообщение в одиночку или в интерактивном взаимодействии с целью повлиять на собеседника в соответствии с его намерениями (Levelt, 1999). Для производства сообщения говорящему необходимо иметь доступ к различным типам информации, среди которых процедурное, декларативное, ситуационное знание, а также то, что в ходе взаимодействия сказал сам говорящий и другие говорящие. Конечное сообщение представляет собой концептуальную структуру, состоящую из лексических понятий, то есть понятий, для которых в языке есть слова. Стоит отметить, что не все понятия, которые мы можем представить, являются лексическими (Levelt, 1999). Затем, на этапе грамматического кодирования, формулировщик переводит некоторую концептуальную структуру в некоторую языковую структуру. Лексические понятия в сообщении активируют соответствующие леммы. Лемма включает в себя все, кроме фонологического аспекта слова, то есть концептуальную информацию и морфосинтаксические характеристики. Выбор лемм делает доступными синтаксические рамки, которые должны соответствовать семантическим функциям и аргументам в сообщении. При грамматическом кодировании говорящий использует эту лексико-синтаксическую информацию для построения соответствующей синтаксической схемы, «поверхностной структуры». Так, совершается грамматическое кодирование сообщения, а

после и фонологическое кодирование: говорящий извлекает нужные леммы и располагает их в необходимом порядке. Он получает доступ к морфологическому и фонологическому составу элемента, основному материалу для построения фонологических слов после того, как лемма была выбрана. Извлечение фонологических форм для лемм и построение артикуляторного плана высказывания происходит посредством артикулятора, который выбирает последовательные блоки внутренней речи из артикуляторного буфера и способствует их передаче на исполнение. Выполнение артикуляционной партитуры гортанным и надгортанным аппаратом в конечном итоге дает продукт – внешнюю речь. Важным аспектом модели В. Левелта является самовосприятие: говорящий является своим собственным слушателем. Так, если замечаются проблемы в производимой речи, в частности, проблемы, которые могут иметь коммуникативные последствия, человек может остановиться и исправить себя. В систему понимания речи входит как понимание внешней речи, так и мониторинг, то есть доступ к собственной внутренней речи (Levelt, 1999). Данная система позволяет рассматривать входящую речь в фонологическом, морфологическом, синтаксическом и семантическом аспектах. Так, процесс речепроизводства подразумевает перевод замысла говорящего в речевые единицы конкретного языка (Спиридонов с соавт., 2015).

1.2.2. Сетевая модель Г.С. Делла

Модель Г.С. Делла была построена на основе изучения различных ошибок, встречающихся в речи людей на разных уровнях. Модель включает 4 уровня: семантический, лексический, фонологический и артикуляционный (рис.2).

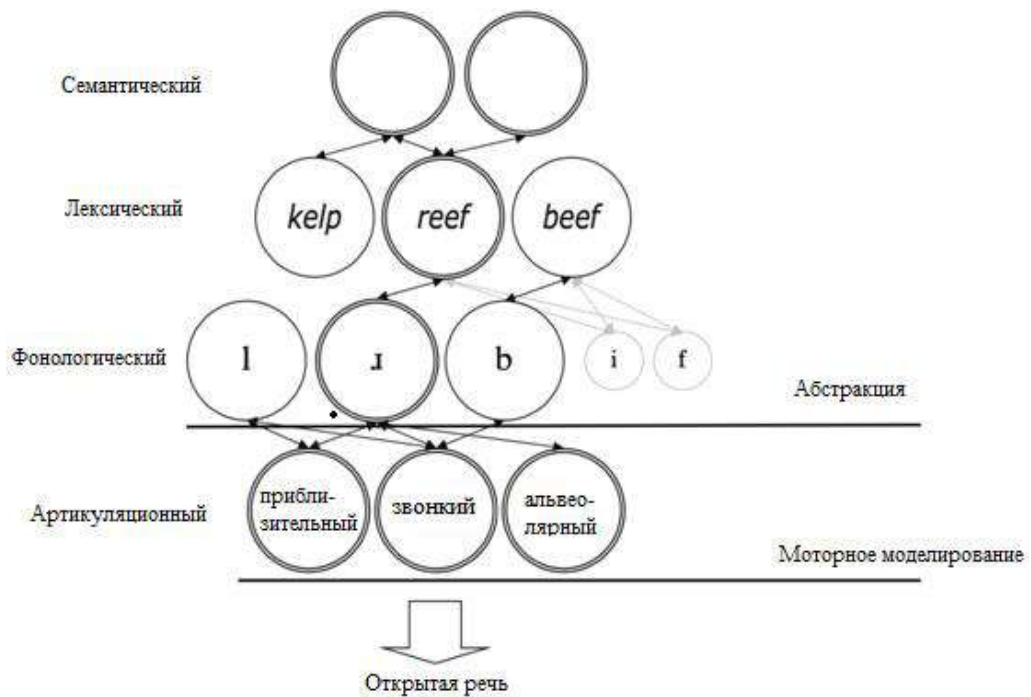


Рис.2. Модель Г.С. Делла (пер. Шевченко А.О.) (Oppenheim, Dell, 2010)

Фонологическое кодирование включает морфемы, слоги, фонемы и признаки (артикуляционные характеристики звуков). Каждый из уровней имеет свои правила отбора необходимых элементов для производства речи. Однако данная модель предполагает сетевое строение, каждый узел представляет отдельные единицы речи и может взаимодействовать с другими на всех уровнях. Так, для произнесения слова активируются все узлы, которые отвечают за семантику, морфемную и фонемную структуру. Эта активность может распространяться на все узлы, пока самый активированный не станет выходной строкой. Из-за этого и возможны различные ошибки во внешней речи (Dell, 1986). **Семантический уровень** – уровень концептуального представления, **лексический уровень** – лексическая сеть, где представлены значения слова, а также морфемы, сегменты и признаки слова. **Фонологический уровень** – уровень фонем. На данном уровне фонемы хранятся в отдельных узлах и связаны одновременно с морфемами, в состав которых они входят, и с отличительными особенностями фонемы в зависимости от её места в слове. Следующий **уровень** – **артикуляционный**, где хранятся моторные особенности произношения звука и его

характеристики: произносится с использованием голоса или нет, положение языка и так далее. Г.М. Оппенгейм и Г.С. Делл (Oppenheim, Dell, 2010) изучили проскальзывания во внутреннем плане с разной степенью артикуляции. В ходе этого исследования авторы пришли к выводу, что артикуляция влияет на внутреннее проговаривание. Это показывает, что внутреннее проговаривание не может быть независимым от артикуляции полностью. Авторы говорят о том, что при внутреннем проговаривании есть только фонологический уровень, на который влияет артикуляция. Каскадная активация означает, что артикуляционные признаки могут быть активированы до выбора фонемы, а наличие обратной связи не позволяет активированным признакам влиять на выбор фонемы. Таким образом, выбранные фонемы могут отражать различное количество активации за пределами уровня фонемы (Oppenheim, Dell, 2010).

1.3. Фонематическая система речи и артикуляция

Фонологическая структура отдельных языков, в том числе и близкородственных, может значительно отличаться друг от друга. Известный филолог В.Я. Плоткин описал различительные артикуляторные параметры для ста различных языков, в том числе для русского (Плоткин, 1993). Переходя непосредственно к процессу речепроизводства, следует отметить, что он представляет собой построение высказывания из различных языковых единиц: предложения состоят из словосочетаний, словосочетания из слов, слова из слогов, а слоги – из фонем. Человеческий мозг оперирует фонемами и слогами, однако важно отметить, что разные языки могут иметь различные кодирующие механизмы (Воробей, 2020).

Фонему принято считать минимальной смысловозначительной единицей языка. Под фонемой понимается абстрактная единица, служащая для формирования языковых знаков: различение единиц лексикона является ее смысловозначительной функцией. Фонема служит строительным

материалом для формирования звуковых оболочек морфем и слов, что выражает ее конституирующую функцию (Касаткин, 2001).

В своих исследованиях, посвященных изучению механизмов речи, Н.И. Жинкин констатирует, что при многократной встрече разных слов с разным звуковым составом, которые дифференцируются на фонемы, формируется решетка фонем. Решетка фонем или фонематическая решетка – системное хранилище элементов, из которых может быть составлено слово с помощью операции отбора. Однако важно учитывать, что для того, чтобы сочетания фонем являлись нормативными для данного языка, надстраивается решетка морфем. Морфемная решетка ограничивает количество звуковых сочетаний при акте отбора, подбирая только нормативные для языка. В свою очередь, сам процесс произнесения слога или слова состоит из переходов от кода речедвижений к фонемной решетке и затем к морфемной решетке (Жинкин, 1958; Морковина и соавт., 2023).

И.А. Бодуэн де Куртенэ установил, что фонема остается целостной и неделимой единицей, в то же время является суммой отдельных антропофонических представлений: физиологических артикуляционных действий (кинем) и акустических результатов артикуляционных действий (акусм). Также ученый ввел термин кинакема, обозначающий единицу двусторонней артикуляционно-перцептивной работы (Бодуэн де Куртенэ, 1963).

Г.А. Климов в своей работе «Фонема и морфема» (Климов, 2009) определил фонему как единицу звукового строя языка, которая служит для опознавания и различения морфем, в чей состав она входит как минимальный сегментный компонент. По мнению Г.А. Климова (2009), фонема – инвариантная единица речи.

В.Я. Плоткин представил отличное от других исследователей мнение в вопросе понимания фонемы. В своих работах он выделил три квантовые фонологические единицы, которые обеспечивают механизм рождения звуковых оболочек для системы знаковых единиц языка – кинакема, фонема и

слог. Кинакема представляет собой языковую единицу, придающую лингвистическую организацию нервно-мышечной деятельности говорящего и соответствующую мышечно-нервной деятельности слушающего (Плоткин, 1993; Вартанов, Швырев, 2016). При этом кинакема – единица субфонемного уровня, неразрывно связанная со звуком. Она служит для его реализации и различения, но в данном случае это является не непосредственно звуком, а квантом деятельности (Плоткин, 1993; Вартанов, Швырев, 2016). Под фонемой исследователем понимается блок кинакем, стандартный квант звука. Звук производится определенной совокупностью артикуляционных действий и воспринимается как совокупность их корреляционных перцептивных детерминант. Кинакема – это кванты звукопроизводства и звуковосприятия. Фонема же, таким образом, представляет собой структуру, состоящую из внутренних квантов, кинакем, она обусловлена бинарностью кинакем. Каждое артикуляционное действие должно иметь определенные оппозиции, или же команды на воздержание от данного артикуляционного действия (Плоткин, 1993; Вартанов, 2008). Второй бинарностью кинакемных единиц является их разделение на константные и вокалические. Исследователь отмечал, что в онтогенезе речи фонема усваивается раньше кинакемы. Распад фонемы на комплекс кинакем происходит при приобретении необходимых артикуляторных навыков в соответствии с образцом. В.Я. Плоткин считал, что, в первую очередь, усваивается фонема, а уже потом происходит процесс ее распада на кинакемы, которые формируются речевым аппаратом путем приобретения необходимых артикуляторных навыков в соответствии с образцом (Плоткин, 1991).

Важной единицей, которая образуется с помощью фонем, является слог – набор фонем; последовательность речевых движений, образующаяся в результате одной управляющей команды (Селиверстов, 1997). В.Я. Плоткин под слогом понимал единицу языкового потока, синтагматически объединенные фонемы. Данные компоненты создают механизм, отвечающий за рождение звуковых оболочек для системы знаковых единиц языка. Слог

является минимальной суперсегментной единицей: один звук слога – слогообразующий, представляющий вершину слога, остальные звуки, соответственно, неслогообразующие. Обязательный слоговой элемент, выраженный гласным звуком, называется ядром или вершиной, часть слога, которая предшествует вершине, есть его инициаль, а часть, стоящая после вершины, – финаль. В соответствии с наличием или отсутствием инициали выделяются прикрытые и неприкрытые слоги, а в соответствии с наличием или отсутствием финали различаются открытые и закрытые слоги (Князев, 2011). Вновь обращаясь к работам В.Я. Плоткина, стоит отметить, что слог является главной единицей фонологического анализа. По его теории, внутреннее строение слога представляет собой набор контрастных различимых звуков.

Отметим, что в основе речи лежат речедвигательный и речеслуховой анализаторы. Полноценный слух, работа слухового анализатора, является необходимым условием для развития речи. На основе хорошей работы слухового анализатора в дальнейшем развивается фонематический слух и фонематическое восприятие.

Речедвигательный анализатор состоит также из трех отделов: периферический отдел, проводниковый отдел, центральный отдел.

Речевой аппарат, относящийся к периферическому отделу речедвигательного анализатора, состоит из трех отделов – артикуляционного, фонаторного и дыхательного, – которые образуют периферический отдел речедвигательного анализатора. При помощи дыхательного аппарата, расположенного в грудной клетке, создается воздушный поток, необходимый для образования звуков. Основным органом фонаторного отдела речевого аппарата является гортань, в которой происходит преобразование воздушного потока. Органы артикуляции располагаются в ротовой полости. Выделяются подвижные органы артикуляции, к которым относятся язык, губы, мягкое небо и нижняя челюсть, и неподвижные – твердое небо и верхняя челюсть. Артикуляция заключается в том, что перечисленные органы образуют щели

или смычки, возникающие при приближении или прикосновении языка к небу, альвеолам, зубам, а также при сжатии губ или прижатии их к зубам (Филичева с соавт., 1989).

В процессе усвоения артикуляционной базы важны кинестезии – ощущения положения и перемещения частей периферического аппарата в пространстве через сигналы проприорецепторов. Речь происходит на выдохе, при фонации голосовые складки сомкнуты (Филичева с соавт., 1989). Смыкания и размыкания голосовых складок продолжаются до тех пор, пока не прекратится давление голосообразующей выдыхательной струи (Филичева с соавт., 1989; Морковина и соавт., 2023). Громкость и отчетливость речевых звуков создаются благодаря резонаторам. При производстве звонких и фрикативных звуков, в том числе глухих, задействуются голосовые складки (гортань), зубы, язык и губы.

Сила выдыхаемого воздушного потока также влияет на тональность произносимого звука. При необходимости произнесения звука с низкой тональностью нужно увеличение объема выдыхаемого воздуха. Для того, чтобы речь человека была членораздельной и понятной, движения речевых органов должны быть физиологичными и точными. Вместе с тем эти движения должны быть автоматизированными, то есть такими, которые осуществлялись бы без специальных произвольных усилий (Морковина и соавт., 2023).

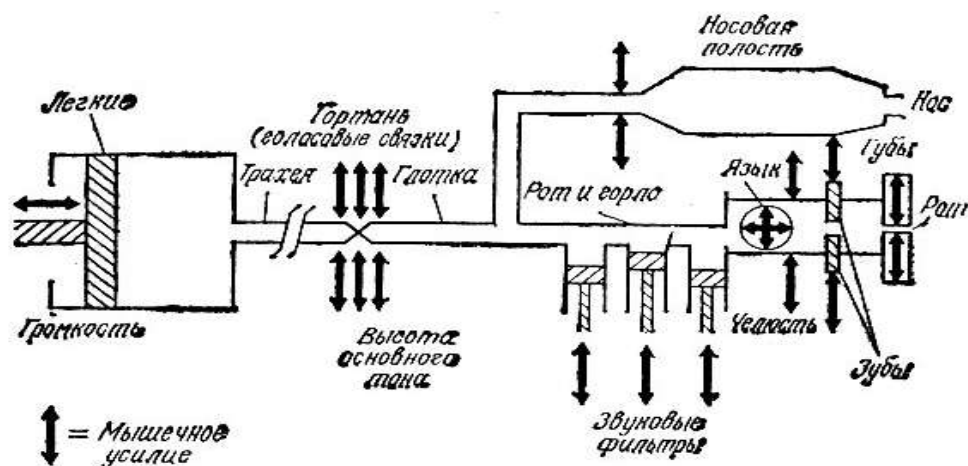


Рис.3. Механическая модель голосового тракта (Кейтер,1985)

На рис. 3 изображена механическая модель голосового тракта (Кейтер,1985). Показан процесс генерации речи, начинающийся с выдоха воздуха из легких, в результате чего создается воздушный поток через голосовой тракт. Согласно схеме, формирование звука происходит либо благодаря колебаниям голосовых связок (мышечное усилие влияет на высоту основного тона), либо с помощью ограничения воздуха языком, губами и зубами. Различные функции голосового тракта связаны с соответствующими речевыми эффектами (Кейтер,1985). Согласно положениям В.Я. Плоткина о том, что фонема состоит из кинаем, которые формируются речевым аппаратом путем приобретения необходимых артикуляторных навыков, важно понимать механизм реализации их в речи с помощью артикуляционного аппарата.

По механизму образования звуки делятся на гласные, характеризующиеся фонацией и свободным проходом воздушного потока через ротовую полость, и согласные, способные реализовываться с участием фонации или без – при прохождении воздуха здесь встречаются преграды (смычки, щели, последовательность смены щели и смычки).

Произнесение фонемы происходит следующим образом: после выдоха воздуха из легких через голосовой тракт образуется воздушный поток, формирующий звучание посредством голосовых органов. При производстве звонких и фрикативных звуков (например, «С») задействуются голосовые складки (гортань), зубы, язык и губы. Громкость речи задается в зависимости от колебаний голосовых складок при определенном количестве выдыхаемого воздуха. Сила выдыхаемого воздушного потока также влияет на тональность произносимого звука: при необходимости произнесения звука с низкой тональностью нужно увеличение количества выдыхаемого воздуха.

Данные характеристики описывают распределение частотных составляющих импульса на выходе из гортани в полосе звуковых частот (максимум распределения смещен в область частот в 200-300 Гц). Частоты представляют собой гармоники, звучащие вследствие колебаний голосовых

складок и резонирующие в носовой и ротовой полостях, а также в горле. При напряжении голосовых складок частота составляющих спектра повышается, что отражается в повышении тональности. Если в звучащей речи содержатся сильные ударения, частота может повышаться до 480 Гц. Форма волны звуковых колебаний и их спектральная характеристика, полученная посредством анализатора спектра, изображены на рис. 4.

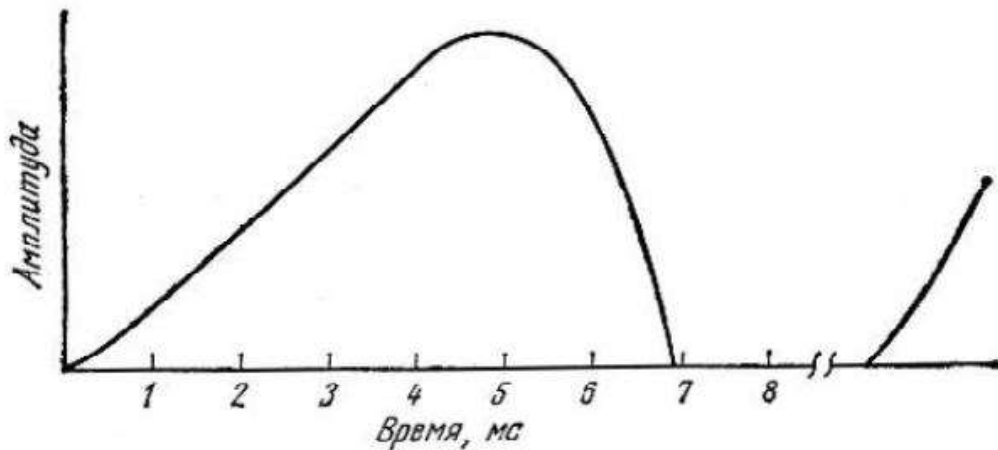


Рис. 4. Форма импульса на выходе из голосовой щели (Кейтер, 1985)

Возникновение глухих звуков, например, «т», «п» и «ш», происходит в результате резонансов и ограничений потоков выдыхаемого воздуха, при участии губ, языка и зубов. При ограничении звука губами возникает звук «ф», зубами и языком — «з». Различия в звучании обосновываются различной силой резонанса, возбуждаемого белым шумом, при ограничении потока воздуха. Источник звука в данном случае находится у выхода из ротовой полости, а следовательно, резонансные полости, находящиеся в гортани, не оказывают значительного влияния на звук, который создается посредством зубов и языка (Кейтер, 1985).

Таким образом, процесс речеобразования имеет сложную физиологическую основу. Речь возникает благодаря совместной работе сложной системы органов.

1.4. Анализ мозговых механизмов речевой деятельности

1.4.1. Конвенциональные речевые зоны

В терминологии, предложенной Ю. Конорски, эту систему называют «центральной двигательной системой поведения» (Конорски, 1970). Определяющим фактором для установления содержания программы двигательной единицы являются особенности синтетической деятельности коры мозга. Определение нейрофизиологической структуры речевой единицы является ключевым моментом в речеобразовании. В своих исследованиях, посвященных двигательному анализатору, А.Р. Лурия уделял особое внимание такой характерной черте импульсов, участвующих в программе двигательного навыка, как «адресованность» (Лурия, 1969). Н.А. Бернштейн также отмечал важность характеристики «целевой точности импульса» и связывал с ней решение двигательной задачи (Бернштейн, 1947). Программа реализации навыка содержит все характеристики, необходимые для его реализации, в том числе и характеристики групп мышц, принимающих участие в данном процессе и последовательность их осуществления. Таким образом, процесс речепроизводства является результатом взаимосвязанной суммарной работы корковых отделов головного мозга и подкорковых двигательных узлов. В речепроизводстве важное место занимает и речевой аппарат, состоящий из тесно связанных регулирующей (центрального речевого аппарата) и исполняющей (периферического речевого аппарата) систем.

Кора головного мозга не является однородной структурой и может быть разделена на множество областей, различающихся на основе их цитоархитектуры. Эти различные области являются функциональными единицами коры головного мозга, так как они получают входные данные от определенных корковых областей и подкорковых структур, вычисляют информацию и отправляют свои выходные данные в области, с которыми они связаны. Это имеет отношение к пониманию нейронной основы языка, потому что вся когнитивная обработка является ничем иным, как сложным

функциональным взаимодействием между сетями корковых и подкорковых областей (Petrides, 2016). Для речепроизводства требуется активация зоны Вернике и зоны Брока, где разрабатывается программа артикуляции, передаваемая затем в моторную кору (Kemmerer, 2015). Зону Брока принято считать центральным отделом речедвигательного анализатора. С точки зрения архитектуры она состоит из двух отдельных областей: области 44 и области 45. Область 44 лежит на *pars opercularis*, а область 45 – на *pars triangularis* нижней лобной извилины.

При восприятии речи звуковые сигналы поступают в первичную слуховую кору и подвергаются в ней обработке (Дружинин, 2003). Центральный отдел речеслухового анализатора – зона Вернике. В зоне Вернике происходит распознавание фонем. Кроме того, важным является развитие фонематического слуха и восприятия. Развитие фонематического восприятия происходит постепенно, параллельно с формированием произношения (Филичева с соавт., 1989). Нарушение в структуре одного ведет к нарушению в работе другого. Критическая роль суперолатеральной области височной доли в доминирующем полушарии для слуховой обработки лингвистической информации была одним из самых ранних концептуальных достижений в формировании понимания языковых областей мозга и заметно фигурировала в теоретических работах Вернике. Исследование поражений, электростимуляция и нейровизуализация подтверждают, что центральная и задняя части верхней височной извилины, верхняя височная борозда и смежная кора средней височной извилины имеют решающее значение для процессов речевой обработки, неотделимых от внутренней речи (Petrides, 2016). В процессе внутренней речи важная роль принадлежит обработке слуховой информации, которая происходит в ряде областей коры в нижней части боковой борозды и прилежащей верхней височной извилине. Первичная слуховая кора, которая находится в центральной части извилины Гешля, имеет типичный кортикальный вид всех первичных сенсорных областей. Эта основная кортикальная область окружена рядом других кортикальных

областей, как в случае ранних стадий визуальной и соматосенсорной кортикальной обработки.

Зоны Брока и Вернике располагаются в непосредственной близости к Сильвиевой борозде. Существует мнение, что область вокруг этой борозды в левом полушарии целиком связана с организацией речевого процесса и не разделяется на отдельные функциональные островки (Ушакова, 2019).

К. Бродман выделил различия между двумя областями прецентральной извилины, областями 4 и 6. Зона 4 занимает большую часть извилины дорсально, близко к средней линии, но отступает и остается ближе к центральной борозде вентрально. Другими словами, в вентральной части прецентральной извилины область 4 в значительной степени ограничена передней частью центральной борозды и занимает только небольшую часть коронки извилины. Область 6 Бродмана охватывает переднюю часть прецентральной извилины (перед областью 4) и проходит впереди верхней прецентральной борозды, покрывая каудальные части верхней и средней лобной извилины, а также каудальную часть медиальной поверхности верхней лобной извилины. Средне-нижняя часть прецентральной извилины связана с артикуляцией. Повреждения в передней части прецентральной извилины (области 6 и 4 Бродмана) усиливают потерю просодии или неправильное соединение фонем, тогда как повреждения в задней части прецентральной извилины (область 4 Бродмана) приводят к искажению фонем. Более того, пациенты с "афазией Брока" всегда имеют вышеуказанные артикуляционные проблемы при нарушениях дорсального потока (Fujii et al., 2016).

Сравнительно недавно был описан лобный кривой пучок (FAT) – внутридолевой тракт, который соединяет верхнюю и нижнюю лобные извилины. В исследовании Э. Чжун с соавт. (Zhong et al., 2022) предположили, что данный тракт будет сильнее связан с двигательными аспектами речи, чем с языковыми. FAT принимает участие во многих речевых процессах и нарушениях, таких как моторные речевые нарушения, заикание, расстройства беглости речи, оперкулярный синдром. Это подтверждает, что FAT связан с

двигательными речевыми аспектами. Кроме того, авторами получены данные о том, что повреждения ФАТ связаны также с апраксией речи (Zhong et al., 2022).

Островковая доля располагается в боковой борозде и связана с лобно-теменной оперкулярной извилиной. Это область коры головного мозга, которая играет важную роль в двигательных аспектах речи. Островковая доля состоит из пяти извилин, которые расположены в форме веера. Вентрально островковая доля сливается с височной долей и роstralно с орбитальной областью лобной доли. Показательно, что участки коры, отвечающие за воспроизведение слов, обладающих различными семантико-лексическими значениями, активируют различные зоны. Например, генерация слов, обозначающих цвет, селективно активирует область в вентральной височной доле сразу перед областью, участвующей в восприятии цвета, тогда как генерация слов действия избирательно активировала область в вентральной височной доле сразу перед областью, участвующей в восприятии движения (Breznitz, 2008).

На исследование речи также повлияло обнаружение «зеркальных нейронов» – данное открытие стало важным этапом в понимании процессов человеческой речи (Arbib, Rizzolatti, 1996). Дж. Риццолатти в 1990-х изучал активность головного мозга макака, выполняющих целенаправленное действие. Удивило исследователей то, что некоторые популяции моторных командных нейронов генерировали разряды тогда, когда животное не действовало само, а наблюдало за таким же действием другой обезьяны или человека. «Зеркальные нейроны» возбуждаются в случае обнаружения индивидом какого-либо действия со стороны других индивидов, а также при воспоминании человеком действия. Подобного рода зеркальные системы присутствуют в различных отделах мозга, в том числе и в зоне Брока. В своем труде М. А. Арбиб и Дж. Риццолатти отмечают, что такие системы участвовали в формировании у индивида поведенческого подражания, что сыграло важную роль в формировании звукового подражания, являющегося

неотъемлемым элементом процесса возникновения речи. «Комплексное подражание» М. А. Арбиб и Дж. Риццолатти считают необходимостью при усвоении слов и артикуляторных движений (Arbib, Rizzolatti, 1996).

Существует подход описания процессов артикуляции при внутреннем проговаривании через упреждающий заряд. Термином «упреждающий» (сопутствующий) разряд обозначается моторная команда, схожая с командой моторного действия, но не приводящая к действию непосредственно. Поведенческой особенностью упреждающего разряда является перцептивный захват – задержка в восприятии. Эксперимент М. Скотта с соавт. (Scott et al., 2013), проведенный с целью доказать предположение, что более активное участие артикуляции должно вызывать упреждающий разряд большей силы, доказал, что более сильное задействование артикуляционного аппарата вызывает более сильный упреждающий разряд.

1.4.2. Таламус и мозжечок как вторичные посредники речевой деятельности

П. Индефрей, В. Левелт (Indefrey, Levelt, 2004) провели большой мета-анализ, где рассматривали зоны, которые также могут быть вовлечены в производство и восприятие речи, кроме основных речевых зон. Так, кроме зон коры больших полушарий в производстве и восприятии речи может участвовать и ряд подкорковых областей, которые были обозначены как вторичные речевые посредники. Однако роль вторичных посредников не определена для внутренней речи.

П. Индефрей, В. Левелт (Indefrey, Levelt, 2004) отмечают, что вентральные билатеральные моторные зоны и сенсорные зоны, правая дорсальная моторная зона, правая дополнительная моторная область (SMA), левый и правый медиальные части мозжечка, билатеральный таламус и средний мозг, вероятнее всего, участвуют в планировании и выполнении артикуляционных движений. Кроме того, левая вентральная прецентральная извилина, двусторонние средние передние извилины, левая веретенообразная

извилины, левый таламус и правый медиальный мозжечок чаще обнаруживались в заданиях с внешней артикуляцией, чем в заданиях с внутренним произнесением (Indefrey, Levelt, 2004).

В своем исследовании А.А. Форд с соавт. (Ford et al., 2013) получили результаты, которые подробно описывают структуры и организации цепи связей "области Брока и базальные ганглии" у человека. В исследовании было показано, что зона Брока взаимодействует с передней скорлупой, выступающей в роли входного ядра в базальные ганглии.

Одна из предполагаемых функций, выполняемая связью между зонами базальных ганглиев и зоны Брока, – усиление активации для наиболее контекстуально подходящих фонем и их соответствующего моторного программирования. Авторы предполагают, что функциональные сети, включающие задние базальные цепи ганглиев зоны Брока, могут быть задействованы для активации соответствующих фонем и их артикуляторных программ для производства выбранного ответа. Две функционально различные области в зоне Брока, передняя часть, вовлеченная в семантический процесс, и задняя часть, вовлеченная в фонологию и синтаксис, проецируются на перекрывающуюся область в передней скорлупе. Сближение этих входов может иметь особое функциональное значение для языка. А.А. Форд с соавт. (Ford et al., 2013) предполагают, что, как только наиболее подходящий семантический ответ был выбран, а потом усилен в цепи базальных ганглиев треугольной части, этот выбор передается в сеть оперкулярной части для усиления активации соответствующего лексико-фонологического представления. Это гарантирует, что желаемый семантический ответ будет сформулирован с использованием соответствующих фонем. Перекрытие этих схем на уровне передней скорлупы и последующих частей контура, включая вентральный передний таламус, может гарантировать, что соответствующие семантические и лексико-фонологические представления точно работают во время выбора слов. Авторы показывают, что входы из двух соседних областей коры (треугольной и оперкулярной частей) сходятся в одной и той же области

передней скорлупы. Исследователи считают, что такое расположение могло бы служить для одновременного усиления соответствующих лексических и семантических представлений во время выбора слова (Ford et al., 2013).

А.А. Богсали с соавт. (Bohsali et al., 2015) в своей статье стараются понять, можно ли определить прямые связи между областью Брока и вентральным передним ядром таламуса и подушкой таламуса (пульвинаром) с помощью диффузионно-взвешенной трактографии. Полученные результаты выявили шесть разных путей, соединяющих область Брока с таламусом. Эти результаты позволяют предположить существование структурного механизма, через который область Брока и таламус могут взаимодействовать. Однако точные функции и роль этих корковых и подкорковых областей остаются неясными. Наиболее распространенным элементом языковых нарушений при поражении вентральной передней или пульвинарной областей является затруднение поиска слов с семантической парафазией. Авторы предположили, что кортикоталамокортикальные связи между треугольной и оперкулярной частями играют роль в сопряжении семантических конструкций с их соответствующими лексическими репрезентациями. Можно предположить, что треугольная и оперкулярная части связаны через кортико-кортикальные пути. Тем не менее, симптомы таламической афазии указывают на то, что таламус играет уникальную роль в интеграции семантических и фонологических представлений во время производства слов (Bohsali et al., 2015).

В настоящее время есть много работ, посвященных изучению роли мозжечка в процессах речи. Существует большое количество исследований (в частности, Starowicz-Filip et al., 2021; Ackermann et al., 2007; Wildgruber et al., 2001), связанных с изучением влияния поражений мозжечка и его полушарий на речь. Так, мозжечковая дизартрия вызывает нарушения не только в моторной сфере, но и в речевой. У людей с мозжечковой дизартрией речь замедленная, толчкообразная, с резким повышением тона к концу фразы – скандированная речь. В речи присутствуют нарушения ритмики (Бабина с

соавт., 2016). А. Старович-Филип с соавт. (Starowicz-Filip et al., 2021) в своем исследовании в ходе анализа отдельных когнитивных областей получали данные о том, что у пациентов с поражением левого полушария мозжечка наблюдается снижение преимущественно внимания и визуально-пространственных функций, а у пациентов с правосторонним поражением наблюдаются более серьезные когнитивные нарушения, в том числе нарушены вербальные способности (Starowicz-Filip et al., 2021). Х. Аккерман с соавт. (Ackermann et al., 2007) в своем метаанализе высказали предположение о том, что правое полушарие мозжечка участвует в процессах временной организации последовательности слогов в привычном темпе во внутренней речи, регулирует длину слогов и участвует в процессах предвосхищения эффектов коартикуляции (взаимовлияния звуков речи друг на друга в речевом потоке). Так, мозжечковые нарушения могут повреждать когнитивные операции, зависящие от внутренней речи (например, повторение) или препятствовать лингвистическому кодированию (Ackermann et al., 2007).

Д. Вильдгрубер с соавт. (Wildgruber et al., 2001) провели исследование с целью определения коррелятов контроля речевой моторики: они использовали фМРТ и внутреннее проговаривание слога «та» с несколькими разными частотами. Был получен результат, что в процессах повторения слогов с определенной частотой во внутреннем проговаривании мозжечковый ответ быстрее в диапазоне 4,0 и 5,5 Гц, а в базальных ганглиях обнаружались высокие показатели активации в диапазоне от 2,5 до 4,0 Гц. Это подтвердило их предположение о разном влиянии различных корковых областей и подкорковых структур на моторный контроль речи (Wildgruber et al., 2001).

1.4.3. Модели мозгового взаимодействия при восприятии и порождении речи

Классическая модель описания речи – это модель двойного потока. Слуховая кора отделяет речевые звуки от неречевых, после чего речевая информация с помощью двух потоков – вентрального и дорсального – поступает в определенные отделы головного мозга. Вентральный (нижний)

поток соединяет звуки со значениями и отвечает за понимание высказываний. Дорсальный (верхний) поток ведет к зонам, где слова превращаются в артикуляторные движения, а именно к нижней теменной и нижней лобной доле (Воробей, 2020). А.Д. Фридеричи и С.М. Гиерхан (Friederici, Gierhan, 2013) делят эти пути еще на два. Так, получается два нижних (вентральных) и два верхних (дорсальных). Таким образом, речевые центры и пути образуют четыре нейронные речевые сети. Две сети занимаются предложениями: одна – словосочетаниями и простыми предложениям, другая – сложными предложениями. Третья сеть обрабатывает значения, а четвертая важна для повторения услышанного (Friederici, Gierhan, 2013).

М. Фудзии с соавт. (Fujii et al., 2016) представил модель (рис. 5), несколько отличающуюся от существовавших ранее. В нее входят зоны Брока и Вернике и дугообразный пучок (AF). В модели представлен дорсальный поток, связанный с фонологической обработкой, а вентральный поток, в свою очередь, связан с семантической обработкой.

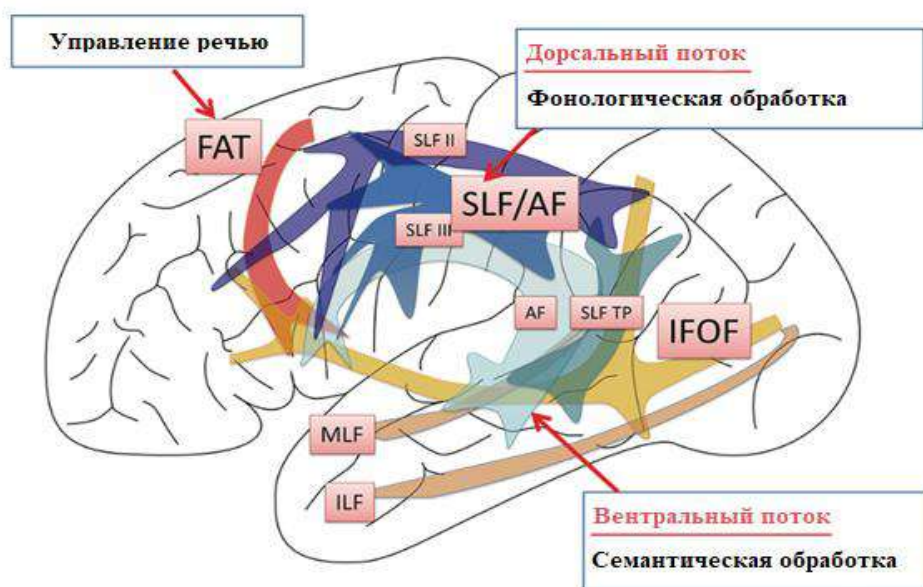


Рис. 5. Модель двойного потока (пер. Шевченко А.О.) (Fujii et al., 2016).

Дорсальный поток связан с фонологической обработкой через SLF. SLF/AF состоит из нескольких субкомпонентов в двух слоях: SLF II, SLF III и SLF TP в поверхностном слое и классический AF в глубоком слое.

Вентральный поток связан с семантической обработкой. Предполагается, что эта сеть состоит из MLF и ILF, а также IFOF. FAT, вероятно, связан с инициацией и спонтанностью речи.

На начальном этапе в модели двойного потока М. Фудзии с соавт. (Fujii et al., 2016) происходит превращение звука в фонему и ее запоминание. Этот этап происходит в первичной слуховой области, которую связывают со слуховой кратковременной памятью. После начинается этап семантической обработки, когда серия фонем, воспринятая ранее, передается в среднюю височную извилину, где получает доступ к значению слова. Задняя часть средней височной извилины, участвующая в доступе к лексикону, связана со слуховым восприятием, выбором слова и фонематической обработки голосового кодирования. Далее происходит более сложный процесс организации синтаксиса предложения, понимания семантики контекста и подбора слов, исходя из конкретного контекста. Важно отметить, что в этот процесс вовлечены в том числе зоны лексического поиска. В данном процессе участвуют многие зоны головного мозга: височная доля (кроме верхней височной извилины), нижняя теменная доля и нижняя лобная извилина. Предполагается, что такая большая сеть поддерживается **нижним лобно-затылочным пучком (IFOF)**, который состоит из двух слоев. Верхний слой, соединяющий верхнюю теменную долю, экстрастриарную затылочную кору, область Вернике и веретенообразную область в затылочно-височном пересечении с **IFG (треугольной и оперкулярной частью)**, которые связаны с речевыми процессами, и глубокий слой, начинающийся от областей, верхнего слоя и заканчивающийся в орбитофронтальной коре, средней лобной извилине и дорсолатеральной префронтальной коре, и, вероятно, связан как с речью, так и с другими когнитивными функциями. **Боковая префронтальная кора** вовлечена в процесс синтаксиса. **Нижнюю теменную долю (IPL)**, угловую извилину связывают с семантическим процессом, а супрамаргинальную извилину (SMG) считают более вовлеченной в фонологический процесс.

Фонологический процесс осуществляется дорсальным потоком, лобная, теменная и височная доли связаны между собой **SLF (верхний продольный**

пучок) /AF (дугообразный пучок). С этим потоком связаны также преобразование звука голоса в фонему, лексический поиск, повторение и артикуляция. Супрамаргинальная извилина и задняя часть верхней височной извилины, связанные с SLF/AF, сопряжены с фонологическим процессом. Реализации речи в дополнительной моторной области (SMA) и средней части фронтальных лобных отделов связаны со спонтанностью и инициацией речи. Авторы уточняют, что описанные этапы могут проходить как последовательно, так и одновременно с динамическими взаимодействиями во многих областях коры (Fujii et al., 2016).

На основе современных исследований локализации поражений и функциональной нейровизуализации понимания и производства речи Дж. Биндер с соавт. (Binder et al., 2015) представили следующую схему:

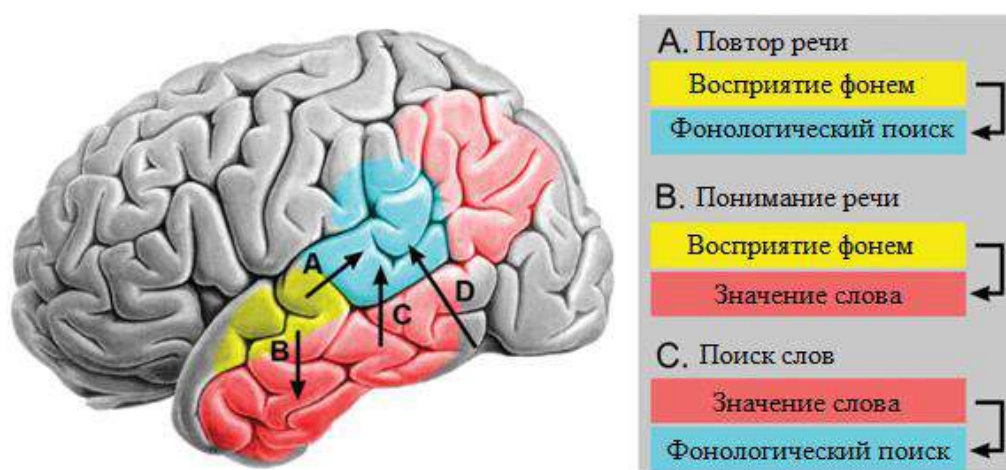


Рис. 6. Функциональная модель основных речевых систем (пер. Шевченко А.О.) (Binder et al., 2015)

Желтым цветом обозначена билатеральная система восприятия фонем речи. **Синим цветом** – область Вернике, поддерживающая преартикуляторный фонологический поиск. **Красным цветом** маркированы височные и теменные компоненты распределенной системы представления значения слова (семантики).

Путь А – для повторения, кроме того, к повторению будут относиться переднетеменные и лобные зоны (не показаны), поддерживающие артикуляторную подготовку и исполнение. **Путь В** – понимание произносимых слов, отражающее воспринимаемые последовательности фонем в смысле слов. **Путь С**, включающий представления концепций на фонологические представления, необходим для производства речи. **Путь D** указывает на прямое сопоставление визуальных форм слов с фонологическими представлениями, что необходимо для чтения вслух.

Область Вернике поддерживает процесс фонологического извлечения, то есть фонемы, которые должны быть артикулированы – их порядок представляется мысленно. Этот процесс необходим для производства речи. При повторении (путь А) работает фонологическая система поиска из слуховой системы восприятия фонем. При чтении вход в фонологическую систему поиска происходит из системы визуального восприятия букв (путь D). Производство речи начинается с этапа извлечения концепции того, что хочет сказать человек, после чего происходит поиск слов с нужным значением и сопоставление этих слов с фонологическими представлениями (путь С). Так, вход данных происходит из внутренней семантической системы, широко распространенной в ассоциативных зонах в височных, теменных и лобных долях (височные и теменные компоненты показаны красным цветом на рис. 6). Понимание речи – это сопоставление последовательности фонем со значениями слов или же работа слуховой системы восприятия фонем и семантической системы (Binder et al., 2015).

Понимание речи – это высоко распределенная функция, включающая в себя биполушарную систему восприятия фонем и широко распространенную семантическую сеть. При задней афазии могут возникать различные нарушения в путях функциональной модели (рис. 7).

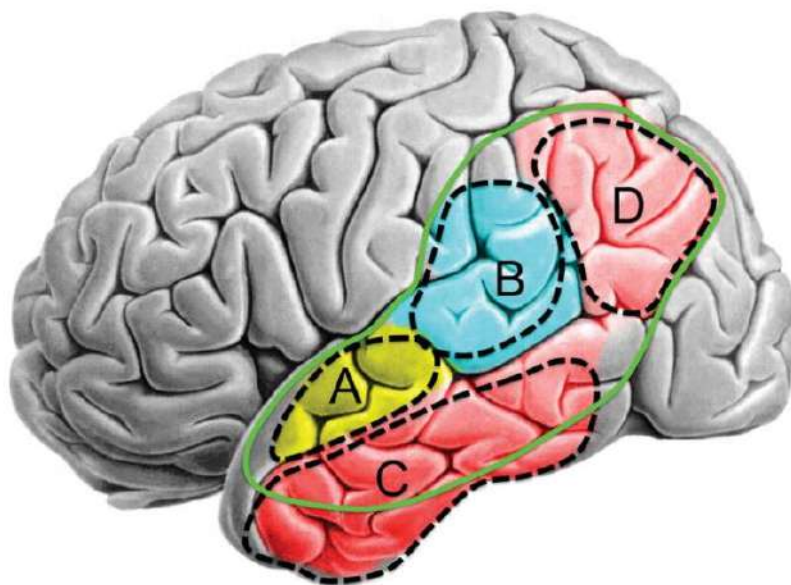


Рис. 7. Корреляты поражения при синдромах задней афазии (Binder et al., 2015). При двустороннем поражении системы *восприятия фонем (А)* происходит словесная глухота. При повреждении системы *фонологического поиска (В)* происходит нарушение поиска слов, фонологические парафазии. *Изолированное повреждение семантической сети (область С или D)* приводит к нарушению понимания речи при сохранной автоматической речевой продукции (Binder et al., 2015).

1.5. Подходы к построению классификаторов внутреннего проговаривания

В современных исследованиях выделяют два вида методов – инвазивные и неинвазивные. Несмотря на заметный прогресс, произошедший в последние годы в использовании неинвазивных методов, необходимо обратить внимание на инвазивные, широко применимые в экспериментах западных и отечественных ученых. Методы инвазивной реконструкции показывают фундаментальную возможность разделить проговариваемые единицы внутренней речи. Недавние исследования (Pasley et al., 2012, Mesgarani et al., 2009, Martin et al., 2014) показали, что при непосредственной регистрации от мозга вживленными электродами (система локальных потенциалов на частоте гамма-ритма) удастся восстановить речевой сигнал и частично дешифровать содержание внутренней речи.

С. Мартин с соавт. (Martin et al., 2014) применяли инвазивные методы исследования для того, чтобы показать, что внутреннее проговаривание может

быть реконструировано на базе моделей, построенных на основе набора данных внешней речи. Для проведения эксперимента использовались внутричерепные записи электрокортикографии у пациентов с эпилепсией, выполнявших задачу чтения вслух или про себя, а также была построена нейронная модель декодирования на высоких гамма-частотах (70-150 Гц) для реконструкции временно-спектральных слуховых особенностей самогенерируемой внешней речи (Martin et al., 2014).

Н. Месгарани с соавт. (Mesgarani et al., 2009) использовали метод реконструкции стимула для изучения того, как сложные звуки кодируются в первичной слуховой коре (ПСК). Этот метод использует линейную спектро-временную модель для отображения ответов нейронной популяции на оценку спектрограммы стимула, что позволяет напрямую сравнивать исходный стимул и его реконструкцию (Mesgarani et al., 2009).

В исследованиях Б.Н. Пасли с соавт. (Pasley et al., 2012) использовались внутричерепные записи из слуховой коры в верхней височной извилине человека, чтобы определить, какая акустическая информация в звуках речи может быть реконструирована из нейронной активности. Декодированные репрезентации речи позволяли считывать и идентифицировать отдельные слова непосредственно по активности мозга во время однократных пробных звуковых предъявлений (Pasley et al., 2012).

Другой инвазивный метод, электрокортикограмма, подразумевает субдуральную имплантацию через небольшие отверстия под твердую оболочку и требует достаточно большой краниотомии. В исследовании М. Ангрик с соавт. (Angrick et al., 2019) был проведен следующий эксперимент: 6 пациентов проговаривали определенные слова, что регистрировалось микрофоном и большим числом сенсоров, имплантированных в кору головного мозга, считывающих мозговую активность. Задачей исследования было превращение данной активности обратно в речь, полагая, что у человека будет наблюдаться одинаковая мозговая активность при внутренней речи. Данная экспериментальная

ситуация схожа с действием моторных интерфейсов: когда человек, например, двигает рукой, исследователь по активности мозга пытается понять, двигает испытуемый правой или левой рукой. Работа описывала считывание активности с большой части коры, в результате чего получался сигнал, который декодировался в мел-спектрограмму, а потом мел-спектрограмма (спектрограмма, в которой частота выражена в мелах) при помощи стандартной нейронной сети Wavenet Vocoder превращала данные в аудиосигнал (Angrick et al., 2019).

В исследовании П. Сан с соавт. (Sun et al., 2020) была создана сеть с большим количеством регуляризации, направленная на декодирование кортикограммы в речь. В сетке выделялось порядка двухсот пятидесяти контактов, однако результаты экспериментов показали, что для декодирования требуется не такое большое число электродов (Sun et al., 2020). Это привело к попытке декодировать речь с помощью небольшого количества электродов в работе А.Т. Петросян с соавт. (Petrosyan et al., 2021): исследование было направлено на декодирование сигналов мозга в отдельные слова (всего – 26 слов). Данное исследование представляло собой попытку декодирования отдельных слов из малого числа электродов. Техническими методами представления речи выступали спектрограмма или мел-спектрограмма. Графические стереосигналы переходили в мел-спектр, а затем мел-спектр преобразовывался в речь. Глобальной целью исследования стало создание минимального инвазивного интерфейса. Было необходимо декодировать дискретные слова, посмотреть синхронные и асинхронные режимы, создать dataset, в котором есть интересные лингвистические признаки, например, обратный порядок слов. Идея, выдвигаемая исследователями, заключалась в существовании некоего нейронального субстрата, неких популяций нейронов, которые являются активными и имеют отношение к артикуляционной функции, а также каким-то образом управляют нашим артикуляционным трактом. Кроме этого, существуют популяции нейронов, которые не имеют отношения к данной выполняемой задаче, и

активность этих популяций регистрируется при помощи сенсоров электродов, в результате чего получается некий вектор в каждом временном интервале N (Petrosyan et al., 2021).

1.5.1. Классификация фонематических стимулов методом ЭЭГ

Практическое применение инвазивной технологии весьма ограничено. В связи с этим возникает задача разработки средств распознавания внутренней речи на основании неинвазивной технологии – электроэнцефалограммы. Если принципиально возможно детектировать внутреннюю речь по электрической активности мозга посредством вживленных электродов, то не исключена возможность успешного ее детектирования и на основе электроэнцефалограммы. Так, практической областью использования нейронауки в изучении внутренней речи стала разработка интерфейсов для распознавания процессов внутреннего проговаривания. Данный пункт был подготовлен на основе совместной статьи Ю. Ю. Гавриленко, Д. Ф. Саада, А. О. Шевченко, Е. А. Ильюшина «Обзор методов распознавания внутреннего проговаривания на основе данных электроэнцефалограммы», в которой был представлен подробный обзор исследований интерфейсов мозг-компьютер.

Данная парадигма исследований показывает, что с помощью прогрессивных математических средств возможно распознавать фонемы благодаря паттернам мозговой активности. Такие исследования предполагают типовой дизайн, который позволяет реконструировать фонемы и слоги, проговариваемые про себя. Анализируя существующие исследования, основанные на распознавании внутреннего проговаривания, можно выделить три типа: распознавание фонем, распознавание слогов и распознавание определенных слов, являющееся наиболее сложной задачей.

Так, Х. Ямагучи с соавт. (Yamaguchi et al., 2015) была создана система распознавания внутреннего проговаривания, основанная на использовании ЭЭГ и ЭМГ (электромиографии). Система включала фазу обучения и фазу

декодирования, представленные в двух экспериментах: в первом были идентифицированы слова из японского эквивалента игры «Камень, ножницы, бумага», во втором слова со значением времен года, которые были преобразованы в последовательность гласных и согласных букв. Точность распознавания составила 65% (Yamaguchi et al., 2015).

Исследование А.Р. Серешке (Sereshkeh et al., 2017) включало разработку двух ИМК: первый был способен различать 10-секундные отрезки внутреннего проговаривания слова «нет» и соответствующие по времени отрезки тишины, а второй возможно 10-секундные отрезки внутреннего проговаривания слова «да» и эквивалентные отрезки с повторением «нет». Каждый из интерфейсов был протестирован с помощью двух онлайн-овых и двух оффлайн-овых сессий на двенадцати испытуемых. Извлечение признаков происходило с помощью фильтра, основанного на быстрой корреляции (FCBF). Точность в работе первого ИМК составила 75.9%, второй ИМК показал точность 69% (Sereshkeh et al., 2017).

Исследование А.К. Порбадник и соавт. (Porbadnigk et al., 2009) ставило целью не только распознавание того, что внутреннее проговаривание может быть распознано, если произносимые слова стоят в блоках. Так, было выявлено существование взаимоотношений между порядком слов и степенью распознавания. Слова, выбранные для мультиклассовой классификации (“go”, “back”, “left”, “right”, “stop”), имели четкое значение, что снижало вероятность разных представлений у испытуемых. Результаты показали, что мультиклассовая классификация не дает результатов, отличных от случайных – точность оставалась около 40%. При бинарной классификации максимальная достигнутая точность составила 87.9% (Porbadnigk et al., 2009).

С. Чиа с соавт. (Chia et al., 2011) исследовали представление фонем: испытуемым было предложено представлять 5 различных фонем, включающих в себя движения челюсти (-aa, ae), языка (-l, -r), небной занавески (-m, -n) и губ (-uu, -ow), а также фрикативные согласные (-s, -z). Применялись методы наивного байесовского классификатора и линейного

дискриминантного анализа. Результаты показали, что сигналы подобного рода могут быть отличимы друг от друга, а для классификации в реальном времени подходят простые линейные классификаторы (Chia et al., 2011).

Те же данные были использованы в работе К. Бригама и Б.В.К.В. Кумара (Brigham, Kumar, 2010), однако в качестве признаков были выявлены коэффициенты авторегрессии, а в качестве алгоритма классификации – метод *k*-ближайших соседей. Целью исследования стало использование представляемой речи для идентификации испытуемого в задаче аутентификации. Так, было выявлено, что использование слогов вместо полных слов устраняет влияние смысла этого слова на мозговые сигналы. Максимальная достигнутая точность работы классификатора составила 44% – метод обработки сигналов показал высокий уровень точности идентификации субъекта, однако точность понизилась при дальнейших сессиях (Brigham, Kumar, 2010).

В статье К.С. Да Салла (Da Salla et al., 2009) был описан эксперимент по распознаванию слогов: во время внутреннего проговаривания представлялась артикуляция при произнесении звуков /a/ и /u/. Так, процесс активации двигательной коры головного мозга в совокупности с воображаемой гласной может быть классифицирован с помощью общих пространственных шаблонов и нелинейного классификатора, однако на точность классификации больше влияют мышечные движения, чем само внутреннее проговаривание. Итоговая точность составила 68-78% (Da Salla et al., 2009).

Внимание исследования М. Мацумото и Дж. Хори (Matsumoto, Hori, 2014) направлено на произнесение звуков /a/, /i/, /u/, /e/, /o/ и на разницу в использовании метода релевантных векторов и метода опорных векторов с Гауссовым ядром. Снижение вычислительной стоимости метода предполагалось за счет использования 19 каналов, общих пространственных фильтров и адаптивного сбора данных. Результатом стала выявленная классификация в 77% и 79%, стоимость реализации первого метода выше и требует больше данных для качественного результата (Matsumoto, Hori, 2014).

В качестве признаков для классификации гласных в исследовании Б. Мин и соав. (Min et al., 2016) были выделены средние значения, дисперсия, стандартное отклонение и асимметрия, а сама классификация основывалась на методе опорных векторов с функцией радиального базиса и с помощью экстремального машинного обучения с разными ядрами. Наибольшая точность классификации была достигнута с помощью метода экстремального машинного обучения – пары гласных /a/ и /i/ были классифицированы с точностью 99.3% (Min et al., 2016).

Нелинейный метод опорных векторов в работе Л. Сармиенто с соавт. (Sarmiento et al., 2014) продемонстрировал точность выполнения 84% и 94%: в исследовании была проведена попытка дифференцирования ментальных состояний при представлении открытых и закрытых гласных без представления артикуляции. Анализ спектра мощности был применен для распознавания шумных сигналов, а возле областей Вернике и Брока был расположен 21 электрод (Sarmiento et al., 2014).

В экспериментах Р. Камалакканнан и соав. (Kamalakkannan et al., 2014) были использованы гласные буквы /a/, /i/, /u/, /e/, /o/. В качестве параметров для извлечения признаков была выделены оценка средних значений, дисперсии и стандартного отклонения и оценка средней мощности. При классификации использовалась нейронная сеть с двумя скрытыми слоями, обученная с помощью метода обратного распространения ошибки. Состоящий из четырех узлов входной слой соответствовал четырем признакам, а состоящий из пяти узлов выходной слой соответствовал пяти распознаваемым гласным. Средняя точность составила 44% (Kamalakkannan et al., 2014).

Несмотря на большое количество проведенных исследований, проблематика разработки интерфейсов безмолвного доступа достаточно обширна. В первую очередь следует отметить выбор числа и места расположения электродов: из-за того, что анатомические особенности строения головы и мозга каждого человека уникальны, расположение электродов на голове одинаковым образом не может гарантировать

одинаковую точность для двух разных людей, а при малом числе электродов на устройстве для регистрации ЭЭГ произведенная классификация будет обладать очень низкой точностью. Важную проблему представляет собой классификация в реальном времени: речевые интерфейсы на основе распознавания внутреннего проговаривания будут удобными в реальной жизни при условии большой скорости классификации, что может быть препятствием при разработке интерфейсов на основе ЭЭГ, поскольку они требуют проведения нескольких попыток для распознавания языковой единицы. Кроме этого, трудности представляют положительные результаты классификации для большего набора звуков, так как все проведенные исследования ограничены очень небольшим числом речевых стимулов. В завершение хочется отметить проблему отсутствия единой точки зрения о конкретной методологии эксперимента, гарантирующей постоянную точность в результатах: на данный момент исследования различаются по виду предъявляемого стимула, по инструкциям для испытуемых, а также по методикам классификации, и каждый из экспериментов обладает своими преимуществами и недостатками.

1.5.2. Миографические подходы к классификации внутреннего проговаривания

Классиком изучения процессов внутренней речи при помощи ЭМГ является А.Н. Соколов (1967). Результаты его исследований показали, что при чтении более сложных грамматических фраз речедвигательные потенциалы проявляются наиболее сильно. Это было ярко продемонстрировано при сравнении чтения на родном языке и на недостаточно изученном иностранном языке: при произнесении русский текст читался без сильных речедвигательных напряжений, в то время как текст иностранного образца вызывал заметное речедвигательное напряжение в органах артикуляции. При анализе речедвигательных напряжений во время слушания было замечено, что

они проявляются при концентрации и фиксации на речи говорящего и при затруднениях в понимании речи другого (Соколов, 1967).

В исследовании Э. Лопес-Ларраз с соавт. (Lopez-Larraz et al., 2010) был представлен интерфейс беззвучной речи на основе электромиографических (ЭМГ) сигналов, регистрируемых в лицевых мышцах. Целью этого интерфейса стало распознавание изолированных прямых слогов испанского языка. Все слоги были разделены на пять групп в соответствии с анатомическим происхождением артикуляции согласного. Сигналы электромиографии отражали электрическую активность мышц во время артикуляции. Электроды ЭМГ были размещены на мышцах только одной стороны лица. Биполярные электроды размещались в направлении волокон лицевой мышцы с периодичностью в один сантиметр. Полученные данные использовались для обучения классификатора для этих слогов. Сигналы от каждого артикулированного слога были преобразованы в вектор признаков, компоненты которого представляли различные глобальные характеристики. Все отобранные признаки были инвариантны к сдвигу времени, чтобы процесс классификации не был чувствителен ко времени произнесения. Полученные результаты показали успешность и эффективность системы при распознавании новых артикулированных слогов (Lopez-Larraz et al., 2010).

М. Хан и М. Джахан (Khan, Jahan, 2018) провели исследование по распознаванию фонем беззвучной речи. Электроды ЭМГ располагались на шее и под подбородком испытуемого. Полученные результаты были проанализированы с помощью трех классификаторов, средняя точность классификации фонем находилась в диапазоне 75-80%. Авторы сделали вывод, что признаки на основе wavelet дают более высокую точность по сравнению с коэффициентами AR. В целом wavelet-преобразование показало большую точность в распознавании и, вероятно, играет важную роль в дифференциации фонемного рисунка беззвучной речи на основе ЭМГ (Khan, Jahan, 2018).

1.6. Постановка исследовательской проблемы, гипотезы и схема исследования

Исходя из теоретического обзора главы 1, пункта 1.1, Внутреннее проговаривание – форма скрытой речевой активности, близкая к внешней речи (Азимов, Щукин, 2009, с.41). Внутренне проговаривание противопоставляется внутренней речи отсутствием свернутости (Азимов, Щукин, 2009, с.41), а также преобладанием знакового компонента (Ерчак, 2017). Многие исследователи отмечают скрытую артикуляцию при внутреннем проговаривании. Так, в исследованиях Г.М. Оппенгейм и Г.С. Делл (Oppenheim, Dell, 2010) получили вывод о том, что внутреннее проговаривание не может быть независимым от артикуляции, которая влияет на фонологический уровень (на котором хранятся фонемы и их отличительные признаки при включении в слог) внутреннего проговаривания. А.Н. Соколов отмечал наличие речедвигательного компонента (зачаточной артикуляции) при внутреннем проговаривании (Соколов, 1984). Также М. Скотт с соавт. (Scott et al., 2013) описывает процесс скрытой артикуляции через упреждающий заряд. Исследование М Скотта и соавт. (Scott et al., 2013) показало, что более сильное задействование артикуляционного аппарата вызывает более сильный упреждающий разряд. При исследовании механизмов внутреннего проговаривания возникает вопрос о специфике активации моторных зон в процессе внутреннего проговаривания, такой вопрос строится на основе исследований А.Н. Соколова, который показал, что артикуляционная активность присутствует при внешнем и внутреннем проговаривании. В процессе внутреннего проговаривания должно происходить или «подавление» конечной артикуляции (которая и отражается в озвучивании), или реализация ее за счет иных мозговых механизмов. Артикуляцию связывают с разными конвенциональными речевыми зонами (глава 1, пункт 1.4.1), а также с мозжечком, таламусом, скорлупой (Indefrey, Levelt, 2004; Ackermann et al., 2007; Ford et al., 2013; Starowicz-Filip et al., 2021) (глава 1, пункт 1.4.2) и взаимодействием этих зон и структур между собой,

однако эти исследования затрагивали внешнюю речь и в меньшей степени внутреннее проговаривание. Кроме того, многие авторы (Indefrey, Levelt, 2004; Marvel, Desmond, 2010; Correia et al., 2020 и др.) связывают мозжечок и подкорковые структуры с артикуляцией во внешней речи, исходя из этого активность мозжечка должна быть выше при внешнем проговаривании, а не при внутреннем проговаривании. С этими положениями связаны две наши первые гипотезы исследования, говорящие о том, что, несмотря на активность конвенциональных речевых зон, при внутреннем проговаривании будет существовать иная система, обеспечивающая обратную связь, а также при сравнении внутреннего и внешнего проговаривания будет присутствовать разница между степенью включенности подкорковых структур и мозжечка.

В исследованиях А.Н. Соколов отмечал наличие артикуляционной реакции при восприятии стимулов и при их внутреннем проговаривании, этот феномен ученый связывал с ориентировочным рефлексом. В его работах резкие амплитудные всплески были выявлены при начальной команде, после чего амплитуды снижались до фонового уровня. В опыте О.П. Капустник дети показывали условную реакцию при произнесении слова «звонок», идентичную той, что была получена после выработки двигательного условного рефлекса на звук звонка (Капустник, 1930). В экспериментах Э.П. Смоленской использовалось вспыхивание красного, белого и желтого света в качестве положительного сигнала и вспыхивание их в обратном порядке в качестве тормозного сигнала: после выработки данных условных реакций испытуемые показывали их при словесных обозначениях раздражителей (Смоленская, 1934). В рамках своего эксперимента Л.И. Котляревский показал, что при выработке условных рефлексов на первосигнальные комплексные раздражители, в состав которых входят компоненты различной интенсивности, вербализуются только сильные компоненты (Котляревский, 1934). Опыты Л.А. Шварц были основаны на замене слов, на которые были выработаны условные рефлексы, синонимами и омонимами: условный фотохимический рефлекс снижения чувствительности

глаз производился как на само слово, так и на его синоним (Шварц, 1948; Соколов, 1967). Так, возникает интерес к изучению условий инициализации, в частности к инициализации условным стимулом. Воспринимаемый стимул обладает рядом характеристик, поэтому возникает интерес к изучению отличий процесса кодирования при внутреннем проговаривании и к тому, в какой момент кодирование заканчивается. Кроме того, интересно проверить, сохраняется ли при ориентировочной реакции информация о воспринимаемом стимуле, так как задающий аудиальный стимул для инициализации проговаривания задает аудиальный образ проговариваемой речевой единицы (включая такие характеристики, как темп, ритм, интонация), а при визуальной или условной подаче стимула аудиальный внутренний образ формируется испытуемым самостоятельно, без влияния аудиальных факторов. Поэтому для определения характеристик кодирования внутреннего проговаривания можно использовать стратегию исследования с помощью инициализации связанного с событием потенциала. С чем и связана третья гипотеза исследования, заключающаяся во влиянии условий инициализации на ранние компоненты связанного с событием потенциала при внутреннем проговаривании.

Также становятся интересными процессы построения высказывания при произношении только фонем или слогов, так как многие схемы речепроизводства на данный момент словоцентричны (глава 1, пункт 1.2). Это исследование в области внутреннего проговаривания фонем и слогов является актуальным, так как внутреннее проговаривание связано с доминированием знакового компонента и в меньшей степени связано с образами (Ерчак, 2017). При этом спецификой внутреннего проговаривания становится и отражение аудиального образа во внутреннем плане. В связи со словоцентричностью большинства моделей невозможно разделить семантическое и несемантическое кодирование речевой деятельности. Так, несемантическое кодирование должно включать артикуляционные паттерны, систему обратной связи, активацию моторных и слуховых зон. Результаты исследования возможно сравнить с исследованиями с семантическим

кодированием для вычленения зон, которые вовлечены только лишь в семантическое кодирование.

При внутреннем проговаривании возникает необходимость формирования образа произносимой единицы на внутреннем плане, что приводит к полимодальности процесса внутреннего проговаривания. Полимодальность – уникальное свойство внутреннего проговаривания, которое вовлекает в процесс проговаривания зоны разных модальностей для построения образа произносимой речевой единицы. Связано это с личным опытом человека, ассоциациями, навыками письма и чтения. Таким образом, система внутреннего проговаривания должна включать не только конвенциональные речевые зоны. Исходя из этого, возникает интерес изучения о различиях в структуре кодирования и в системе иерархических моделей речепроизводства (Levelt, 1999; Dell, 1986; Indefrey, Levelt, 2004; Oppenheim, Dell, 2010) (глава 1, пункт 1.2).

В настоящее время существует ряд иностранных исследований (Chia et al., 2011; Da Salla et al., 2009; Matsumoto, Hori, 2014 и др. (глава 1, пункт 1.5.1)), которые с помощью неинвазивных методов предпринимали попытки построить классификаторы фонем или слогов на примере других языков. В данной диссертационной работе возможность построения классификатора фонем и слогов русского языка не только является актуальностью работы, но и позволяет доказать, что респондент действительно внутренне проговаривает заданный стимул, что говорит о релевантности исследования, а также дает возможность для развития технологий нейроинтерфейсов для людей, у которых нет возможности использовать внешнюю (озвученную) речь. Произнесение лишь фонем также представлено в речевой деятельности человека посредством ввода несемантических единиц, которые объединяются в слова и предложения, но по отдельности не являются смысловыми. Например, в ситуациях проговаривания по буквам для записи фамилий, при наборе текста, в работе с интерфейсами «BCI» и пр. Таким образом, четвертая гипотеза исследования заключается в возможностях классификации

отдельной фонемы или отдельного слога благодаря наличию уникального пространственно-временного паттерна связанного с событием потенциала.

Схема исследования. Для ЭЭГ-исследования был выбран классический ход эксперимента, связанный с классификацией внутренне проговариваемых единиц. Для точности эксперимента испытуемые должны быть сосредоточены на задании и не должны испытывать дискомфорт – помещение изолируется от громких звуков. Могут предъявляться визуальные стимулы, представляющие собой сообщение-инструкцию на экране, или звуковые. В качестве результата выступает получение набора сигналов. Плюсом такой схемы эксперимента является дополнительная валидизация наличия внутреннего проговаривания, так как использование прогрессивных статистических моделей позволяет предугадать проговариваемые единицы (с определенной долей вероятности). Такой дизайн эксперимента позволяет провести исследования как с помощью ЭЭГ, так и с помощью метода фМРТ. При этом полученные данные могут лечь в основу как описания механизмов внутреннего проговаривания, так и оптимизации и теоретического описания парадигмы «классификаторов» и ВСІ.

Обоснования выбора стимулов. В большинстве исследований по реконструкции речи стимульным материалом выступают фонемы, потому что они позволяют универсализировать иные единицы речевого ввода. Чаще всего в подобных исследованиях выбирают небольшое количество фонем и слогов (от 1-4 фонем и слогов) (Da Salla et al., 2009; Matsumoto, Hori, 2014; Min, Kim, Park, Lee, 2016). Кроме того, метод связанных с событием потенциалов также ограничивает нас в количестве потенциально возможных стимулов в связи с частыми повторениями и длительностью экспериментов, что также является ограничением при подборе стимульного материала для исследования. В качестве стимульного материала были отобраны 7 фонем русского языка: две фонемы гласные (А – [а], У – [у]), и пять согласных (Б – [б], Ф – [ф], Г – [г], М – [м], Р – [р]). Фонемы и слоги были выбраны как единицы, не имеющие значения, таким образом, исключается семантическое кодирование для

исследования лишь процессов внутреннего проговаривания. Семантическое кодирование вовлекло бы во многом персональный опыт, эмоциональное содержание речи и ассоциации.

Выбранные фонемы имеют ряд различительных признаков, которые могут помочь при декодировании. Согласно артикуляционной классификации, выделяют следующие признаки гласных звуков: ряд – горизонтальное положение языка в ротовой полости при произнесении звука, степень подъема языка к небу; подъем – вертикальное положение языка в ротовой полости при произнесении звука, места подъема языка в ротовой полости; лабиализация – участие губ при произнесении звука. Для согласных учитываются следующие признаки (Филичева и др. 1989, Колесникова и др., 2023):

1. Место образования – место, где активные и пассивные органы артикуляции создают преграду для получения шума. Учитывается активный и пассивный орган артикуляции.

2. Способ образования – способ преодоления преграды, которая образовалась органами артикуляции. Смычка – преграда, полностью перекрывающая проход воздушной струи. Щель – ограничивает проход воздушной струи.

3. По наличию вибрации голосовых складок выделяют звонкие – сочетание голоса и шума, вибрируют голосовые складки; сонорные – преобладание голоса над шумом, вибрируют голосовые складки; глухие – в образовании участвует только шум, голосовые складки не вибрируют.

4. По месту резонирования – место, где происходит основное резонирование звука.

В работе не учитывается признак подъема спинки языка, так как мягкие фонемы не брались.

Таблица 1. Характеристики выбранных фонем

Фонема	Артикуляционная классификация (Филичева и др., 1989; Колесникова и др. 2023)	Классификация В. Я. Плоткина (1993)
А – [а]	Средний ряд Нижний подъем Нелабиализованная	Вокалическая подсистема Ширококорастворность
У – [у]	Задний ряд Высокий подъем лабиализованный	Вокалическая подсистема Огубленность Узкокорастворность
Б – [б]	1. Губно-губной 2. Взрывной 3. Звонкий 4. Ротовой	Консонантная подсистема Преградная категория Смычная оппозиция Активная предцентральность Шумность
Ф – [ф]	1. Губно-зубные 2. Щелевые 3. Глухой 4. Ротовой	Консонантная подсистема Активная предцентральность
Г – [г]	1. Заднеязычный 2. Смычный 3. Звонкий 4. Ротовой	Консонантная подсистема Преградная категория Смычная оппозиция Активная зацентральность Шумность
М – [м]	1. Губно-губной 2. Смычно-проходной 3. Сонорный 4. Носовой	Консонантная подсистема Преградная категория Проточная оппозиция Активная предцентральность Сонорная оппозиция
Р – [р]	1. Переднеязычный (язычно-альвеолярный) 2. Дрожащий 3. Сонорный 4. Ротовой	Консонантная подсистема Фональная категория Сонорная оппозиция

Слоги составлялись из выбранных фонем. Всего десять слогов: БА – [ба], ФА – [фа], ГА – [га], МА – [ма], РА – [ра], БУ – [бу], РУ – [ру], МУ – [му], ФУ – [фу], ГУ – [гу]. Были выбраны прикрытые слог, т.е. начинающиеся с согласного. В процессе артикулирования первой группы слогов (слоги с [А]) артикуляционные движения для произнесения гласного начинаются на

момент произнесения согласного и не влияют на него в отличие от второй группы слогов (слоги с [У]). Влияние фонемы [У] на согласные приводит к огублению при произнесении. Таким образом, вводится оппозиция в слоги, которая может быть полезна при построении классификаторов. Также при восприятии таких слогов для индификации согласного происходит анализ всего слога (Ганиев, 2012). Кроме того, все выбранные фонемы относятся к разным группам согласно артикуляционной классификации и классификации В.Я. Плоткина. Такой подход не только показывает активность одной фонемы, но и позволяет в процессе усреднения ССП этих фонем говорить об общих тенденциях для всех форм артикуляционных паттернов фонем русского языка, а следовательно, позволяет строить усредненную модель внутреннего проговаривания при подобном рода дизайне эксперимента (ССП).

ГЛАВА 2. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ

2.1. Методика исследования внутреннего проговаривания методом электромиографии

Испытуемые. В исследовании принимали участие 7 испытуемых: 4 женского пола и 3 мужского пола в возрасте от 18 до 23 лет (средний возраст – 19,6 лет). Запуск внутреннего проговаривания происходил на основе аудиального стимула (фонемы).

Оборудование. В качестве программы для записи и просмотра ЭМГ, а также для предварительного удаления артефактов вызванных потенциалов использовалась Brain Sys (Brain Win). Регистрация электрической активности мозга проводилась монополярно, системы «Нейро-КМ» (компания «Статокин», Россия). Для предъявления стимулов использовалась программа Presentation (версия 18.0 фирмы Neurobehavioral Systems, Inc.).

Процедура. Электроды для записи располагались в двух местах над верхней губой слева и на гортани. Схема расположения электродов представлена в Приложении 1 (схема 1). Длительность исследования 15-20 минут, исследование состоит из двух серий: первая серия – внешнее проговаривание; вторая серия – внутреннее проговаривание. В начале серии исследования участникам предлагалась инструкция: «В данном задании Вам будут предъявляться фонемы. Они будут демонстрироваться в случайном порядке через наушники. Постарайтесь не совершать резких движений, смотрите только на монитор и постарайтесь не отвлекаться от предъявляемых стимулов. После специального звукового сигнала и появления фиксационного креста на экране компьютера Вам необходимо проговорить заданную фонему вслух/про себя (для второй серии) до исчезновения фиксационного креста. Проговаривая фонему постарайтесь сохранить заданную длительность и интонацию». Стимулы предъявлялись в случайном порядке, 70 повторений на каждый стимул. Исследование соответствует стандартным схемам исследования, описанных в схожих статьях (рис. 8):

1. Аудиальная подача стимула – фонемы (700 мс).
2. Пауза в 500 мс.
3. Представление фиксационного креста и звукового сигнала, который являлся стартовой командой внутреннего проговаривания или внешнего проговаривания (озвучивания) (250 мс).
4. Внешнее или внутреннее проговаривание заданной случайной фонемы (1500 мс).
5. Отдых (1600 мс).

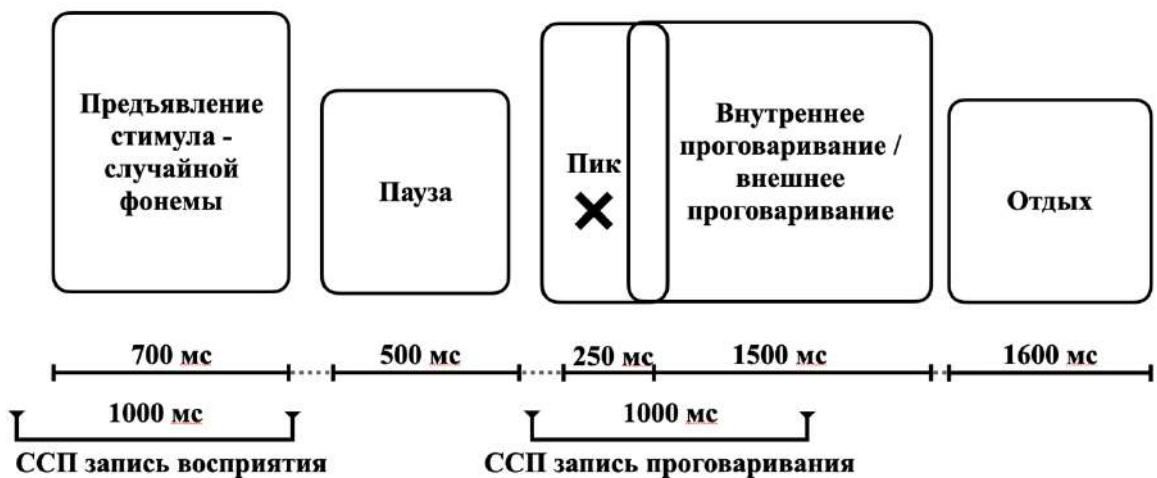


Рис. 8. Схема исследования внутреннего проговаривания методом электромиографии

Синхронизация СПП внутреннего проговаривания. Задавалась посредством предъявления звукового сигнала и фиксационного креста.

Стимулы. Предъявлялись следующие стимулы: 7 фонем (А – [а], Б – [б], Ф – [ф], Г – [г], М – [м], Р – [р], У – [у]).

2.2. Результаты исследования внутреннего проговаривания методом электромиографии

Для апробации работы было проведено исследование в парадигме исследований А.Н. Соколова с целью выявления артикуляционной активности при внутреннем проговаривании с помощью метода ЭМГ. Наличие артикуляционной активности подтверждает наличие внутреннего проговаривания в дизайне эксперимента.

В результате проведенного исследования были построены усредненные потенциалы по всем испытуемым в процессе внутреннего проговаривания фонем, внешнего проговаривания (озвучивания) и восприятия фонем без их повторения. Для усреднения и построения графиков использовалась программа Statistica 8, длительность усредненного потенциала составляла 200 мс. до стимула и 800 мс. после стимула. Ориентировочная реакция начинается от 0 мс. и представлена в смещении пиков. Внешнее проговаривание фонем обусловлено резким сдвигом электрической активности от 200 мс. до 700 мс., при этом в отличие от внутреннего проговаривания представлена обратная форма ЭМГ ССП на двух электродах (рис. 9). Так, EMG 2 (черная сплошная линия) представляет плавную негативную активацию с компонента n200.

Анализируя усредненные показатели проговаривания про себя фонем, среди компонентов можно выделить негативную волну на 200 мс, которая смещается в позитивный компонент в интервале 300-400 мс. Анализ компонентов артикуляционного ответа при восприятии фонем (в сравнении с проговариванием) показывает сдвиг на более ранние компоненты относительно процесса внутреннего проговаривания и повторяют волну, но уже от 150 мс (Шевченко, Вартанов, 2023).

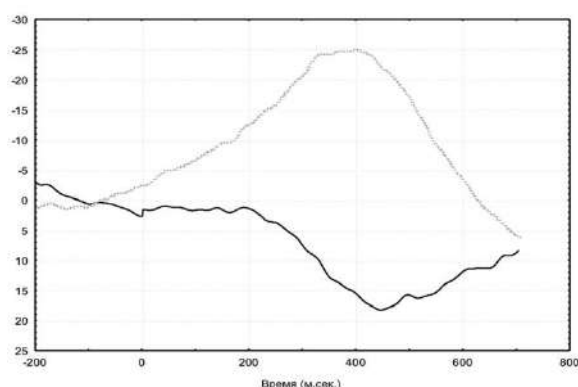


Рис. 9. Усредненный моторный ответ 7 участников во время произнесения фонем вслух; EMG 1 – серая пунктирная линия, EMG 2 – черная сплошная линия (Шевченко, Вартанов, 2023)

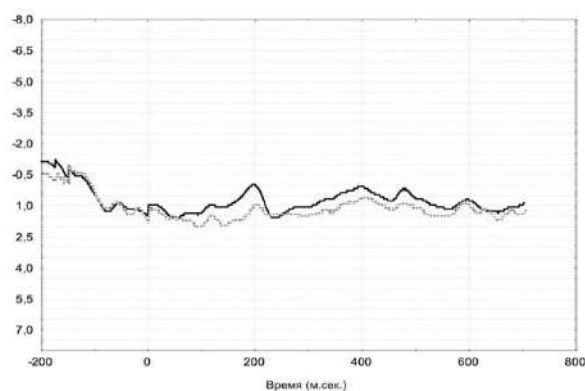


Рис. 10. Усредненный моторный ответ 7 участников во время произнесения фонем про себя; EMG 1 – серая пунктирная линия, EMG 2 – черная сплошная линия (Шевченко, Вартанов, 2023)

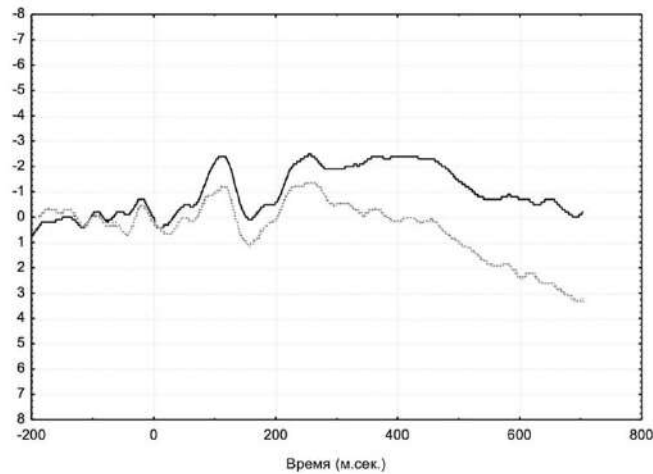


Рис. 11. Усредненный моторный ответ 7 участников во время восприятия фонем; EMG 1 – серая пунктирная линия, EMG 2 – черная сплошная линия (Шевченко, Варганов, 2023)

2.3. Методика исследования внутреннего проговаривания методом электроэнцефалографии

Исследование внутреннего проговаривания на аудиальную и визуальную инициализацию

Испытуемые. В исследовании приняли участие 25 испытуемых: 15 испытуемых женского пола и 10 испытуемых мужского пола в возрасте от 18 до 28 лет (средний возраст – 20,8 лет). Образование испытуемых – неоконченное высшее или высшее. Все испытуемые не имели в анамнезе черепно-мозговых травм, неврологических и психоневрологических заболеваний. У всех участников ведущая рука правая.

Оборудование. В качестве программы для записи и просмотра ЭЭГ, а также для предварительного удаления артефактов вызванных потенциалов использовалась – Brain Sys (BrainWin). Регистрация электрической активности мозга проводилась монополярно с помощью 19-канального электроэнцефалографа «Нейро-КМ» (компания «Статокин», Россия). Electrodes were placed according to the international 10-20% system with two mastoids. For stimulus presentation, the Presentation program (version 18.0, Neurobehavioral Systems, Inc.) was used.

Процедура. После стимула, задающего фонему, предъявлялся специальный звуковой стимул, обозначающий начало проговаривания, а

также появлялся фиксационный крест. Звуковые стимулы – заранее записанные на диктофон фонемы – предъявлялись через наушники. Стимулы предъявлялись в случайном порядке, 70 повторений на каждый стимул. Были проведены следующие серии эксперимента, перед каждой предлагалась инструкция:

1. Контрольная серия – испытуемому демонстрировались фонемы и слоги в случайном порядке зрительно и аудиально, задачи повторить не было. Инструкция: «В данном задании Вам будут демонстрироваться случайные фонемы или слоги. Постарайтесь не совершать резких движений, смотрите только на монитор и постарайтесь не отвлекаться от предъявляемых стимулов».

2. Эксперимент с инициализацией внутреннего проговаривания на основе визуальных стимулов: фонем и слогов, которые необходимо было прочитать, а после повторить. Данная серия проводится с открытыми глазами. Инструкция: «На экране будет появляться стимул: фонема или слог. Постарайтесь не совершать резких движений, смотрите только на монитор. Вам необходимо прочитать про себя заданный стимул. После появления фиксационного креста и звукового сигнала необходимо внутренне проговорить представленный ранее стимул. Постарайтесь сделать это до исчезновения фиксационного креста».

3. Эксперимент с инициализацией внутреннего проговаривания на основе аудиального стимула (фонемы и слоги). В этой серии стимул задается звуковой записью, участнику его надо только повторить. Испытуемый проговаривает с открытыми глазами. Инструкция: «В наушниках будет подаваться стимул случайные фонема или слог. После звукового сигнала и появления фиксационного креста Вам необходимо его внутренне проговорить. Постарайтесь не совершать резких движений, смотрите только на монитор».

Каждая серия происходила согласно стандартной схеме эксперимента, описанной в схожих исследованиях (рис. 12):

1. В течение 700 мс испытуемому демонстрируется визуальный либо звуковой стимул, который является инструкцией к внутреннему проговариванию (определяет, какую именно фонему нужно будет проговорить про себя), данный этап задает образ на дальнейшее проговаривание (инициализирует проговаривание).
2. После предъявления стимула следует пауза длительностью 500 мс.
3. После паузы на экране появляется фиксационный крест и звуковой сигнал в виде пика (250 мс). На фиксационном кресте должен фокусировать свой взгляд испытуемый в процессе проговаривания, и это также является стартовой командой внутреннего проговаривания соответствующей фонемы, которое должно завершиться в течение времени, пока он отображается на экране (1500 мс)
4. Фиксационный крест перестает отображаться на экране и наступает отдых (1600 мс), в течение которого испытуемый не выполняет внутреннее проговаривание и готовится к следующему стимулу.

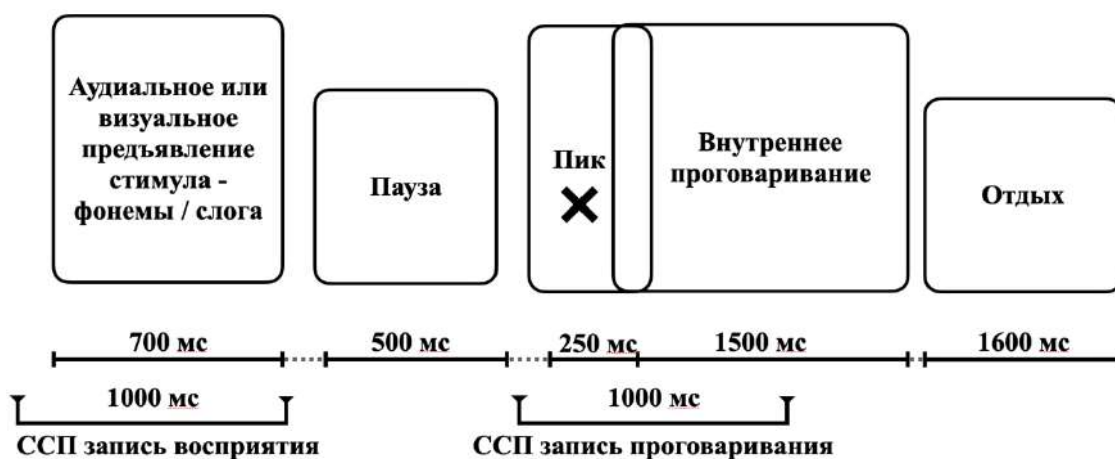


Рис. 12. Схема экспериментальных серий исследования внутреннего проговаривания методом электроэнцефалографии

Общее время исследования от 80 до 90 минут. Каждая серия длилась по 12 минут, между каждой серией испытуемому давалась возможность отдохнуть, выпить воды.

Синхронизация СПП внутреннего проговаривания. Задавалась посредством предъявления звукового сигнала и фиксационного креста.

Стимулы. В соответствии с главой 1, подпункта 1.6 в качестве предъявляемых стимулов были выбраны семь фонем русского языка (А – [а], Б – [б], Ф – [ф], Г – [г], М – [м], Р – [р], У – [у]) и десять слогов, составленных из этих фонем (БА – [ба], ФА – [фа], ГА – [га], МА – [ма], РА – [ра], БУ – [бу], РУ – [ру], МУ – [му], ФУ – [фу], ГУ – [гу]). Стимулы предъявлялись в случайном порядке.

Анализ. Анализ полученных данных проводился с помощью языка статистического программирования R (версия 3.6.3) (RCoreTeam.2019) в рамках интегрированной среды разработки RStudio (версия 1.2.5003) (RStudioTeam 2019). В рамках данной работы использовались пакеты `data.table` для эффективной работы с данными (Dowle, 2019) и `ERP` для статистического анализа связанных с событием потенциалов (Causeur, 2020). Основные два метода, которые использовались в рамках анализа, – это функциональный дисперсионный анализ и попарные сравнения с контролем ошибки I рода (поправки Бэнджамини-Хохберга (Benjamini et al., 1995)). Первый метод имеет такую же интерпретацию и использование, как и «классический» дисперсионный анализ – его основное отличие в том, что он оценивает различие не средних, а кривых. По этой причине функциональный дисперсионный анализ прекрасно подходит для вызванных потенциалов (любой ВП является кривой) и был адаптирован с целью обнаружения значимых различий в пакете `ERP` (Causeur, 2019). Хотя функциональный дисперсионный анализ может обнаруживать сам факт значимого различия, он не способен определять его точную локализацию. В пакете `ERP` реализовано несколько процедур распознавания сигнала, однако ввиду сравнительно малого размера выборки многие из них могут быть излишне консервативны. По этой причине в данной работе используется классический метод попарных сравнений ВП – оценивается значимость различия между каждой совпадающей точкой в кривых ВП. Такой подход связан с большими рисками

ошибки I рода из-за множественных сравнений, поэтому получившиеся р-значения были скорректированы с помощью поправки Бэнджамини-Хохберга, используемой в пакете ERP по умолчанию (Benjamini et al., 1995). Кроме того, дизайн эксперимента является внутригрупповым, а этапы эксперимента являются зависимыми друг от друга, поэтому была учтена вариабельность внутри результатов испытуемых. Пакет ERP не имеет отдельного функционала для дисперсионного анализа с повторными измерениями, поэтому авторы пакета рекомендуют кодирование переменной испытуемых контрастами, сумма которых дает 0. Именно этот подход и был реализован в работе (Sarmiento et al., 2014).

Исследование инициализации условным стимулом внутреннего проговаривания

Многие исследователи (Капустник, 1930; Смоленская, 1934; Котляревский, 1934; Шварц, 1948) занимались изучением скрытых речевых реакций. К попыткам обнаружения скрытых речевых реакций с помощью условно-рефлекторной методики обращался А.Н. Соколов, который в работе «Внутренняя речь и мышление» описал ряд экспериментов на условный импульс (Соколов, 1967).

Испытуемые. В исследовании приняли участие 12 испытуемых в возрасте от 17 до 28 лет (средний возраст 19 лет). Образование испытуемых – неоконченное высшее или высшее. Все испытуемые не имели в анамнезе черепно-мозговых травм, неврологических и психоневрологических заболеваний. У всех участников ведущая рука правая.

Процедура. Предварительно испытуемый заучивал ассоциации японского слова (японским языком никто не владел) с определенной фонемой, в данном случае японские слова выступали квазисловами. Стимулы участникам предъявлялись в случайном порядке. Начало проговаривания определялось специальным стимулом, следующим после первого стимула, задающего проговариваемую фонему. Для предъявления звукового стимула на испытуемых были надеты наушники. Данное исследование состояло из одной

серии с инициацией внутреннего проговаривания на основе условного стимула, предъявляемого аудиально, длительность 12 минут. Участникам предлагалась следующая инструкция: «Перед проведением исследования Вам предлагалось выучить квазислова, связанные с определенной фонемой. В этом исследовании Вам будут предъявляться квазислова. Они будут демонстрироваться в случайном порядке через наушники. После специального звукового сигнала и появления фиксационного креста на экране компьютера Вам необходимо проговорить ассоциированную со стимулом фонему про себя до исчезновения фиксационного креста. Постарайтесь не совершать резких движений, смотрите только на монитор и постарайтесь не отвлекаться от предъявляемых стимулов».

Схема эксперимента соответствует схеме исследования внутреннего проговаривания на аудиальную и визуальную инициализацию (рис.12).

Стимулы. В качестве проговариваемых стимулов были выбраны семь фонем русского языка: А – [а], Б – [б], Ф – [ф], Г – [г], М – [м], Р – [р], У – [у]. В качестве условного стимула были выбраны японские слова, респонденты японского не знали, поэтому эти слова выступали квазисловами.

Анализ. Анализ полученных данных так же, как в предыдущем исследовании, проводился с помощью языка статистического программирования R (версия 3.6.3) (RCoreTeam.2019) в рамках интегрированной среды разработки RStudio (версия 1.2.5003) (RStudioTeam 2019). Также используется классический метод попарных сравнений ВП – оценивается значимость различия между каждой совпадающей точкой в кривых ВП.

2.4. Результаты исследования внутреннего проговаривания методом электроэнцефалографии

2.4.1. Результаты исследования внутреннего проговаривания методом электроэнцефалографии на визуальную и аудиальную инициализацию

При сравнении проговаривания на всех экспериментальных этапах с контрольной частью функциональный дисперсионный анализ показал значимые различия во всех четырех случаях (визуальное предъявление, аудиальное предъявление, фонемы, слоги), которые были необъяснимы вариабельностью внутри испытуемых ($p < 0.001$). Проверка эффекта взаимодействия переменной канала и переменной этапа показала значимые различия на уровне 0.05 только для аудиальных фонем и слогов ($p = 0.02$ и $p = 0.005$, соответственно). При визуальной подаче фонем и слогов взаимодействие является статистически незначимым ($p = 0.28$ и $p = 0.22$, соответственно). Это означает, что пространственное распределение СПП по каналам изменяется относительно контрольного этапа только в случае аудиальной подачи фонем/слогов; при визуальной подаче это распределение остается относительно константным. Эффект канала – характеристики СПП различаются между каналами. Эффект этапа – характеристики СПП различаются между условиями эксперимента. Таким образом, если нет взаимодействия, то эффект канала и эффект этапа меняются независимо друг от друга. При этом между каналами амплитуда СПП может различаться, но характер этих различий одинаковый вне зависимости от условия. Если взаимодействие есть, то какие-то характеристики каналов меняются в зависимости от условия. Что касается локализации различий, то наиболее значимые на уровне 0.05 различия обнаруживаются на каналах С3 и F3 (для аудиальной подачи различия также видны на канале F7). Они характеризуются позитивной волной в районе 120-170 мс и негативной волной (на С3 и F3) в районе 200-280 мс. При этом позитивная волна на С3 и F3 является статистически значимой вне зависимости от модальности стимула, а

негативная волна статистически значима только для аудиальных стимулов.

Пример подобного различия можно видеть на рисунке 13.



Рис. 13. Различия между связанными с событием потенциалами (аудиальная инициализация фонем минус контроль); регионы значимых различий выделены темно-синим цветом; серый – 95% доверительный интервал, тёмно-серый – 99% доверительный интервал; ССП построен на 25 участниках.

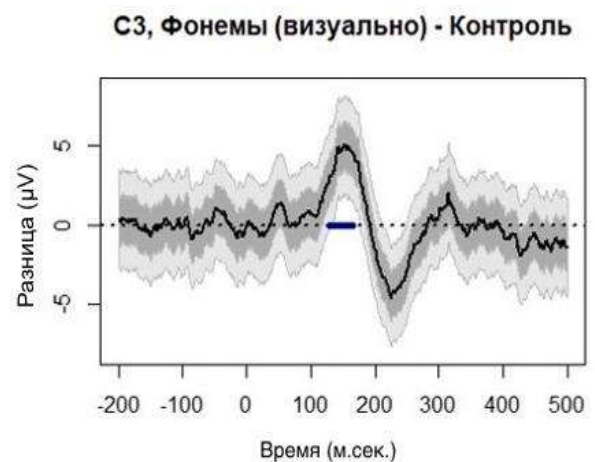


Рис. 14. Различия между связанными с событием потенциалами (визуальная инициализация фонем минус контроль); регионы значимых различий выделены темно-синим цветом; серый – 95% доверительный интервал, тёмно-серый – 99% доверительный интервал; ССП построен на 25 участниках.

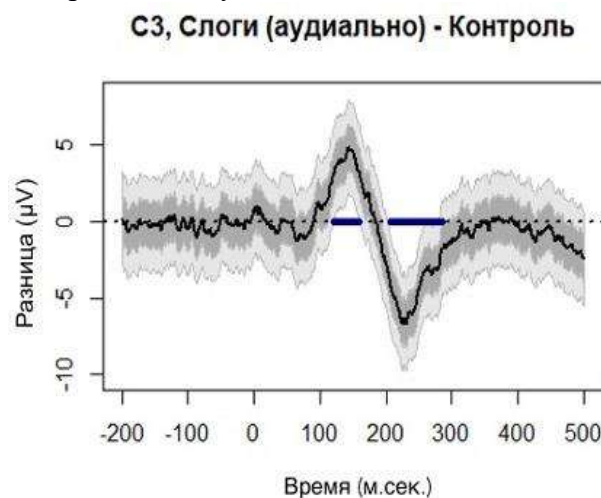


Рис. 15. Различия между связанными с событием потенциалами (аудиальная инициализация слогов минус контроль); регионы значимых различий выделены темно-синим цветом; серый – 95% доверительный интервал, тёмно-серый – 99% доверительный интервал; ССП построен на 25 участниках.

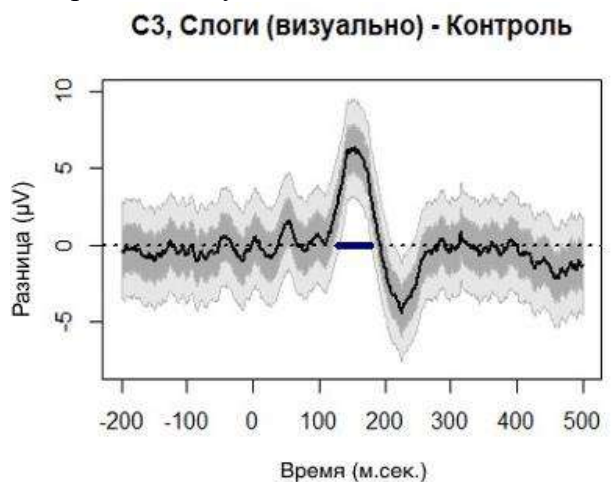


Рис. 16. Различия между связанными с событием потенциалами (визуальная инициализация слогов минус контроль); регионы значимых различий выделены темно-синим цветом; серый – 95% доверительный интервал, тёмно-серый – 99% доверительный интервал; ССП построен на 25 участниках.

Результаты выделения разностного связанного с событием потенциала на внутреннее проговаривание и восприятие

Относительно сравнения восприятия стимулов и их проговаривания: эффекты по каналам не зависят от того, визуальная была подача стимулов или нет ($p=0.11$ для фонем, $p=0.9$ для слогов). Однако эффекты по каналам зависят от того, воспринимает ли человек стимул или проговаривает ($p<0.001$), и этапы восприятия и проговаривания в целом тоже различаются ($p<0.001$). Если посмотреть серии со слогами в среднем, то наиболее выраженными различиями вновь являются каналы C3 и F3, причем при аудиальной подаче стимула более выражена негативная волна с латенцией в 200 мс (рис. 17), а при визуальной – позитивная волна с латенцией в 100 мс. В случае фонем ситуация похожа, однако существует некоторое отличие: при аудиальной подаче стимула также наблюдается значимая позитивная волна, но негативная при этом еще более выраженная и еще более длинная по сравнению со слогами (рис.18).

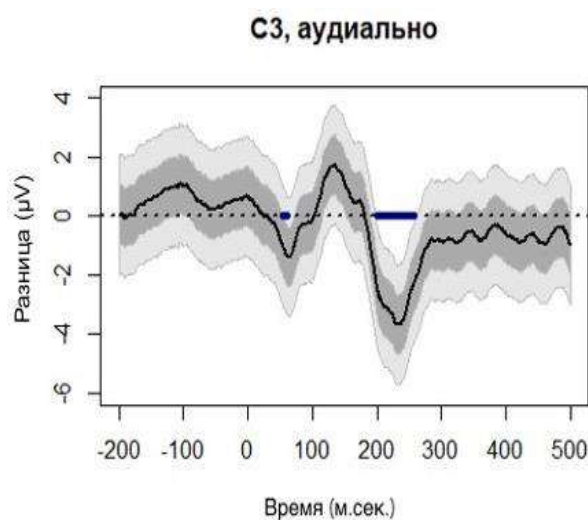


Рис. 17. Различия между связанными с событием потенциалами (проговаривание слогов минус аудиальное восприятие); регионы значимых различий выделены темно-синим цветом; серый – 95% доверительный интервал, тёмно-серый – 99% доверительный интервал; ССП построен на 25 участниках.

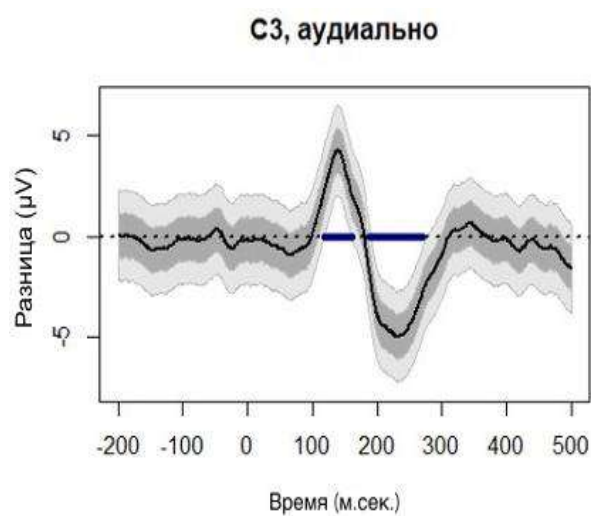


Рис. 18. Различия между связанными с событием потенциалами (проговаривание фонем минус аудиальное восприятие); регионы значимых различий выделены темно-синим цветом; серый – 95% доверительный интервал, тёмно-серый – 99% доверительный интервал; ССП построен на 25 участниках.

Если же анализировать различия по отдельным фонемам и слогам, то в случае фонем различия наблюдаются для всех фонем, но чаще всего при их аудиальной подаче и на каналах СЗ и FЗ. В случае слогов сколько-либо выраженные различия наблюдаются только относительно слогов БУ, ГУ, ГА, ФА, ФУ. Для слога ФА также больше всего видны различия между визуальной и аудиальной подачей стимула (позитивная волна в первом случае, негативная во втором, примерно та же латенция, что и до этого).

Сравнение внутреннего проговаривания и восприятия, локализация связанного с событием потенциала

Способ базируется на анализе динамики и уровня корреляции изменений сигнала по отведениям, но с добавлением данных, рассчитанных на основе расстояний от исследуемой точки до скальповых электродов (RU Patent № 2021116955/1 автор Vartanov A.V.). Используется модель источника низкочастотной (до 32 Гц) ЭЭГ и закон линейного убывания электрического потенциала от расстояния. С помощью факторного анализа главных компонент (PCA) и ортогонального вращения однозначно находится один фактор, который заведомо существует в объединенном массиве экспериментальных и искусственно сгенерированных данных. Каждый потенциальный источник находится независимо от других, поэтому предложенный способ не требует оценки возможного числа источников. Полученные и денормализованные факторные значения для экспериментальной ЭЭГ можно интерпретировать в качестве электрической активности «локального поля» (RU Patent № 2021116955/1 автор Vartanov A.V.; Vartanov, 2022). Наиболее близким к такой процедуре является метод пространственной фильтрации (представлен в US Patent №5263488, авторы VanVeen, Joseph, Несох, 23.11.1993), а также его последующее развитие в виде группы методов локализации источников, объединенных общим названием «формирование диаграммы направленности» («beamforming»). На основе локализации ЭЭГ активности для каждого отдельного испытуемого была осуществлена процедура усреднения

связанного с событием потенциала по всей группе испытуемых по всем предъявленным фонемам с оценкой 95% доверительного интервала для четырех анализируемых случаев: восприятия и мысленного проговаривания с учетом модальности инициализации (Шевченко, Вартанов, 2022).

В результате была исследована активность в 33 точках, выбранных по атласу MNI152 в центре следующих структур: *Brainstem*, *Mesencephalon*, *Hypothalamus*, *Caputn.Caudati L*, *Caputn.Caudati R*, *Medula Oblongata*, *G. Cingulate Medialis*, *Globus Pallidus Medialis L*, *Globus Pallidus Medialis R*, *Corpus Amygdaloideum L*, *Corpus Amygdaloideum R*, *Anterior Cingulate BA32*, *Dorsomedial prefrontal cortex BA9 L*, *Dorsomedial prefrontal cortex BA9 R*, *Hippocampus L*, *Hippocampus R*, *Insula L BA13*, *Insula R BA13*, *Parietal cortex BA7 L*, *Parietal cortex BA7 R*, *Putamen L*, *Putamen R*, *Supramarginal gyrus BA40 L*, *Supramarginal gyrus BA40 R*, *Thalamus L*, *Thalamus R*, *VI BA17 L*, *VI BA17 R*, *Ventral Striatum BA25*. Полученные ССП на картинках далее представлены следующим образом:

1) Восприятие:

Метка 1 – Аудиальное предъявление стимула (сплошная линия)

Метка 2 – Визуальное предъявление стимула (пунктирная линия)

2) Проговаривание:

Метка 3 – Проговаривание при аудиальной инициализации (сплошная линия)

Метка 4 – Проговаривание при визуальной инициализации (пунктирная линия)

Таким образом, этот анализ дает возможность оценить выраженность компонентов ССП, а метод локализации позволяет оценить ССП в тех источниках, которые задаются исследователем.

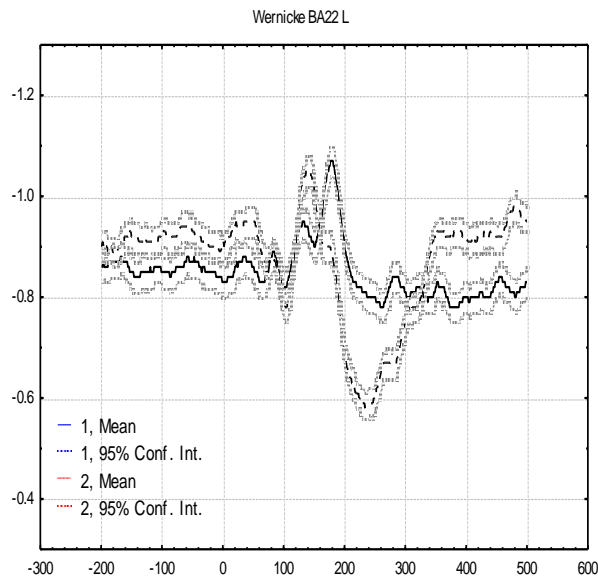


Рис. 19. ССП зоны Вернике BA22L Mean 1 – аудиальное восприятие, Mean 2 – визуальное восприятие (Метка 1 – сплошная линия; Метка 2 – пунктирная линия); ССП построен на 25 участниках.

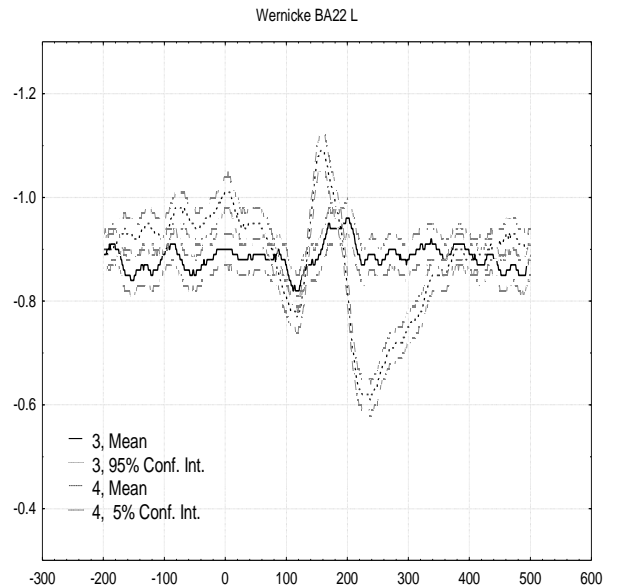


Рис. 20. ССП зоны Вернике BA22L Mean 3 – аудиальное восприятие, Mean 4 – визуальное восприятие (Метка 3 – сплошная линия; Метка 4 – пунктирная линия); ССП построен на 25 участниках.

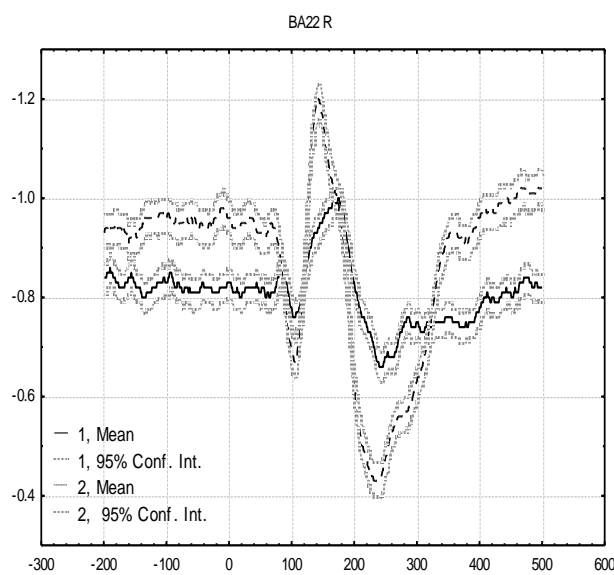


Рис. 21. ССП зоны Вернике BA22R Mean 1 – аудиальное восприятие, Mean 2 – визуальное восприятие (Метка 1 – сплошная линия; Метка 2 – пунктирная линия); ССП построен на 25 участниках.

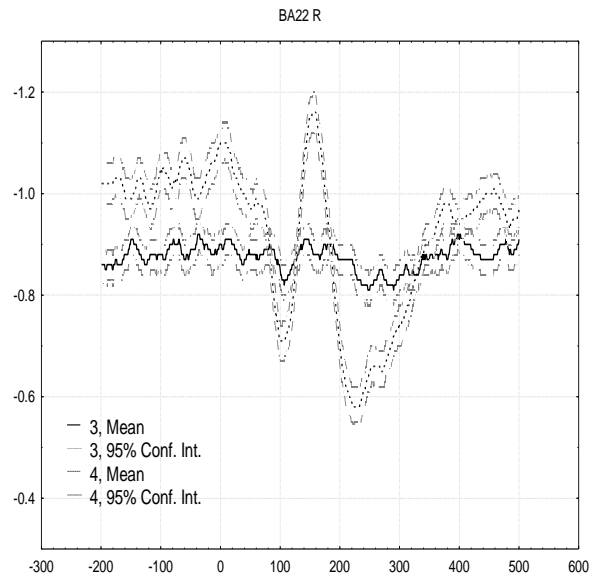


Рис. 22. ССП зоны Вернике BA22R Mean 3 – аудиальное восприятие, Mean 4 – визуальное восприятие (Метка 3 – сплошная линия; Метка 4 – пунктирная линия); ССП построен на 25 участниках.

Так, можно оценить сдвиг по латенции в зоне Вернике, которая связана как с процессами фонематического восприятия, так и с речевой деятельностью.

На рис. 23-26 представлены ССП, локализуемые в супромаргинальной извилине левого и правого полушарий.

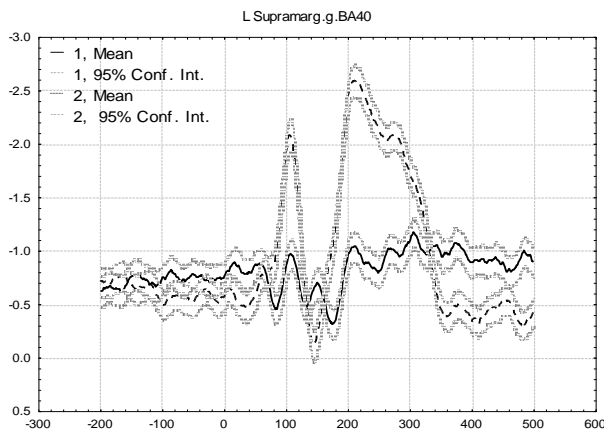


Рис. 23. ССП супромаргинальной извилины BA40L Mean 1 – аудиальное восприятие, Mean 2 – визуальное восприятие (Метка 1 – сплошная линия; Метка 2 – пунктирная линия); ССП построен на 25 участниках.

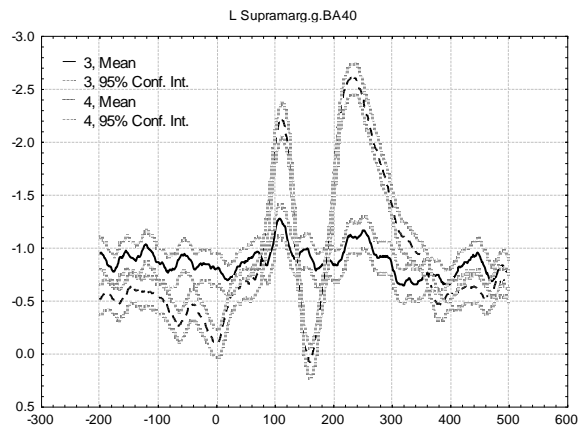


Рис. 24. ССП супромаргинальной извилины BA40L Mean 3 – аудиальное восприятие, Mean 4 – визуальное восприятие (Метка 3 – сплошная линия; Метка 4 – пунктирная линия); ССП построен на 25 участниках.

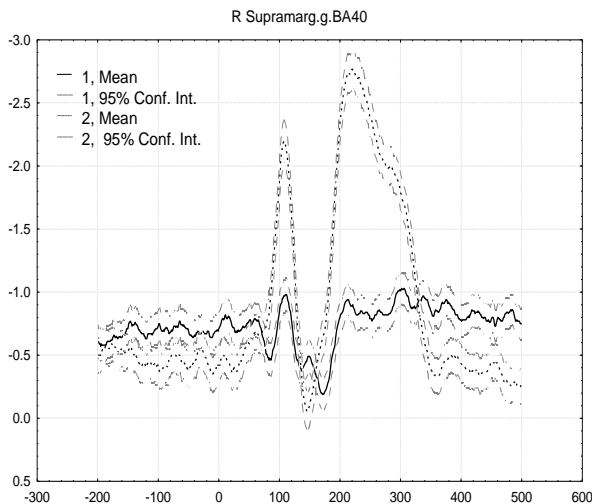


Рис. 25. ССП супромаргинальной извилины BA40R Mean 1 – аудиальное восприятие, Mean 2 – визуальное восприятие (Метка 1 – сплошная линия; Метка 2 – пунктирная линия); ССП построен на 25 участниках.

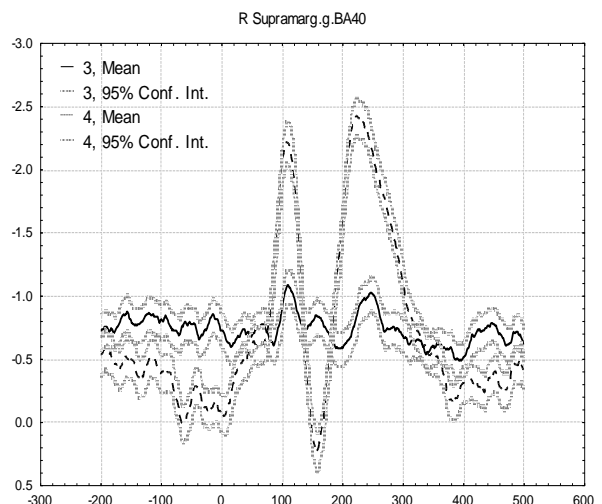


Рис. 26. ССП супромаргинальной извилины BA40R Mean 3 – аудиальное восприятие, Mean 4 – визуальное восприятие (Метка 3 – сплошная линия; Метка 4 – пунктирная линия); ССП построен на 25 участниках.

Различия ССП при внутреннем проговаривании на разные типы инициализации

Между проговариваниями в зависимости от визуальной/аудиальной подачи стимула также есть значимое различие вне зависимости от типа стимула ($p < 0.001$), но изменения по каналам значимы только для фонем ($p = 0.03$). В случае сравнения всех проговариваний с проговариванием на условный стимул в виде японских слов значимые различия опять были выявлены во всех четырех случаях ($p < 0.001$ для всего аудиального, $p = 0.01$ для визуальных фонем, $p = 0.009$ для визуальных слогов), но значимое взаимодействие канала с этапом нашлось только для аудиальных слогов ($p < 0.001$).

На основании полученных результатов, согласно функциональному дисперсионному анализу, все экспериментальные этапы имеют значимые различия, но при этом пространственное распределение ССП значимо различается только для аудиальных стимулов. Также было проведено сравнение всех вариантов инициализации с проговариванием на условный стимул в виде японских слов, значимые различия опять нашлись во всех четырех случаях, но значимое взаимодействие канала с этапом было установлено только для аудиальных слогов.

Есть также различия по отдельным фонемам и слогам, имеющим сходную структуру компонентов.

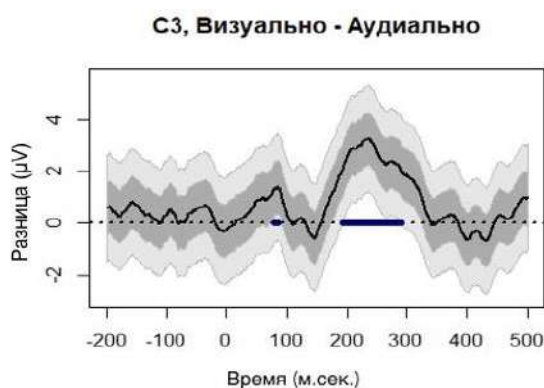


Рис. 27. Различия между связанными с событием потенциалами на проговаривание фонем (визуальная подача стимула минус аудиальная); регионы значимых различий выделены темно-синим цветом; серый – 95% доверительный интервал, тёмно-серый – 99% доверительный интервал; ССП построен на 25 участниках.

2.4.2. Результаты исследования внутреннего проговаривания на инициализацию условным стимулом методом электроэнцефалографии

В случае сравнения всех проговариваний с проговариванием на условный стимул в виде японских слов значимые различия опять нашлись во всех четырех случаях ($p < 0.001$ для всего аудиального, $p = 0.01$ для визуальных фонем, $p = 0.009$ для визуальных слогов), но значимое взаимодействие канала с этапом нашлось только для аудиальных слогов ($p < 0.001$). Результаты в целом крайне похожи на те, которые были получены на этапе с контролем, но с меньшим количеством значимых различий по регионам.

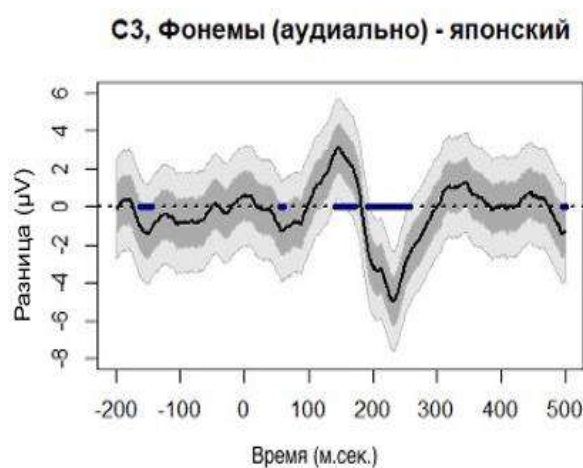


Рис. 28. Различия между связанными с событием потенциалами (аудиальные фонемы минус инициация условным японским стимулом); регионы значимых различий выделены темно-синим цветом; серый – 95% доверительный интервал, темно-серый – 99% доверительный интервал; ССП построен на 25 участниках.

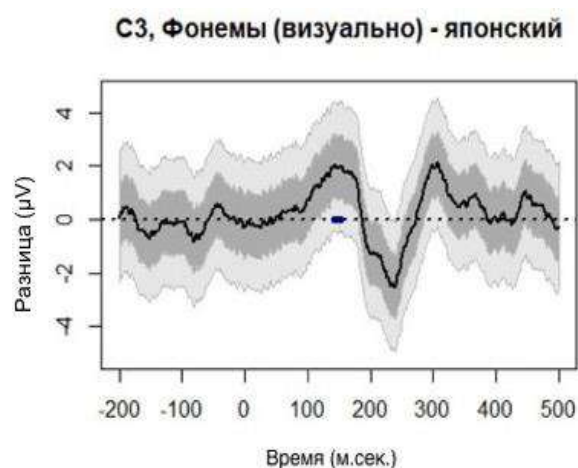


Рис. 29. Различия между связанными с событием потенциалами (визуальные фонемы минус инициация условным японским стимулом); регионы значимых различий выделены темно-синим цветом; серый – 95% доверительный интервал, темно-серый – 99% доверительный интервал; ССП построен на 25 участниках.

2.4.3. Построение классификатора на основе метода опорных векторов (SVM-RE)

Данный пункт был подготовлен на основе совместных исследований (Gavrilenko et al., 2021; Suyuncheva et al., 2021), в которых подробно представлены результаты по реконструкции фонем и слогов с помощью машинного обучения.

Для распознавания внутреннего проговаривания был использован алгоритм извлечения и нормализации признаков из сигнала электроэнцефалограммы, полученного во время внутреннего проговаривания.

Выбор каналов производился при учете локализации данных процессов. В соответствии с пунктом 1.5 главы 1, а также с результатами главы 2 пункта 2.3.1 извлечение признаков проводилось на основе записанных с этих каналов данных – F7, F3, T3, C3 (Gavrilenko et al., 2021).

В процессе проведения исследования была проведена процедура фильтрации с применением полосового фильтра Баттерворта (частота среза – 3-30 Гц) для избавления от слишком высоких частот в записанном сигнале, не связанных с внутренним проговариванием (Gavrilenko et al., 2021).

Для предварительной обработки данных сигнал был разделен в соответствии с временными участками, во время которых осуществлялся процесс внутреннего проговаривания. Для выполнения данной процедуры во время записи сигнала были поставлены метки, соответствующие моментам времени предъявления стимула и протеканию процесса внутреннего проговаривания. Из полученной записи были выбраны участки длительностью 350 мс: со 150 мс по 550 мс. Далее участки были представлены в виде 7 или 10 векторов, где вектор соответствует каждой фонеме или слогу.

К соответствующим фонемам или слогам-векторам было применено вейвлет-преобразование с использованием в качестве функции комплексного вейвлета Морле. Данная функция является скалярным произведением синусоиды на гауссиан. В исследовании использовался именно этот вейвлет, поскольку другие исследования с его применением были успешными. В семействе содержалось 40 вейвлетов (с увеличением количества циклов от 4 до 10 и с сопутствующим увеличением частоты) для обнаружения активности с длинными и короткими периодами активации. Таким образом, были получены матрицы распределения энергии сигнала по частотам и во времени.

Распределение мощности сигнала электроэнцефалограммы происходит в соответствии со степенным законом – мощность на низких частотах

соответствует большей величине, чем на высоких. Для устранения данного эффекта была применена нормализация по базовой линии. Данные были преобразованы на частотах к единому масштабу, а все численные значения были приведены к единой метрике (Cohen, 2014), Частотно-временная динамика сигнала была отделена от фонового сигнала, записанного до внутреннего проговаривания фонем. Для этого был произведен расчет процентного изменения, которое может быть рассмотрено как изменение мощности относительно базовой линии.

Реализация предложенного алгоритма извлечения и нормализации признаков из сигнала электроэнцефалограммы, полученного во время внутреннего проговаривания, была осуществлена при использовании программы MATLAB (выпуск R2019a) с расширением Signal Processing Toolbox и функцией `edf Read` (Gavrilenko et al., 2021).

Для расчета вейвлет-преобразования применяется теорема о свертке: при выполнении определенных условий преобразование Фурье свертки двух сигналов равнозначно скалярному произведению их преобразований Фурье. При использовании данной теоремы операция свертки вейвлета и сигнала во временной области когерентна их скалярному произведению в частотной области. Переход из временной в частотную область реализован посредством применения к сигналу операции преобразования Фурье и обратного преобразования Фурье. Таким образом, при использовании вейвлет-преобразования получены коэффициенты мощности сигнала на определенной частоте в определенный момент времени.

На полученных признаках была проведена процедура попарной классификации семи фонем. Был использован метод SVM с сигмоидным ядром. Ниже представлены таблицы точности попарной классификации внутреннего проговаривания фонем и слогов во время звукового и визуального предъявления стимулов для всех испытуемых (для каждого из испытуемых таблицы попарной классификации представлены в Приложении 2 (табл. 1 – 8)), признаки которого были извлечены предложенным методом.

Таблица 5. Пример попарной классификации слогов(визуально) с помощью SVM, средние показатели реконструируемости в процентах

	ФА	ГА	МА	РА	БУ	РУ	МУ	ФУ	ГУ
БА	56%	61%	65%	67%	65%	65%	65%	68%	56%
ФА		62%	63%	61%	60%	62%	60%	62%	61%
ГА			66%	57%	66%	66%	65%	65%	62%
МА				54%	63%	60%	70%	62%	65%
РА					60%	66%	67%	61%	66%
БУ						58%	55%	60%	59%
РУ							64%	61%	61%
МУ								65%	62%
ФУ									58%

Таблица 6 показывает результаты попарной классификации внутреннего проговаривания фонем на условный стимул (результаты попарной классификации внутреннего проговаривания фонем на условный стимул по каждому испытуемому представлены в приложении 2 (табл. 9, 10)).

Таблица 6. Пример попарной классификации фонем на условный стимул с помощью SVM, средние показатели реконструируемости в процентах (Suyuncheva et al., 2021)

	Б	Ф	Г	М	Р	У
А	60%	62%	63%	60%	62%	57%
Б		64%	59%	67%	68%	70%
Ф			70%	68%	62%	64%
Г				68%	58%	61%
М					72%	64%
Р						69%

Мы получили следующую статистику классифицируемости:

- среднее значение классификатора фонем, предъявляемых аудиально – 62%;
- среднее значение классификатора фонем, предъявляемых визуально – 62%;
- среднее значение классификатора слогов, предъявляемых визуально – 61%;

- среднее значение классификатора слогов, предъявляемых аудиально – 63%;
- среднее значение классификатора фонем при инициации на условный стимул (японские слова) – 64%.

2.4.4. Построение классификатора на основе сверточных нейронных сетей

Данный пункт был подготовлен на основе совместного исследования (Seleznyev et al., 2023), в котором представлены результаты по реконструкции фонем и слогов с помощью свёрточных нейронных сетей.

Каждая фонема представляет собой частотно-временную спектрограмму для одного из каналов записи. Так, например, на рис. 30 представлена спектрограмма, где 200 мс предшествуют проговариванию фонемы, а 600 мс соответствуют ее мысленному проговариванию один раз.

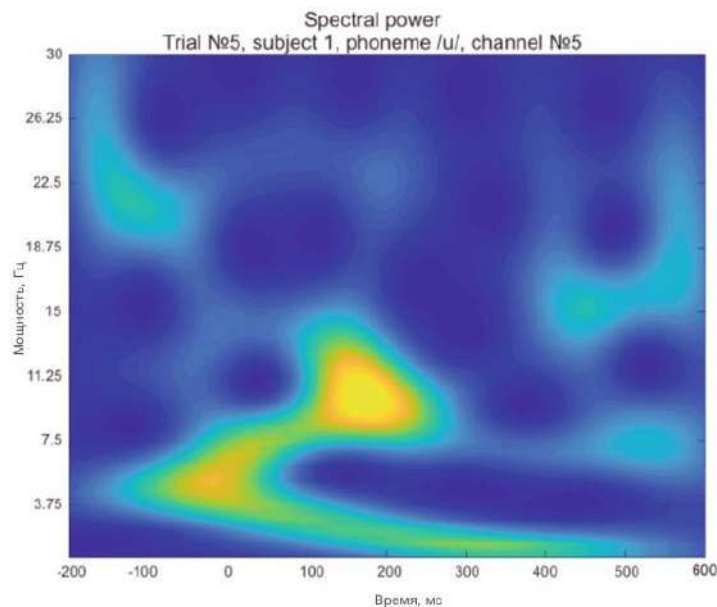


Рис. 30. Пример представления фонемы «У» в частотно-временной спектрограмме построенный по одному испытуемому.

Так, получается, что задача классификации временных рядов будет сводиться к классификации изображений, полученных после представления фонем в частотно-временной спектрограмме. При решении такой задачи

хорошие результаты показывают сверточные нейронные сети. Их преимущество проявляется в их основе на ядрах свертки, которые позволяют выделить отдельные признаки объекта или даже виды признаков и сосредоточить на них внимание классификатора. Такой подход дает возможность решить проблему снижения процента неверно классифицированных примеров для данных других испытуемых (Seleznyev et al., 2023).

Базовая архитектура построенной сверточной нейронной сети

Архитектурой выступает сверточная нейронная сеть, состоящая из L сверточных и подвыборочных слоев и R полносвязных слоев без учета выходного слоя (рис. 31).

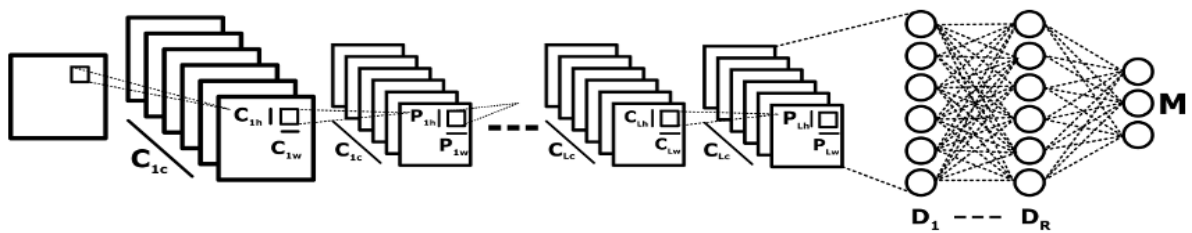


Рис. 31. Архитектура сверточной нейронной сети (Селезнев с соав., 2023).

C_{1c} , C_{1w} , C_{1h} задают число каналов i сверточного слоя, а также ширину и высоту сверточного ядра для i сверточного слоя. P_{1w} , P_{1h} задают ширину и высоту подвыборочного ядра для i подвыборочного слоя. Общее число пар сверточных и подвыборочных слоев составляет L . D_r задает число нейронов полносвязной сети, M задает число выходных нейронов всей сети. Число полносвязных слоев без учета последнего слоя – R . Кроме того, применяют дополнительные Dropout слои с целью решения проблемы переобучения и проблемы затухающих и растущих градиентов. Доля «исключаемых» параметров для каждого сверточного слоя i задается значением Q_i . Доля «исключаемых» параметров для каждого полносвязного слоя r задается значением S_r (Seleznyev et al., 2023).

Было выделено 4 подзадачи, каждой из которых соответствует своя сверточная сеть (Селезнев с соавт., 2023):

1. Бинарная классификация для каждого испытуемого для всех возможных неповторяющихся пар фонем.
2. Многоклассовая классификация с числом классов равным числу фонем для каждого испытуемого.
3. Бинарная классификация для всех неповторяющихся пар фонем для объединения данных всех испытуемых.
4. Многоклассовая классификация с числом классов равным числу фонем для объединения данных всех испытуемых.

Такое разделение необходимо из-за трудностей построения архитектуры сети, которая может одинаково эффективно решать каждую задачу одновременно.

Число выходных нейронов сети M зависит от рассматриваемой подзадачи классификации: в подзадачах 1 и 3, $M = 1$, что соответствует вероятности выбора одной из двух фонем для бинарной классификации, в подзадачах 2 и 4, $M = 7$, что соответствует числу фонем в задаче многоклассовой классификации. В качестве функций активации каждого слоя во всех 4 подзадачах используется ReLU, в дополнение к ней в 1 и 3 подзадачах используется сигмоидальная функция активации. Применяется метод пакетного градиентного обратного распространения ошибки с использованием функции потерь Cross Entropy Loss для многоклассовой классификации и Binary Cross Entropy Loss в случае бинарной классификации (Seleznyev et al., 2023).

Итоговый алгоритм (Seleznyev et al., 2023):

1. Сверточные сети с использованием сверточных и подвыборочных слоев, общие параметры сети задаются в зависимости от типа задачи.

2. Обучение алгоритма за счет пакетного алгоритма градиентного обратного распространения ошибок, заданного числом эпох использования постепенного снижения коэффициента обучения.
3. Для увеличения числа обучающих примеров предлагается аугментировать исходные данные.

Методика проведения экспериментального исследования

В рамках экспериментального исследования была проведена оценка эффективности решения задачи классификации 7 разных фонем для 12 разных людей для каждой задачи исследования. Кроме того, каждая подзадача имела гиперпараметры алгоритма для обучения (Селезнев с соавт., 2023).

Для решения задач бинарной классификации (1 и 3 подзадача) производится оценка точности попарной классификации для каждой пары фонем за исключением совпадающих фонем. Для решения задач многоклассовой классификации (2 и 4 подзадача) предлагается оценить среднюю и наибольшую точность классификации для каждого отдельного человека (подзадача 2) и общую точность классификации в случае объединения всех данных (подзадача 4) (Seleznyev et al., 2023).

Результаты экспериментального исследования

Средняя точность попарной классификации всех фонем по всем запускам и испытуемым составляет 73.98% (среднее значение всех ячеек Таблицы 7). Максимальная средняя по всем испытуемым и запускам точность попарной классификации, составляющая 80.60% (максимальное значение среди ячеек Таблицы 8) для фонем Б и Ф (Seleznyev et al., 2023).

Таблица 7. Средняя попарная классификация по всем запускам и испытуемым (Seleznyev et al., 2023)

	А	Б	Ф	Г	М	Р	У
А		76.75%	75.06%	73.17%	78.12%	78.14%	75.33%
Б			80.60%	77.71%	72.96%	73.66%	70.70%
Ф				75.22%	76.57%	67.23%	74.82%
Г					66.09%	75.60%	71.46%
М						76.36%	67.55%
Р							70.52%
У							

Таблица 8. Лучшая попарная классификация в среднем по всем запускам, испытуемый 1 (Seleznyev et al., 2023)

	А	Б	Ф	Г	М	Р	У
А		94.87%	54.65%	82.89%	83.70%	82.89%	86.59%
Б			91.25%	92.86%	89.53%	94.29%	85.90%
Ф				82.05%	77.66%	76.92%	74.42%
Г					80.95%	97.14%	89.47%
М						88.10%	77.78%
Р							84.21%
У							

Для задачи 3 результаты приведены в Таблице 9, общая средняя точность за все запуски составляет 76.44%, тогда как максимальная средняя за все запуски точность составляет 93.02%. Несмотря на полученную среднюю и максимальную точности, для некоторых пар фонем точность бинарной классификации близка к случайному угадыванию (Seleznyev et al., 2023).

Таблица 9. Точность попарной классификации для всех данных (Seleznyev et al., 2023)

	А	Б	Ф	Г	М	Р	У
А		63.46%	93.02%	76.84%	92.87%	90.75%	86.20%
Б			53.52%	79.86%	71.69%	74.83%	78.34%
Ф				80.87%	82.75%	88.77%	63.02%
Г					59.01%	58.88%	66.41%
М						87.00%	79.42%
Р							77.74%
У							

В задаче 2 средняя точность на всех запусках составляет 61.98%, что связано с неудачными начальными инициализациями сетей, в некоторых случаях приводившими к необучаемости сети, а также с малыми размерами обучающих выборок для отдельных испытуемых. Максимальная средняя точность по всем запускам среди всех испытуемых составляет 80.34% (усреднение точности для каждого испытуемого по всем запускам). Таким образом, алгоритм обучения можно комбинировать с алгоритмом мультистарта с последующим выбором лучшей сети среди обученных для каждого испытуемого в отдельности (Seleznyev et al., 2023).

В задаче 4 многоклассовой классификации на данных всех испытуемых точность классификации на лучшем из 3 запусков составляет 80.18%. Обученную на данных всех испытуемых сеть в последствии можно применять в качестве предобученной сети для более точного дообучения под отдельных испытуемых. Обученная на данных всех испытуемых сеть позволяет изначально значительно снизить величину метрики FAIL и повысить величину метрики MATCH, таким образом позволяя решить поставленную ранее задачу оптимизации (Seleznyev et al., 2023).

2.5. Исследование внутреннего проговаривания методом функциональной магнитно-резонансной томографии

Испытуемые. В исследовании приняли участие 30 респондентов (9 мужского пола и 21 женского пола в возрасте от 20 до 30 лет, средний возраст – 24 года). Все участники имели ведущую правую руку, не имели неврологических и психических отклонений, что подтверждается предварительным опросом. Все участники подписали информированное согласие после объяснения протокола эксперимента.

Оборудование и параметры получения МР-изображений. Функциональная МРТ выполнялась на Siemens Magnetom Skyra 3 TMRI Machine (Siemens Medical Systems, Erlangen, Germany). Для каждого участника были получены T1-взвешенные структурные изображения (TR = 2200 мс, TE

= 2,48 мс, FA = 6°, матрица = 256 × 256, FoV = 230 мм, 176 срезов, толщина среза = 1 мм). Для получения функциональных T2*-взвешенных (41 объём) изображений использовалась импульсная последовательность градиент-эхо EPI (TR = 3000 мс, TE = 30 мс, FA = 90°, матрица 128 × 128, FoV = 260 мм). Движение головы было минимизировано с помощью специально подобранного подголовника.

Процедура. Перед началом каждого этапа испытуемые слушали стандартизированную инструкцию. Серии были следующими: простое прослушивание стимульного материала, прослушивание с последующим повторением, также была записана фоновая серия, во время которой испытуемый спокойно лежал в томографе. Продолжительность каждой серии составляла 2 минуты (всего эксперимент занимал около 8 минут). Длительность каждого стимула составляла 1500 мс, затем после паузы (500 мс) следовал звуковой сигнал, который обозначал начало проговаривания. Стимулы были произнесены женским голосом. Предъявление производилось через специальные наушники. Испытуемые получали следующие инструкции: (1) «Пожалуйста лежите спокойно, не совершайте резких движений»; (2) «Пожалуйста лежите спокойно, не совершайте резких движений. Вам предлагается прослушать стимульный материал. Повторять его не надо»; (3) «Вам будут демонстрироваться в случайном порядке через наушники стимулы. После специального звукового сигнала Вам необходимо проговорить заданный стимул вслух/про себя (для четвертой серии) до повторного звукового сигнала. Проговаривая заданный стимул постарайтесь сохранить заданную длительность и интонацию». Все этапы проводились в одну сессию. Дизайн фМРТ-исследования был не блочным. В течение всего эксперимента глаза испытуемых были закрыты.

Стимулы. Для исследования особенностей внутреннего проговаривания были подготовлены серии проговаривания фонем и серии проговаривания слогов. В качестве стимульного материала представлялись:

- 7 фонем (А – [а], Б – [б], Ф – [ф], Г – [г], М – [м], Р – [р], У – [у]);

- 10 слогов (БА – [ба], ФА – [фа], ГА – [га], МА – [ма], РА – [ра], БУ – [бу], РУ – [ру], МУ – [му], ФУ – [фу], ГУ – [гу]).

Анализ фМРТ-изображений. Анализ данных выполняли с использованием пакета SPM12 ((SPM12, 2020) The Wellcome Centre for Human Neuroimaging, UCL Queen Square Institute of Neurology). Проводилась предобработка полученных изображений. Производилась коррекция времени среза (slicetiming) 41 срез, TR – 3, TA – 2,927 и контрольный срез (referenceslice) – 20. Производилось нормализация со значениями [3:3:3] и сглаживание (smoothing) FWHMс значением [8:8:8]. Для каждого испытуемого проводился поворот срезов анатомии и полученных функциональных данных для унификации полученных результатов. Для получение валидных результатов были удалены первые 5 срезов в связи с потенциальным артефактами.

Далее производился анализ групповой статистики с помощью одновыборочного t-критерия Стьюдента. После t-статистики были преобразованы в нормальное стандартное отклонение (z). Воксельный порог активации соответствовал $p_{FWEuncorr} < 0.001$.

2.6. Результаты исследования внутреннего проговаривания методом функциональной магнитно-резонансной томографии

Сравнения внешнего и внутреннего проговаривания

В результате проведенного анализа были выделены воксели, значимо изменявшие свою активность в зависимости от реализуемых испытуемыми задач. В данном исследовании серии сравнивались не с фоном, поэтому выделение отличий происходило на основании сравнения серий внешнего и внутреннего проговаривания. Исходя из этого, могли быть представлены не все структуры, связанные с процессами речи, а лишь те, где есть различия в сравнении серий. Разность кластеров, полученная при вычитании из данных одной серии данных другой серии представлена в таблице 10.

Таблица 10. Значимые фМРТ-различия в кластерах при сравнении внешнего и внутреннего проговаривания по данным 30 участников. Координаты представлены в пространстве MNI

Кластер	Объём кластера	T	X {mm}	Y {mm}	Z {mm}	L/R	Структура	BA
Внешнее проговаривание фонем в сравнении с внутренним проговариванием фонем								
1	135	3.82	-12	53	-4	L	Передняя префронтальная кора	10
Внешнее проговаривание слогов в сравнении с внутренним проговариванием слогов								
2	270	3,86	42	-43	-14	R	Веретенообразная извилина	37
3	135	3,7	-3	-4	30	L	Вентральная передняя поясная кора	24
Внутреннее проговаривание фонем в сравнении с внешним проговариванием								
4	1674	4,46	-3	-49	-46	L	Мозжечок	
		4,11	-3	-52	-55	L	Мозжечок	
		3,87	0	-49	-37	L	Мозжечок	
5	297	4,22	54	-1	-25	R	Средняя височная извилина	21
6	1080	4,19	30	-55	-43	R	Мозжечок	
7	324	3,88	33	-73	-52	R	Мозжечок	
Внутреннее проговаривание слогов в сравнении с внешним проговариванием слогов								
8	10422	6.65	-45	-49	-46	L	Мозжечок	
		6.16	-33	-61	-49	L	Мозжечок	
		5.91	-36	-52	-55	L	Мозжечок	
9	1566	5.53	51	5	-31	R	Височный полюс	38
10	1647	5.25	36	-49	-52	R	Мозжечок	
		4.61	27	-49	-52	R	Мозжечок	
		3.78	39	-46	-43	R	Мозжечок	
11	1890	5.19	-21	-67	29	L	Верхняя теменная долька (VisMotor)	7
		3.50	-24	-58	35	L	Верхняя теменная долька (VisMotor)	7
		3.45	-33	-79	26	L	Нижняя затылочная извилина (VisAssoc)	19
12	3780	4.87	-54	-25	20	L	Супрамаргинальная извилина	40
		4.83	-54	-37	11	L	Верхняя височная извилина	22

		4.30	-63	-31	35	L	Супрамаргинальная извилина	40
13	594	4.74	-6	-28	-46	L	Мост	
14	324	4.55	59	-22	50	R	Первичная соматосенсорная кора	1
		3.88	63	-22	41	R	Супрамаргинальная извилина	40
15	432	4.42	-54	8	-31	L	Височный полюс	38
16	648	4.42	-45	-70	29	L	Угловая извилина	39
17	135	4.14	-57	-58	-1	L	Веретенообразная извилина	37
18	189	3.98	-9	-70	62	L	Верхняя теменная доля (VisMotor)	7
19	459	3.94	-39	-70	17	L	Нижняя затылочная извилина (VisAssoc)	19
20	486	3.79	57	-49	44	R	Угловая извилина	39

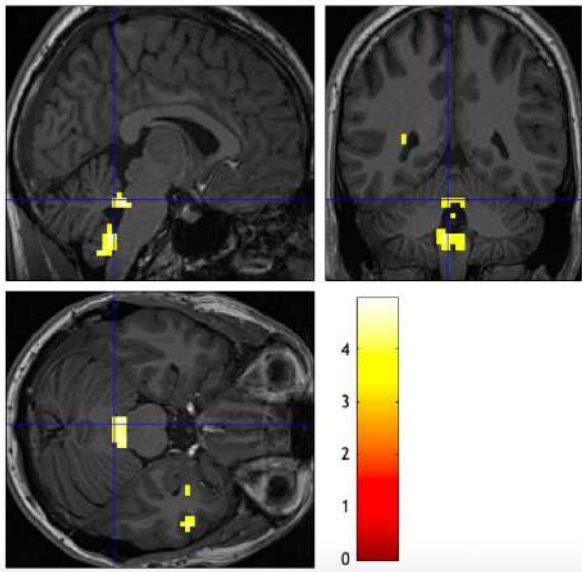


Рис. 32. ФМРТ изображение значимой разности активности зон головного мозга в серии сравнения внутреннего проговаривания фонем с внешним проговариванием (градиентом показаны зоны, связанные с внутренним проговариванием фонем). Изображение построено по усредненным данным 30 участников

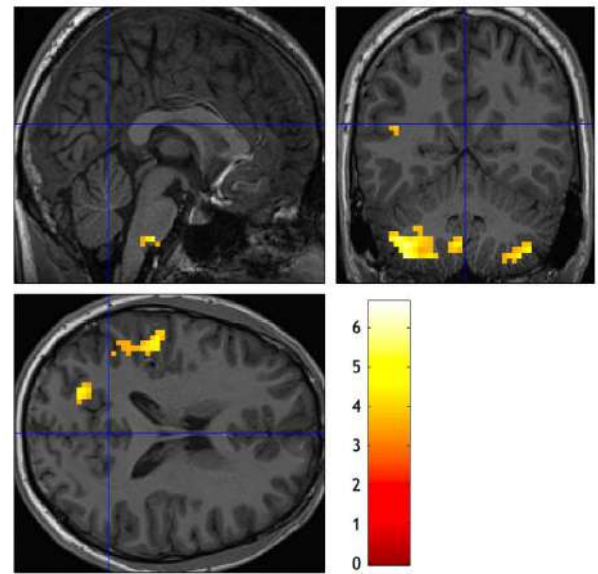


Рис. 33. ФМРТ изображение значимой разности активности зон головного мозга в серии сравнения внутреннего проговаривания слогов с внешним проговариванием (градиентом показаны зоны, связанные с внутренним проговариванием слогов). Изображение построено по усредненным данным 30 участников

Результаты сравнения восприятия и внутреннего проговаривания фонем и слогов

В результате проведенного анализа были выделены структуры, значимо связанные с процессами восприятия и намеренного проговаривания. Воксельный порог активации соответствовал $p_{FWEuncorr} < 0.001$. При этом показателем вклада в описание модели являются T-значения, относительно которых и отфильтрована таблица 11. ВА – номер цитоархитектонического поля по Бродману. Полученные координаты были интерпретированы с помощью сервиса MNI (MNI, 2020).

Таблица 11. Значимые фМРТ-различия в кластерах при сравнении внутреннего проговаривания фонем и слогов и их восприятия по данным 30 участников. Координаты представлены в пространстве MNI

Кластер	Объём кластера	T	X {mm}	Y {mm}	Z {mm}	L/R	Структура	ВА
Внутреннее проговаривание слогов по сравнению с восприятием								
1	6021	6.10	-36	-52	-55	L	Мозжечок	
		6.10	-36	-73	-52	L	Мозжечок	
		5.59	-45	-55	-46	L	Мозжечок	
2	2916	4.74	-51	-28	11	L	Первичная слуховая кора	41
		4.26	-51	-40	14	L	Верхняя височная извилина	22
		4.00	-51	-31	20	L	Супрамаргинальная извилина	40
3	675	4.73	51	11	-37	R	Височный полюс	38
		3.59	54	8	-19	R	Височный полюс	38
4	513	3.65	63	-34	38	R	Супрамаргинальная извилина	40
5	513	4.46	-65	-31	35	L	Супрамаргинальная извилина	40
		3.61	-65	-34	26	L	Супрамаргинальная извилина	40
6	1107	4.18	-30	-76	26	L	Угловая извилина	39
		3.50	-21	-67	29	L	Верхняя теменная доля (VisMotor)	7
7	945	4.30	-42	-70	14	L	Нижняя затылочная извилина (VisAssoc)	19

		4.11	-48	-79	14	L	Нижняя затылочная извилина (VisAssoc)	19
8	162	3.57	39	-73	41	R	Угловая извилина	39
Внутреннее проговаривание фоном по сравнению с восприятием								
9	33588	8.44	-36	-67	-49	L	Мозжечок	
		7.98	-39	-52	-52	L	Мозжечок	
		6.12	-45	-49	-46	L	Мозжечок	
10	29187	7.14	51	-49	32	R	Угловая извилина	39
		6.34	51	-55	41	R	Угловая извилина	39
11	4644	7.03	51	-7	-31	R	Нижняя височная извилина	20
		5.72	45	14	-37	R	Височный полюс	38
		5.58	51	5	-31	R	Височный полюс	38
12	17037	5.32	-33	-64	40	L	Верхняя теменная долька (VisMotor)	7
		5.24	-39	-49	38	L	Угловая извилина	39
13	3672	5.54	-21	-76	13	L	Первичная зрительная кора	17
		4.37	-39	-82	20	L	Нижняя затылочная извилина (VisAssoc)	19
14	3132	5.27	-51	2	-31	L	Височный полюс	38
		4.68	-42	11	-34	L	Височный полюс	38
15	4050	5.22	45	32	2	R	Треугольная часть	45
		5.17	45	44	-16	R	Орбитальная (глазничная) часть	47
		5.09	48	20	5	R	Оперкулярная часть	44
16	5103	5.14	27	-49	-46	R	Мозжечок	
		4.98	36	-52	-46	R	Мозжечок	
		4.51	30	-73	-55	R	Мозжечок	
17	3051	4.61	-33	26	29	L	Дорсолатеральная префронтальная кора (дорсальная)	9
18	162	4.32	-36	-85	29	L	Нижняя затылочная извилина (VisAssoc)	19
19	135	4.32	39	26	51	R	Фронтальные глазные поля	8
20	243	4.13	63	-31	14	R	Верхняя височная извилина	22
21	270	4.05	51	8	-13	R	Височный полюс	38
22	405	4.00	33	17	35	R	Фронтальные глазные поля	8

23	162	3.92	-57	-34	11	L	Верхняя височная извилина	22
24	297	3.91	9	41	35	R	Дорсолатеральная префронтальная кора (дорсальная)	9
25	324	3.90	-42	-61	-7	L	Веретенообразная извилина	37
26	135	3.83	-54	-28	23	L	Супрамаргинальная извилина	40
27	189	3.82	-51	-43	17	L	Верхняя височная извилина	22
28	189	3.80	-33	-79	-7	L	Нижняя затылочная извилина (VisAssoc)	19

На основе полученных координат были построены изображения с выделенными цветом зонами активации намеренного внутреннего проговаривания в сравнении с фонематическим восприятием (рис. 34 и 35).

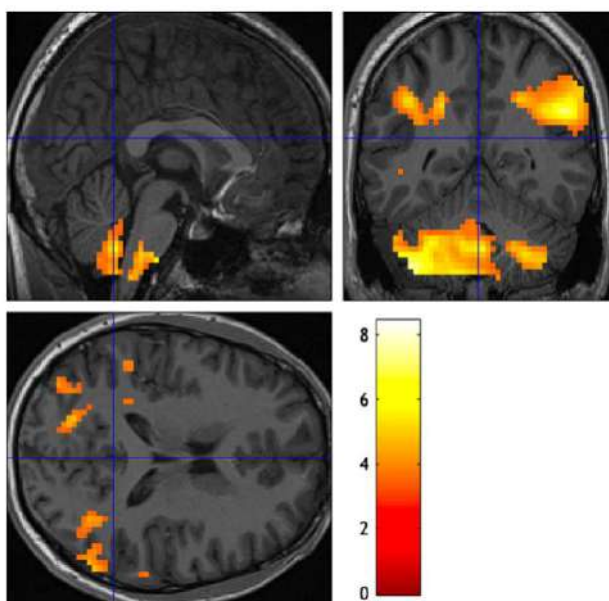


Рис. 34. ФМРТ изображение значимой разности активности зон головного мозга в серии сравнения внутреннего проговаривания фонем с восприятием (градиентом показаны зоны, связанные с внутренним проговариванием фонем). Изображение построено по усредненным данным 30 участников

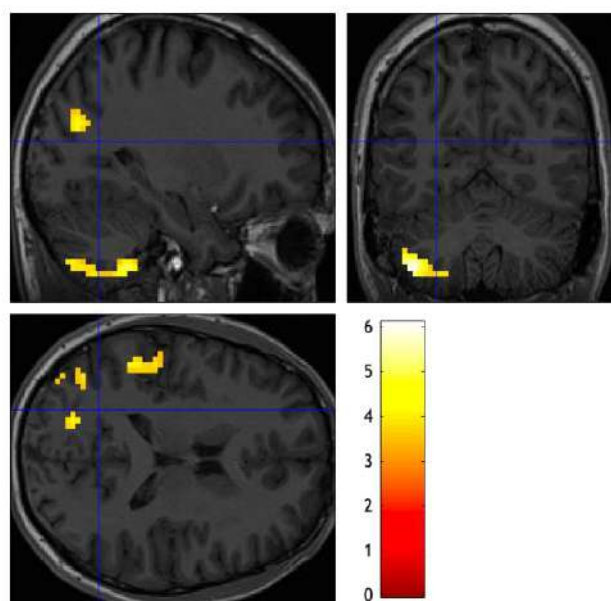


Рис. 35. ФМРТ изображение значимой разности активности зон головного мозга в серии сравнения внутреннего проговаривания слогов с восприятием (градиентом показаны зоны, связанные с внутренним проговариванием слогов). Изображение построено по усредненным данным 30 участников

2.7. Построение фМРТ-связности с помощью анализа независимых компонент

Для оценки связанности и выделения компонентов внутреннего проговаривания, также был использован метод анализа независимых компонент, который позволяет нивелировать потенциальные недостатки неблочного дизайна в связи, с тем, что производится оценка группового эффекта и связанности без индивидуального сравнения серий. Предобработанные данные были загружены в Group ICA-T версии 4.0 (GIFT, 2017). Для оценки независимых переменных использовались следующие параметры: количество компонент 20. Для корректирования модели использовался алгоритм MDL (FWHM) с параметрами [5:5:5] в Group ICA-T. Медианное количество найденных компонент равно 21. Алгоритм, используемый для анализа, – Fast ICA, преимуществом которого является устойчивость при нехватке фМРТ-данных, что позволяет его использовать в парадигме rt-фМРТ (Soldati et al., 2013). Сети отбирались на основе пространственной корреляции с маской речевой сети по атласу Stanford's Functional Imaging in Neuropsychiatric Disorders (Legacy FINDLAB Data and Functional ROIs, 2011). Полученные координаты были интерпретированы с помощью сервиса MNI (MNI, 2020).

Полученные сети были загружены в SPM12 (SPM12, 2020), после чего было проведено сравнение речевых сетей при разных задачах с поправкой FWE на уровне 0,05. Полученные сети были сравнены с данными пункта 2.6 для выявления общих языковых тенденций.

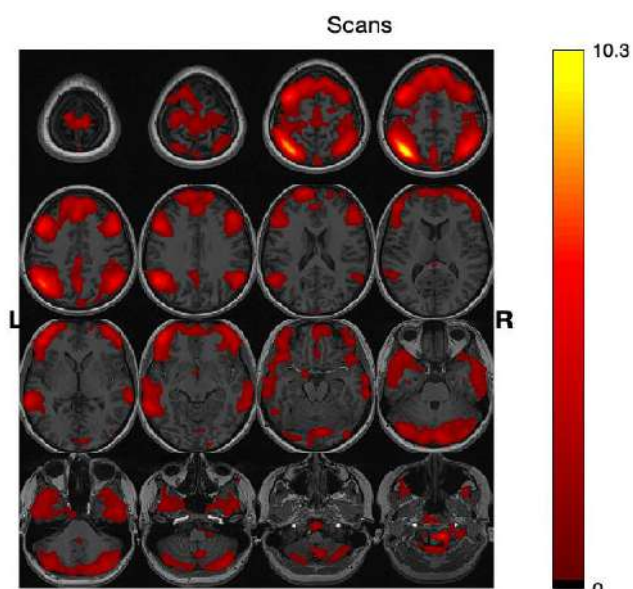


Рис. 36. Сеть, выделенная по фМРТ-данным, полученным при выполнении задачи внутреннего проговаривания фонем. Коэффициент корреляции 0.13864 со Стэнфордской языковой rs-фМРТ. Градиентом показаны зоны, связанные с внутренним проговариванием фонем.

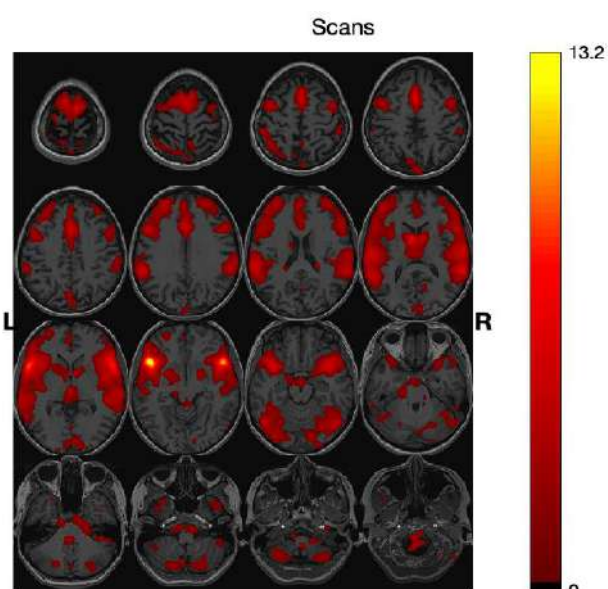


Рис. 37. Сеть, выделенная по фМРТ-данным, полученным при выполнении задачи внешнего проговаривания фонем. Коэффициент корреляции 0.18575 со Стэнфордской языковой rs-фМРТ. Градиентом показаны зоны, связанные с внешним проговариванием фонем.

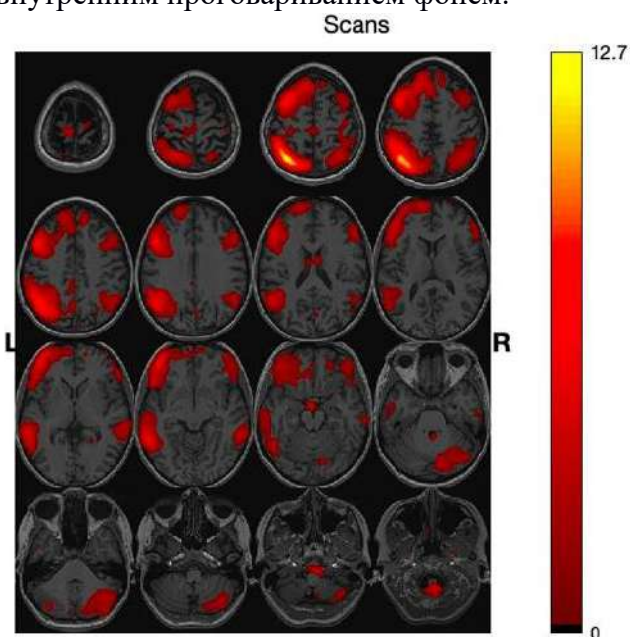


Рис. 38. Сеть, выделенная по фМРТ-данным, полученным при выполнении задачи внутреннего проговаривания слогов. Коэффициент корреляции 0.16318 со Стэнфордской языковой rs-фМРТ. Градиентом показаны зоны, связанные с внутренним проговариванием слогов.

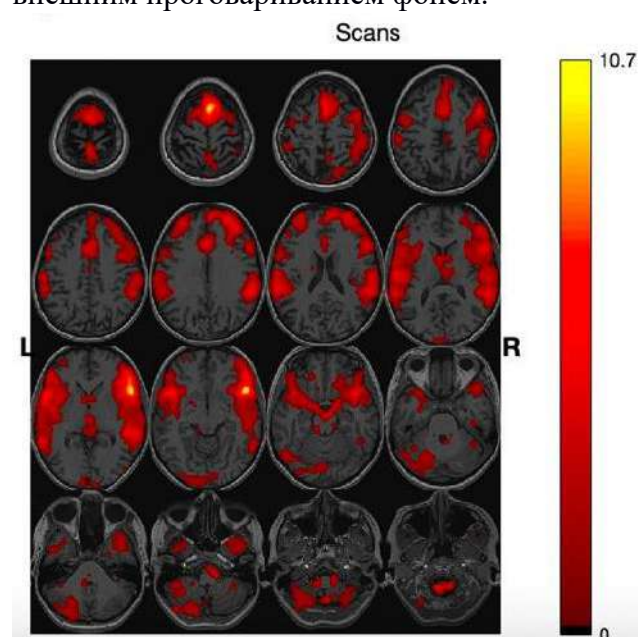


Рис. 39. Сеть, выделенная по фМРТ-данным, полученным при выполнении задачи внешнего проговаривания слогов. Коэффициент корреляции 0.1223 со Стэнфордской языковой rs-фМРТ. Градиентом показаны зоны, связанные с внешним проговариванием слогов.

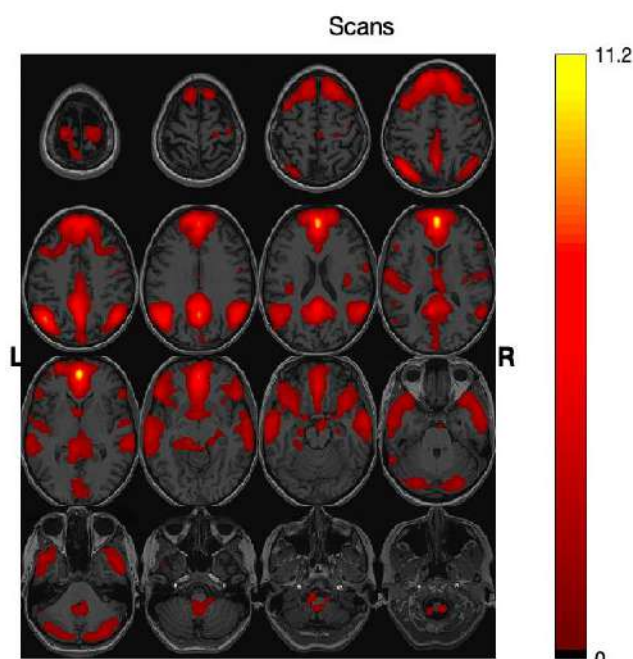


Рис. 40. Сетей, выделенная по фМРТ-данным, полученным при выполнении задачи аудиального восприятия фонем и слогов. Коэффициент корреляции 0.15171 со Стэнфордской языковой rs-фМРТ. Градиентом показаны зоны, связанные с восприятием фонем и слогов.

Ограничением анализа независимых компонент является выборочность компонентов, связанных с речевой задачей, поэтому не все различия могут быть отражены в полученных сетях. Так, для описания речевых и аудиальных процессов мы будем учитывать результаты полученных коннективити (приложение 3,4), и результаты, полученные при групповом сравнении (глава 2, пункт 2.6). Для оценки полимодальности в связи с унификацией речевой сети мы будем использовать лишь результаты группового сравнения. Получены следующие общие активации, которые совпадают и с пунктом 2.6, и с полученными таблицами связности (приложение 3). Так, при сравнении *внутреннего проговаривания фонем и слогов с их внешним проговариванием* отмечались общие активации в левой и правой верхней теменной доле (VisMotor) (BA7), правой первичной соматосенсорной коре (BA1), билатеральная активация угловой извилины (BA39) и билатеральная активация мозжечка. В серии сравнения *внешнего проговаривания фонем и слогов с их внутренним проговариванием* отмечались общие активации левой вентральной передней поясной коры (BA24), левой супрамаргинальной

извилины (BA40), правой веретенообразной извилины (BA37). В серии сравнения *внутреннего проговаривания фонем и слогов с их восприятием* отмечались общие активации левой и правой супрамаргинальной извилины (BA40), правой угловой извилины (BA39), левой верхней теменной дольки (VisMotor) (BA7), правой части фронтальных глазных полей (BA8), билатерально дорсолатеральной префронтальной коры (дорсальной) (BA9).

Особенностью полученных данных является различие динамики ряда конвенциональных речевых зон: в *серии сравнения внутреннего проговаривания фонем и слогов с их внешним проговариванием* были активны билатеральная премоторная кора и дополнительная моторная кора (SMA) (BA6), правая первичная моторная кора (BA4), левая орбитальная часть (BA47); в серии сравнения *внешнего проговаривания фонем и слогов с внутренним проговариванием* наблюдалась активация билатеральной премоторной коры и дополнительной моторной коры (SMA) (BA6), правой орбитальной части (BA47), правой оперкулярной части (BA44), а также билатеральная активация инсулы (BA13), а в серии сравнения *внутреннего проговаривания фонем и слогов с их восприятием* активны были левая оперкулярная часть (BA44), билатеральная премоторная кора и дополнительная моторная кора (SMA) (BA6).

ГЛАВА 3. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

3.1. Выделение связанного с событием потенциала внутреннего проговаривания и условия различных инициализаций

Для того, чтобы выявить артикуляционную активность внутреннего проговаривания, было проведено исследование с регистрацией ЭМГ. Еще в трудах А.Н. Соколова отмечалось наличие артикуляционной реакции при восприятии стимулов и при их внутреннем проговаривании, однако он интерпретировал их как связь с ориентировочным рефлексом. В исследованиях А.Н. Соколова были представлены электромиограммы, которые фиксировались в моменте прослушивания испытуемыми условий задачи и числовых рядов, необходимых для запоминания. Резкие амплитудные всплески были выявлены при начальной команде, после чего амплитуды снижались до фонового уровня. А.Н. Соколов делает вывод, что первоначальные амплитудные всплески являются проявлением общей ориентировочной реакции организма, которая возникает в результате образования доминирующего очага возбуждения в слуховом анализаторе коры мозга. Усиление речедвигательных раздражений А.Н. Соколов объясняет тем, что ориентировочное торможение речедвижений осуществляется благодаря сильному напряжению мышц, которые поддерживают речевые органы в неподвижном состоянии. Однако следует отметить, что значительные речедвигательные импульсы также отмечаются в случаях, когда слушание закрепляется фиксированием или логической переработкой, наиболее отчетливо данный факт наблюдается при восприятии иностранного языка (Соколов, 1967). Наличие артикуляционной активности подтверждает факт осуществления внутреннего проговаривания, а соотношение временных реакций позволяет определить те временные интервалы, которые, вероятнее всего, связаны с внутренней речью, но уже в рамках ЭЭГ-данных.

Согласно последним исследованиям И. Берент с соавт. (Berent et al., 2023), по категоризации звуков (разделение и дифференциация звука в потоке

речи) авторы пришли к выводам о том, что при восприятии какого-либо звука, должен быть смоделирован его целостный образ (все особые признаки звука: артикуляционный уклад, участие голоса и т.д.), а не выделен конкретный признак. Авторы исследования говорят о том, что моторное моделирование звука опосредует его категоризацию. Так, если происходит нарушение в функционировании области головного мозга, отвечающей за работу круговой мышцы рта, происходит нарушение понимания всех особенностей звука, даже тех, которые не связаны с работой этой мышцы (например, участие голоса). Кроме того, И. Берент с соавт. (Berent et al., 2023) отмечают, что восприятие речевых категорий повторяет артикуляцию.

Полученные результаты ЭМГ-исследования хорошо согласуются с результатами исследований А.Н. Соколова (1967), однако интерпретация выявленных феноменов иная. Полученные данные скорее свидетельствуют не только о том, что происходит ориентировочная реакция, но и о содержательной функции артикуляторной активности во внутреннем проговаривании и восприятии, о первично-задающей функции. В нашем исследовании наибольшая выраженная активность в процессе внутреннего проговаривания была на электроде, расположенном над губой, что согласуется с исследованиями А.Н. Соколова. Схожая артикуляторная активность была получена и при восприятии фонем. Говоря о фонематическом восприятии, можно отметить участие автоматических артикуляторных движений, сопровождающих процесс восприятия фонем и слогов. Восприятие по огибающей миограмме и по ее амплитуде сходно с внутренним проговариванием. Так, временной интервал, связанный с внутренним проговариванием, начинается в промежутке с 200 мс до 600 мс от окончания стартового сигнала.

В результате нашего исследования можно сделать вывод, что разные типы инициализации будут влиять на ориентировочную реакцию, о которой писал А.Н. Соколов. Эта реакция будет отражаться в ранних компонентах ССП ЭЭГ. При анализе полученных ЭЭГ-данных было показано, что при

различных формах инициации и восприятия главным дифференциальным элементом ССП являются ранние и средние компоненты. Этот процесс может быть связан с моторной предартикуляционной активностью, и лучше всего это проявляется на латенции от 100 до 300 мс, что можно трактовать как его моторное непроизвольное усвоение и подготовку к проговариванию. Кроме того, было проведено сравнение всех вариантов инициализации (разные типы команды стимула) с проговариванием на условный стимул (в виде японских слов), значимые различия снова обнаружились во всех случаях. Так, в статье А.В. Вартанов, А.Р. Суюнчева, А.О. Шевченко (2021) было отмечено, что в поздних компонентах не было выявлено зависимости от того, визуальной или аудиальной являлась подача. Значимое различие ранних и средних компонентов вне зависимости от типа инициализации ($p < 0,001$) наблюдалось между сериями с проговариванием фонем и слогов: опираясь на эти данные, можно сделать вывод, что задающий стимул влияет на подготовку респондентом проговариваемой единицы (фонемы или слога) на внутреннем плане. Кроме того, различие в ранних компонентах может быть интерпретировано как некое влияние на предобработку, вызванное средовыми факторами, среди которых выделяются характеристики команды-стимула, то есть команды, задающие нужную букву для последующего проговаривания. Так, голос диктора способен задавать ряд характеристик, отражающихся на внутреннем плане респондента и влияющих на произношение – интонацию, скорость, ритмику и т.д.

А.Н. Соколов говорил о том, что существует ориентировочная реакция, которая проявляется в ранних реакциях. Наше исследование и исследование И. Берента с соавт. (Berent et al., 2023) говорят о том, что в ранних компонентах происходит не только ориентировочная реакция, но еще и функция артикуляционной предобработки, которая важна как для планирования артикуляции фонемы или слога, так и для её понимания. Следовательно, ориентировочная артикуляционная предобработка – важный этап в процессах восприятия фонемы и дальнейшего речепроизводства, что

отражено в схеме на рис. 41 – Интегративная модель внутреннего проговаривания и восприятия фонем и слогов (глава 3, пункт 3.5) – и на этот этап влияют разные условия инициализации.

3.2. Сравнение внутреннего проговаривания фонем и слогов с их восприятием

Восприятие – менее произвольный процесс, чем внутреннее проговаривание. Так, проговаривание связано с формированием образа на внутреннем плане, а восприятие связано с внешним аудиальным образом. Сравнивая восприятие и проговаривание, можно исследовать важную характеристику внутреннего проговаривания – произвольность. Так, проводилось сравнение в рамках *ЭЭГ-исследования*. При восприятии или проговаривании стимулов существуют различия по эффектам каналов (выраженность различий в отдельных каналах). По результатам сравнения восприятия стимула с его намеренным проговариванием на основе анализа связанных с событием потенциалов на височных электродах можно предположить, что основным различием между внутренним проговариванием и восприятием являются более ранние компоненты связанного с событием потенциала (150-200 мс). Это может свидетельствовать как о процессе запуска команды на проговаривание (которой не было при пассивном восприятии), так и об ориентировочной артикуляционной предобработке (которая описана в главе 3.1). При восприятии аудиального стимула, вероятно, возникает схожее внутреннее аудиальное представление, что и при проговаривании. На основании полученных результатов, согласно функциональному дисперсионному анализу, можно констатировать, что все экспериментальные этапы значительно отличаются от контрольной серии.

Кроме того, используя метод локализации ЭЭГ, мы показали (Вартанов, Шевченко, 2022), что в процессе восприятия в большей степени активируется зона Вернике, нежели зона Брока. Как известно, при восприятии речи звуковые сигналы поступают в первичную слуховую кору и подвергаются обработке в зоне Вернике. Для порождения речи при нормальной речи

требуется активация зоны Брока, где разрабатывается программа артикуляции, передаваемая затем в моторную кору (Kemmerer, 2015). Однако, как показали результаты, при внутреннем проговаривании (повторении фонемы) активация в сенсорной зоне Вернике оказалась существенно больше, чем в зоне Брока. В данном эксперименте воспринимались и проговаривались только отдельные фонемы и не было задачи составлять из них слова или предложения. При изучении локализации с помощью метода ЭЭГ (глава 2, пункт 2.3.1) в задаче «Проговаривание» и в задаче «Восприятие» сильно проявила себя как левая, так и правая зона супромаргинальной извилины – зона, ассоциированная с процессами фонологического анализа (Hartwigsen et al., 2010). При этом супромаргинальная извилина при сравнении компонентов ССП в большей степени активировалась при визуальной подаче стимулов. Также, сравнивая это исследование с исследованием К. Стокель и соавт. (Stoekel et al., 2009), где показана связь данной зоны с чтением стимулов и переводом зрительного образа в слуховой, можно говорить о схожести результатов. В работе и при серии восприятия, и при серии проговаривания были активны конвенциональные речевые зоны: оперкулярная часть, верхняя височная извилина, первичная зрительная кора, угловая извилина.

Для точного определения локализации использовался метод фМРТ. На основе полученных фМРТ-данных можно выявить закономерности в восприятии и во внутреннем проговаривании, что подробно описывается в исследовании А.В. Вартанов, А.О. Шевченко (2022).

Сравнение внутреннего проговаривания фонем и слогов с восприятием методом фМРТ

В обеих сериях сравнения внутреннего проговаривания фонем и слогов наблюдалась вовлеченность угловой извилины (ВА39), верхней височной извилины (ВА22), супрамаргинальной извилины (ВА40) и нижней затылочной извилины (ВА19).

При сравнении внутреннего проговаривания фонем и слогов с их восприятием можно отметить вовлеченность конвенциональных речевых зон:

зоны Брока (BA44), а также билатерально угловой извилины. Принято считать, что зона BA44 (оперкулярная часть) вовлечена в фонологические и синтаксические процессы (Ford et al., 2013; Heim et al., 2008). Результаты фМРТ исследования показывают активацию угловой извилины (BA39), которую связывают с семантическими процессами (Binder et al., 2009; Price, Mechelli, 2005; Seghier, 2013), а также с процессами чтения (Price, Mechelli, 2005; Seghier, 2013). Процессы чтения связаны с внутренним проговариванием, поэтому активация угловой извилины в парадигме нашего исследования может быть связана именно с этой функцией зоны BA39.

При сравнении фонем и слогов отмечалась вовлеченность височной области, участие данной области традиционно связывают с речевыми задачами: восприятием, узнаванием и пр. При сравнении внутреннего проговаривания фонем и слогов с восприятием была заметна вовлеченность зоны Вернике (левая и правая BA22). Традиционно левую зону BA22 связывают с процессами слуховой обработки речи, семантической обработкой (Ardila et al., 2015; Binder, 2015). Несмотря на существующую разницу между результатами активации этой зоны в пунктах 2.6 и 2.7, мы полагаем, что эти различия не значительны и зона Вернике присутствует как при исследовании с помощью языковой-коннективити, так и при исследовании с помощью групповых сравнений, а степень различия активации определяется спецификой решаемой задачи. Также отмечалась билатеральная вовлеченность супрамаргинальной извилины (BA40), которую связывают с фонологическими операциями (McDermott et al., 2003; Chou et al., 2006). В исследовании Дж. Биндера с соавт. (Binder et al., 2015), которое описывалось в главе 1, пункте 1.4.3., согласно их функциональной модели основных языковых систем (рис. 6), данная зона входит в область преартикуляторного фонологического поиска и относится к пути А, связанному с повторением. Шустер и соавт. (Shuster et al., 2005) предполагают, что роль этой зоны в речи может заключаться в соотнесении обратной связи от реального речевого действия с внутренней репрезентацией обратной связи, ожидаемой для

данного конкретного высказывания. Данный факт может также подтверждать наличие преартикуляции в процессе внутреннего проговаривания.

При внутреннем проговаривании фонем была обнаружена вовлеченность левой веретенообразной извилины (BA37). Активацию левой зоны BA37 связывают с семантической функцией языка, а также с повторением на визуальный стимул (Ardila et al., 2015б; Flowers et al., 2004; Booth et al., 2002), но в дизайне диссертационного исследования не было визуальной подачи стимула, а также исключено семантическое кодирование, поэтому мы предполагаем, что эта зона в процессах внутреннего проговаривания может быть связана с визуальным представлением или с каким-либо фонематическими процессами. Также отмечалась активация нижней теменной доли (BA7), которую связывают с визуально-моторными функциями, участием в речевых процессах (McDermott et al., 2003), участием в формировании представлений об объектах окружающего мира (Sommerhoff, Gered, 2000).

Стоит отметить, что зона нижней затылочной извилины (BA19) присутствует в пункте 2.6, но её активация не показана в пункте 2.7, что объясняется особенностями построения языкового-коннективи. С вербальным мышлением и решением вербальных задач связывают зону нижней затылочной извилины (VisAssoc) в зоне BA19 (Martynova et al., 2017), кроме того, эту область связывают с чтением и чтением по губам (Calvert et al., 2003).

При внутреннем проговаривании также можно выделить вовлеченность лобных отделов: дорсолатеральной префронтальной коры (BA9), фронтальных глазных полях (BA8). В исследовании была замечена активация левой и правой дорсолатеральной префронтальной коры (BA9), эту зону связывают с беглостью речи (Abrahams et al., 2003). Так, в исследовании С. Абрахамс и соавт. (Abrahams et al., 2003) испытуемым задавался стимул в виде звука, на который испытуемому необходимо было назвать слово. А. Мандал и соавт. (Mandal et al., 2020) в своем исследовании пишут о том, что при

повреждении области дорсолатеральной префронтальной коры происходит снижение способности к распознаванию речевых ошибок. Кроме того, эта область входит в систему самоконтроля речи. Авторы исследования отмечают, что снижение обнаружения фонологических ошибок связано с низкой беглостью речи (Mandal et al., 2020). Так мы можем говорить об участии лобных отделов коры в процессах самоконтроля.

Таким образом, был выделен ряд зон, которые, предполагаем, связаны с проговариванием и характеризуются произвольностью в отличие от восприятия.

3.3. Сравнение внешнего проговаривания и внутреннего проговаривания

Как мы писали ранее, внутреннее проговаривание связывают с внешним проговариванием (озвучиванием). При внешней речи артикуляция более выражена в отличие от внутреннего проговаривания, при котором она присутствует в менее выраженном виде. Для того, чтобы выделить специфические зоны, связанные с внутренним проговариванием, надо сравнить внешнее и внутреннее проговаривание.

Сравнение внешнего проговаривания и внутреннего проговаривания проводилось по результатам фМРТ-данных, так как ЭЭГ-исследование было бы сильно осложнено двигательными артефактами. Мы можем констатировать, что конвенциональные речевые зоны задействованы при сравнении внешней и внутренней речи, при этом они не представлены в таблице сравнений внешней и внутренней речи (Таблица 10), так как они присутствуют в каждой из них.

Во внешней речи наблюдалась активация зоны передней поясной коры (АСС), которая описывается как область подготовки и планирования, а также языковой «инициативы» в работе А. Ардилаи соавт. (Ardila et al., 2015a). Данные исследователи пишут, что эта языковая «инициатива» и моторные команды возможны только во внешней речи, в то время как в речи про себя

эта зона не проявляется (Ardila et al., 2015a). Область vACC связывают с принятием социальных решений (Lockwood, Wittmann 2018), а также с социальным познанием, включая положительное подкрепление (Rigney et al., 2018), поэтому необходимо учитывать роль vACC (вентральная передняя поясная кора) в процессах внешней речи, которая имеет большую социальную составляющую, чем внутренняя речь. В исследовании И.К. Кристоффельса и соавт. (Christoffels et al., 2007) было показано, что зона передней поясной извилины входит в сеть, обеспечивающую слуховую обратную связь во внешней речи, так же, как и ряд подкорковых областей (мозжечок, таламус, базальные ганглии). Авторы исследования предполагают, что мониторинг речи в значительной мере зависит от неспецифических для речи областей, а также что передняя поясная кора и инсула могут быть основными областями в сети слуховой обратной связи в процессах внешней речи. Полученные нами данные подтверждают участие ACC в сети мониторинга при внешней речи.

При сравнении внутреннего проговаривания фонем и слогов с их внешним проговариванием отмечалась билатеральная активация структур мозжечка. Однако стоит уточнить, что при сравнении внутреннего проговаривания фонем с их внешним проговариванием наблюдалась билатеральная активация мозжечка, в случае сравнения внутреннего проговаривания слогов с их внешним проговариванием активация была в правом мозжечке, а при сравнении внешнего проговаривания слогов с их внутренним проговариванием активность была в левом мозжечке. Известно об участии левого мозжечка в речевых операциях, однако данное исследование показывает, что в процессах внутреннего проговаривания участвуют билатеральные структуры мозжечка в отличие от существующих моделей, где описываются только левые структуры мозжечка. Также мы предполагаем, что мозжечок во внутренней речи участвует в системе обратной связи.

Стоит отметить, что при сравнении внутреннего проговаривания слогов с их внешним проговариванием была обнаружена активация правой первичной соматосенсорной коры (BA1), которую связывают с ощущениями

положения языка (Acher et al., 2012). Мы предполагаем, что данная зона связана с артикуляционной предобработкой.

Внутреннее проговаривание может обладать свойством полимодальности, которое вовлекает в процесс проговаривания зоны различных модальностей для построения внутреннего образа произносимой речевой единицы. Связано это с личным опытом человека, ассоциациями, навыками письма и чтения. Мы полагаем, что во внутреннем плане происходит естественная системная связь визуальных и аудиальных отделов, которые и формируют единый образ, даже на уровне фонематических единиц. Все участники исследования владеют навыками письма и чтения, что ассоциативно влияет на внутреннюю репрезентацию аудиальных стимулов. Помимо этого, внутреннее проговаривание может быть связано и с более сложными образами во внутреннем плане. Полагаем, что акт внутреннего проговаривания фонем и слогов способен запустить ассоциативный образ, который отражает не только аудиальные, но и системные ассоциативные процессы. При анализе полученных в ходе исследования результатов были обнаружены активации зон верхней теменной доли (ВА7) в обоих полушариях, билатерально угловой извилины (ВА39), левой нижней затылочной извилины (VisAssoc) (ВА19) и левой веретенообразной извилины (ВА37), которые присутствуют во всех сериях сравнения внутреннего проговаривания. Все эти области относятся к третичным полям коры, а зона веретенообразной извилины и угловой извилины входят в ТРО – зону перекрытия височной, теменной и затылочной коры, которая имеет наиболее сложные интегративные функции (Хомская, 2005). Мы связываем активацию этих зон в обеспечении полимодальности внутреннего образа проговариваемой единицы в связи с отсутствием внешней слуховой обратной связи, как при внешней речи.

Большая часть конвенциональных речевых зон была задействована при внутреннем проговаривании и при внешнем проговаривании. Проводя анализ сетей коннективности, стоит выделить правую первичную моторную кору

(BA4), которая была активна только в серии сравнения внутреннего проговаривания фонем и слогов с внешним проговариванием, а при сравнении внешнего проговаривания фонем и слогов с их внутренним проговариванием наблюдалась билатеральная активация инсулы (BA13).

Также существуют факторы избирательности при языковом кодировании. Известно, что участки коры головного мозга, ответственные за воспроизведение слов с разными семантико-лексическими значениями, активируют разные зоны. Например, порождение слов, обозначающих цвет, избирательно активирует область в вентральной височной доле непосредственно перед областью, вовлеченной в восприятие цвета, в то время как порождение слов, обозначающих действие, избирательно активирует область в вентральной височной доле непосредственно перед областью, вовлеченной в восприятие движения (Breznitz, 2008). Не менее важную роль во внутренней обработке речи играет обработка слуховой информации, которая происходит в ряде областей коры, расположенных в нижней части латеральной борозды и рядом с верхней височной извилиной. Первичная слуховая кора, размещенная в центральной части извилины Хешля, имеет типичный вид коры, характерный для всех первичных сенсорных областей. Эта первичная корковая область окружена рядом других корковых областей, как и в случае ранних стадий обработки зрительной и соматосенсорной коры. Так, специфическая активность мозжечковых структур может указывать на моторную избирательность, возможно, связанную с эмуляцией звукового образа внутренней речи, если предположить, что внутренняя и внешняя речь представляют собой сходный процесс, а внутренняя речь не озвучивается внешней. Роль мозжечка также можно объяснить необходимостью контролировать и подавлять внешнюю речь, хотя могут быть задействованы и другие области мозга – лобная кора и подкорковые ядра (хвостатое ядро).

В целом, активация левополушарной сети с СТГ, МФГ и мозжечковой области во время внутренней речи согласуется с многочисленными исследованиями, посвященными этой теме (Shergill et al., 2001, Alderson-Day

et al., 2016). Кроме того, можно наблюдать ряд структур, которые активируются только при внутренней речи.

3.4. Мозжечок в процессах внутреннего проговаривания

Особую роль во всех сериях экспериментов с внутренним проговариванием сыграла билатеральная активация мозжечка (Вартанов, Шевченко, 2022; Шевченко, Вартанов, 2023, Shevchenko et al., 2022)

Роль мозжечка при внутреннем проговаривании в сравнении с внешним проговариванием. При сравнении внешнего и внутреннего проговаривания специфическая активность мозжечковых структур может указывать на моторную избирательность, которая, возможно, связана с эмуляцией звукового образа внутренней речи. Роль мозжечка также можно объяснить необходимостью контролировать и подавлять внешнюю речь, хотя могут быть задействованы и другие области мозга – лобная кора и подкорковые ядра (хвостатое ядро) (Ford et al., 2013).

В исследовании Б. Алдерсон-Дэй с соавт. (Alderson-Day et al., 2016) наблюдалась активация левой области мозжечка в процессе внутренней речи, тогда как в этом исследовании активация наблюдалась в обоих полушариях мозжечка во время внутреннего проговаривания. Ссылаясь на исследование Ч.Л. Марвел и Дж.Э. Десмонд (Marvel, Desmond, 2010), можно отметить, что они связали активацию мозжечка с поддержкой вербальной рабочей памяти через связь мозжечка с моторной корой. Однако можно выдвинуть и другую гипотезу, предположив, что мозжечок играет роль в обеспечении эмуляции звукового образа внутренней речи, что необходимо в отсутствии реального звука для генерации возбуждения в специфической речевой области (области Вернике), а также построения ритмического рисунка речи. Также это согласуется с исследованием Д. Вильдгрубера с соавт. (Wildgruber et al., 2001), о котором писалось выше в главе 1, пункте 1.4.2. Ученые получили данные о том, что различные корковые и подкорковые области по-разному влияют на моторный контроль речи в зависимости от частоты повторения при

внутреннем проговаривании. Согласно их данным, мозжечок быстрее проявляет ответ в диапазоне 4,0 и 5,5 Гц.

Вероятно, мозжечок включается в модель «обратной динамики» (Ziegler, 2016), что доказывает участие мозжечка в обучении новым паттернам, оценке и компенсаторной реакции на слуховые и соматосенсорные ошибки. Вследствие этого некоторые исследования относят мозжечок к вспомогательным компонентам в модели речепроизводства. Исследования П. Индефри, В. Левелт (Indefrey, Levelt, 2004) демонстрируют, что наиболее часто происходит активация областей, вовлеченных в планирование артикуляционных движений, именно при внешнем произнесении, при этом области, задействованные в процессе слогаобразования (схема 1, глава 1, пункт 1.2), должны активироваться вне зависимости от внешнего или внутреннего произнесения. В работе П. Индефри, В. Левелт (Indefrey, Levelt, 2004) также отмечается вероятная роль вентральных билатеральных моторных зон и сенсорных зон, правой дорсальной моторной зоны, правой дополнительной моторной области (SMA), левой и правой медиальных частей мозжечка, билатерального таламуса и среднего мозга в процессе планирования и выполнения артикуляционных движений. В заданиях с внешней артикуляцией по сравнению с заданиями с внутренним проговариванием наиболее часто обнаруживаются левая вентральная прецентральная извилина (ventral precentral gyrus), двусторонние средние передние темпоральные извилины, левая фузиформная извилина, левый таламус и правый медиальный мозжечок, но, тем не менее, в этом исследовании была отмечена активация как правых, так и левых областей мозжечка в проговаривании слогов и фонем при внутреннем произнесении (Indefrey, Levelt, 2004). Кроме того, в работе И.К. Кристоффельса и соавт. (Christoffels et al., 2007) мозжечок, как и ряд других подкорковых структур (таламус, базальные ганглии), входит в сеть слуховой обратной связи при внешней речи, что также указывает на его роль моторного посредника и участия в обратной связи в речепроизводстве. Авторы исследования

(Christoffels et al., 2007) предположили, что основными областями в системе контроля во внешней речи являются инсула и передняя поясная кора. Мы предполагаем, что мозжечок в системе контроля во внутреннем плане является одной из основных структур. Полученные нами данные показывают, что при сравнении внутреннего проговаривания фонем с их внешним проговариванием наблюдается билатеральная активация мозжечка, а в случае сравнения внутреннего проговаривания слогов с их внешним проговариванием активация была в правом мозжечке. При анализе серии сравнения внешнего проговаривания слогов с их внутренним проговариванием активность была в левом мозжечке. В настоящее время известно об участии левого мозжечка в речевых операциях, однако данное исследование показывает, что в процессах внутреннего проговаривания участвуют билатеральные структуры мозжечка.

Роль мозжечка во внутреннем проговаривании в сравнении с восприятием. При сравнении внутреннего проговаривания фонем и слогов с их восприятием наблюдалась активация левого и правого мозжечка на уровне $pFWЕuncorr.<0,001$. Как написано в главе 1, пункте 1.4.2., поражения левого полушария мозжечка преимущественно ведут к нарушениям внимания и визуально-пространственных функций, в то время как поражение правого полушария мозжечка связывают с нарушениями вербальных навыков (Starowicz-Filip et al., 2021). Результаты показывают, что и левое, и правое полушария мозжечка принимают участие в вербальных процессах, что согласуется с данными П. Индефри и В. Левелт (Indefrey, Levelt, 2004), писавших о том, что и левые, и правые части мозжечка участвуют в процессах планирования и выполнения артикуляционных действий.

Таким образом, в исследовании была отмечена активация как правых, так и левых областей мозжечка в проговаривании слогов и фонем при внутреннем проговаривании. Можно сделать вывод, что вспомогательная роль мозжечка проявляется на уровне не только внешней речи, но и внутренней речи и дифференцируется на уровне фонематических единиц (фонемы, слоги

и т.д.). Ранее отмечалось (глава 1, пункт 1.4.2), что мозжечковые нарушения могут повреждать когнитивные операции, которые зависят от внутренней речи (Askermann et al., 2007). В исследовании Ж.М. Коррейя с соавт. (Correia et al., 2020) было показано, что мозжечок дифференцирует артикуляционные и фонические движения, что говорит о роли мозжечка как моторного посредника и блока обработки сенсорной обратной связи.

Полученные в этой работе результаты показывают, что левый и правый мозжечок принимают участие в процессах внутренней речи.

3.5. Интегративная модель внутреннего проговаривания и восприятия фонем и слогов

Так, анализируя последние исследования (Berent et al., 2023; Correia et al., 2020), наследие А.Н. Соколова и полученные нами результаты, а также перспективность построения ВСІ (где элементы ввода будут фонемы и слоги), приходим к выводу о становлении актуальной задачи по интеграции этих данных и создании модели внутреннего проговаривания. Для данного исследования крайне релевантными являются труды А.Н. Соколова, который тоже отмечал наличие артикуляционной реакции при слушании стимулов, однако в большей мере связывал их с ориентировочным рефлексом. А.Н. Соколов анализировал результаты электромиограммы скрытой артикуляции при слушании речи. В исследованиях А.Н. Соколова были представлены электромиограммы, которые фиксировались в моменте прослушивания испытуемыми условий задачи и числовых рядов, необходимых для запоминания. Резкие амплитудные вспышки были отмечены на сигналах, говорящих испытуемым о начале эксперимента, после чего амплитуды снижались до фонового уровня. А.Н. Соколов делает вывод, что первоначальные амплитудные вспышки являются проявлением общей ориентировочной реакции организма, которая возникает в результате образования доминирующего очага возбуждения в слуховом анализаторе коры мозга (Соколов, 1967). Кроме того, стоит отметить и исследования

Г.М. Оппенгейми Г.С. Делл (Oppenheim, Dell, 2010), которые описывались в главе 1, пункте 1.2.2. Авторы говорят о том, что артикуляция изменяет внутреннюю речь и что внутренняя речь не может быть независимой от артикуляции полностью (Oppenheim, Dell, 2010).

Исходя из полученных результатов и на основании проведенного теоретического анализа, можно представить следующую схему восприятия и проговаривания (рис. 41), которая вдохновлена моделью В. Левелт, представленной в главе 1, пункт 1.2.1 (рис. 1), и включает в себя ориентировочный артикуляционный разряд.

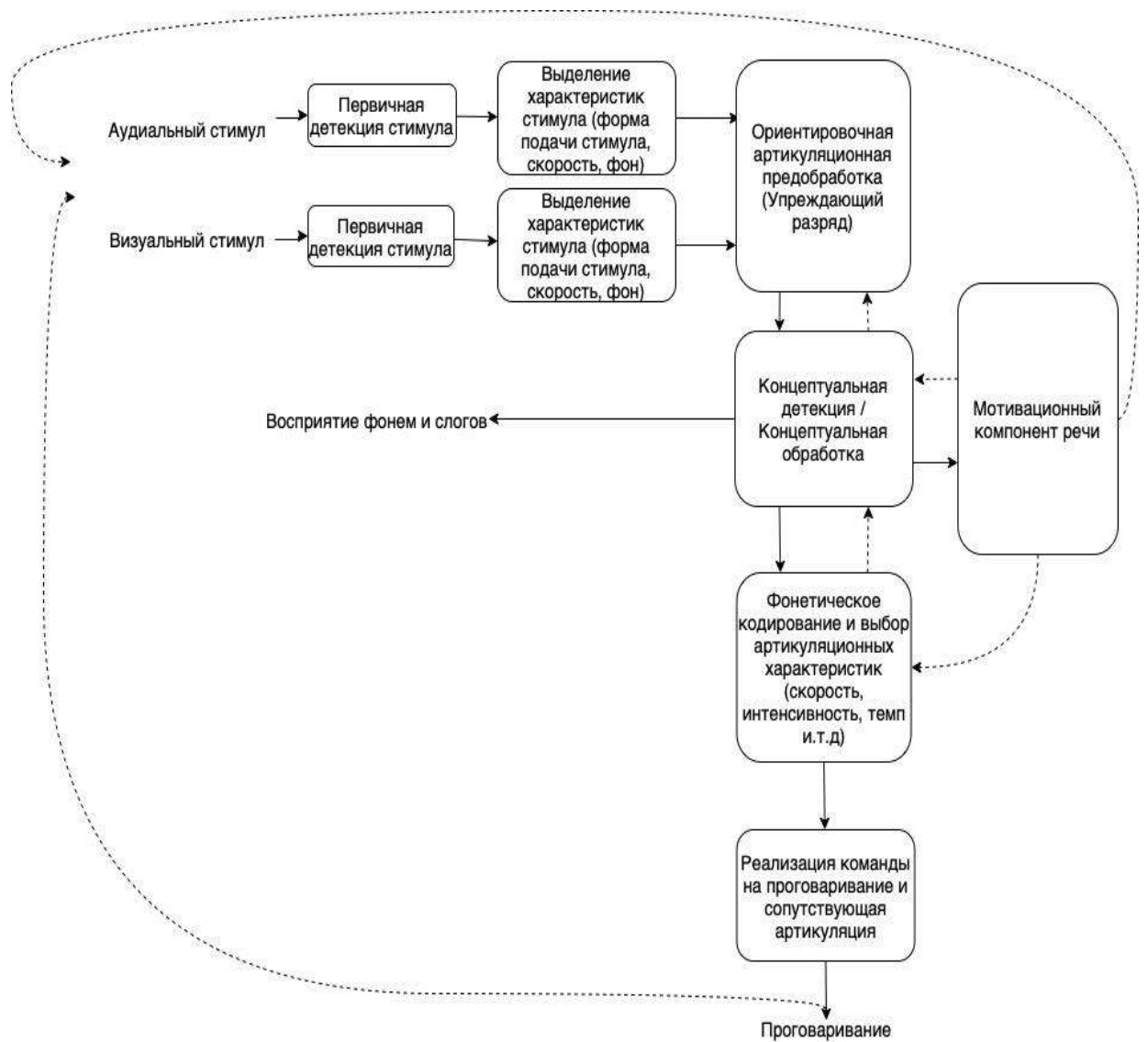


Рис. 41. Интегративная модель внутреннего проговаривания и восприятия фонем и слогов. Пунктирной линией обозначены процессы обратной афферентации

В отличие от схемы В. Левелта (Indefrey, Levelt, 2004) данная схема рассматривает план лишь на уровне фонематических единиц. В таком случае полагаем, что важным условием даже фонематического плана является **мотивационный компонент речи**, так как на этом уровне происходит оценка и планирование потенциального результата. О важности мотива в системе речепроизводства говорили многие отечественные ученые: Л.С. Выготский, А.А. Леонтьев, Т.В. Ахутина и др. Этот блок оказывает влияние как на последующий этап, так и на этап фонологического кодирования. Также стоит отметить, что мотивация влияет на качество, точность и силу при артикуляции, что было показано в исследовании Л. Сармиенто с соавт. (Sarmiento et al., 2014), где испытуемым ставилась задача активно представлять и сильно артикулировать про себя фонему, и при таком условии распознавание улучшалось. Детекция фонем и слогов, воспринимаемых на основе зрительного или слухового входа, кодируется в восприятии по-разному, что было показано в других проведенных исследованиях (Вартанов, Суюнчева, Шевченко, 2021), что согласуется с моделью В. Левелта. **Первичная детекция** стимулов происходит на основе задействованного анализатора. Далее следует **выделение характеристик стимула**, таких как форма подачи, скорость и т.д. Этап **артикуляционной предобработки** осуществляет процесс распада фонемы на составляющие (субфонемные кванты) или команды к моторному действию (кинаемы) (Плоткин, 1993) и запускает упреждающий разряд на определенные мозговые зоны. Важно отметить, что этап артикуляционной предобработки происходит *непроизвольно* и при восприятии, и при намеренном проговаривании. Однако при проговаривании артикуляция не заканчивается на артикуляционной предобработке, а задается в процессах фонологического кодирования. Это можно проследить, исходя из полученных данных ЭЭГ и ЭМГ исследований.

Следующий этап – **концептуальная детекция/концептуальная обработка**, во время этого этапа происходит сбор из отдельных перцептивных признаков единого образа языковой единицы. На данный образ влияет форма

подачи стимула, его тоновые характеристики, а также мотивация, существующая у человека в момент получения стимула. Также на этом этапе большое значение имеет память, так как собранные элементы подбираются и дифференцируются с теми, которые существуют в памяти человека, и на этом уровне определяется будущий образ проговариваемой языковой единицы.

После этапа концептуальной детекции происходит этап *фонетического кодирования*, на котором осуществляется построение программы для артикуляторных действий, связанных с будущим слуховым и моторным образом, который должен получиться в результате произнесения. Также на этом этапе закладываются характеристики к произношению фонем или слогов с учетом влияния мотивационного компонента и образа, сформированного на уровне концептуальной детекции. Так, например, голос диктора может задавать интонацию в задаче проговаривания на основе аудиального стимула. Кроме того, на этом этапе происходит сравнение полученной программы с тем образом, который сформировался в буфере концептуальной детекции для проверки точности. В случае ошибки происходит коррекция программы моторных действий.

При реализации звукового образа и сопутствующей артикуляции полученный образ попадает на блок первичной детекции стимула с целью корректировки целостного образа. При мысленном *проговаривании* (при торможении внешней речи) сигнал с уровня фонетического кодирования поступает обратно на блок концептуальной детекции, что обеспечивает процесс возникновения звукового образа фонемы или слога в отсутствие внешнего стимула. Кроме того, в схеме отражается система обратной афферентации – «процесса коррекции поведения на основе получаемой мозгом информации извне о результатах протекающей деятельности». Термин «обратная афферентация» был введен П.К. Анохиным (Энциклопедический словарь по психологии и педагогике, 2013).

3.6. Построение классификатора как область практико-ориентированных исследований внутреннего проговаривания

Исследования по построению классификаторов (Chia, 2011; Da Salla et al., 2009), подобные этому, доказывают, что в электрической активности, регистрируемой ЭЭГ, заложена информация об артикуляции и о звуковых характеристиках проговариваемых фонем. Однако в других исследованиях не было представлено ни модели, описывающей внутреннее проговаривание в подобного рода экспериментах, ни деталей артикуляционных характеристик. Кроме этого, построение классификатора является валидизацией исследований внутреннего проговаривания, так как прогрессивные математические модели способны вычленивать проговариваемые единицы, что так же, как и ЭМГ, может свидетельствовать о наличии проговаривания про себя испытуемым заданных единиц.

Средняя степень реконструируемости мало зависит от вида инициализации при усреднении. Однако на основе данных, которые представлены в таблицах 1-10, приложения 2, можно сделать вывод, что для каждого испытуемого существует свой оптимальный метод предъявления стимулов, при котором классификация происходит лучше всего. Данный факт свидетельствует об актуальности индивидуальной преднастройки под каждого испытуемого в речевых классификаторах. При том, что общая статистика реконструируемости 61-65%, индивидуальная реконструируемость фонем может превышать 68%. Также в попарных классификациях некоторые стимулы могут классифицироваться от 84 до 90% по отдельным испытуемым, что говорит о хорошей предсказательности у отдельных испытуемых по отдельным стимулам, однако, усредняя большее количество испытуемых, среднее значение точности предсказуемости уменьшается. В попарных классификациях на условный стимул (японские слова) фонем некоторые стимулы классифицируются от 76% до 83% в усреднении, что говорит о хорошей предсказательности фонем по отдельным

стимулам. Результаты классификации можно сопоставить с рядом западных работ. Существуют эксперименты с воображаемой артикуляцией (Sarmiento et al., 2014), в которых было продемонстрировано, что при представлении движений лица во время произнесения нейроны активируются в моторной коре, которая может использоваться для управления протезом. Этот факт часто используется при построении ВСИ, а также предполагает, что воображаемая артикуляция фонемы может помочь повысить точность классификации. Эксперимент по распознаванию фонем /a/ и /и/ был проведен в (Da Salla et al., 2009). Активация нейронов моторной коры вместе с внутренним произнесением гласной была классифицирована с использованием метода общей пространственной структуры и нелинейного классификатора. Полученная точность варьировалась от 68% до 78%.

Сопоставляя полученные данные, можно отметить, что и при построении классификатора методом опорных векторов, и при построении классификатора с помощью сверточных нейронных сетей лучше всего классифицировались фонемы М-Р (65% и 76.36%). Такое высокое значение в реконструкции можно объяснить тем, что обе согласные фонемы достаточно звучные, их резонирование происходит в разных местах ([P] – в ротовой области, [M] – в носовой полости). Кроме того, данный процесс задействует отличные органы артикуляции (так в артикуляции звука [P] активно участвует язык), [P] – вибрирующая фонема, требующая большего артикуляционного импульса, в отличие от фонемы [M], одного из первых согласных звуков, которым овладевает человек. Низкое значение по обеим таблицам показала пара У-М (60% и 67.55%). Такой показатель этой пары может быть вызван тем, что фонемы принадлежат к разным группам ([У] – гласный, [М] – согласный), но имеют сходные характеристики: в артикуляции фонем принимают участие губы, эти фонемы осваиваются человеком в одно время.

Сверточные нейронные сети показали свою эффективность в распознавании мысленно проговариваемых фонем. Они показали либо такой же результат, как при машинном обучении, либо превосходящий его в парной

классификации фонем. Полученная сеть из-за малого количества данных была малого размера и требовала частично ручного подбора гиперпараметров и алгоритма обучения. Такой подход к обучению сети на общих данных с последующим дообучением и множественные старты при эксперименте позволяют оптимизировать сеть (Селезнев с соавт., 2023). Так, в таблице по машинному обучению наименьшее значение наблюдается у фонем А-Р (56%), хотя в таблице по нейронной сети у данных фонем представлен высокий результат (78.14%), а наиболее высокое значение продемонстрировано в таблице по машинному обучению у фонем Г-Р, хотя в таблице по нейронной сети показатель у данных фонем представляет собой средний результат 75.60%. Кроме того, сверточные сети могут осуществлять не только попарную, но и многоклассовую классификацию (Селезнев с соавт., 2023).

ЗАКЛЮЧЕНИЕ

На основании полученных результатов были описаны психофизиологические механизмы процесса внутреннего проговаривания фонем и слогов, и представлен классификатор. Обобщение полученных результатов представлено в оригинальной модели поэтапной обработки процессов внутреннего проговаривания и восприятия. Интегративная модель основана на иерархической модели В. Левелта, теории моторного кодирования речи А.Н. Соколова и современных исследованиях (Berent et al., 2023; Correia et al., 2020). При этом в модели отражены процессы обратной связи (рис. 41), соответствующей пониманию обратной афферентации П.К. Анохина. Представленная модель описывает этапы обработки фонетических единиц и может быть использована в качестве теоретического базиса дальнейшего построения практических приложений. Для проверки наличия внешне проявляемого в подавленном мышечном компоненте процесса подготовки артикуляционного паттерна было проведено исследование с ЭМГ и выявлено наличие соответствующего моторного артикуляционного ответа по миограмме в диапазоне 200 мс для проговаривания и 150 мс для процесса восприятия. Можно допустить, что несмотря на то, что процессы восприятия и проговаривания формально являются процессами разной природы, однако подавленные моторные компоненты присутствуют и при восприятии, и при внутреннем проговаривании.

Для проверки третьей гипотезы о влиянии условий инициализации на ранние компоненты связанного с событием потенциала проводилось сравнение различных типов инициализации с целью оценки временных характеристик и различий в кодировании. По результатам исследования гипотеза была принята. Было выявлено, что наиболее дифференцирующими являются ранние и средние компоненты, что соответствует ориентировочной реакции, описанной А.Н. Соколовым (1967). На основании данных А.Н. Соколова (1984) можно сделать вывод о моторной составляющей

внутреннего проговаривания и об автоматическом моторном ответе при фонематическом восприятии. При выделении дифференцирующего компонента в ЭЭГ наиболее представленным стал компонент на временном интервале 170-300мс. Этот ответ отличался в зависимости от типа инициализации. Так, обнаружены отличия в реакции на аудиальный стимул по сравнению с визуальным, при этом реакция на «условный стимул» (подаваемый аудиально) была больше похожа на визуальный. Схожесть «условного стимула» и визуального объясняется тем, что респонденту приходится самостоятельно восстанавливать образ звука в процессе концептуальной обработки. Кроме этого, можно предположить, что при аудиальной инициализации процесс обработки происходит иначе, чем при визуальной инициализации, так как при аудиальной инициализации многие характеристики звукового образа (например, ритмика, частотные характеристики) уже имеются в отличие от визуальной инициализации, где необходимо их воссоздать самостоятельно. Однако стоит отметить, что процессы концептуальной обработки, влияющие на артикуляционный ответ, не мешают алгоритмам машинного обучения и нейронным сетям выявлять и дифференцировать паттерны. Из этого следует, что на ориентировочную реакцию влияет тип инициации, однако в самой ориентировочной реакции содержится информация о первичной детекции, концептуальной детекции и артикуляционной предобработке, что соответствует исследованию И. Берента с соавт. (Berent et al., 2023).

При проверке второй гипотезы, которая заключается в том, что мозжечок больше представлен в процессах внешнего проговаривания, были получены результаты, опровергающие её. Так, результаты нашего исследования показывают, что при сравнении внутреннего проговаривания слогов с их внешним проговариванием активация была в правом мозжечке, а при сравнении внешнего проговаривания слогов с их внутренним проговариванием активность была в левом мозжечке. Известно об участии левого мозжечка в речевых операциях, однако данное исследование

показывает, что в процессах внутреннего проговаривания участвуют билатеральные структуры мозжечка. Мозжечок участвует в процессах планирования артикуляции и фонирования (Indefrey, Levelt, 2004; Ackermann et al., 2007; Correia et al., 2020), а также в процессах контроля (Wildgruber et al., 2001; Ackermann et al., 2007). Судя по всему, мозжечок не только осуществляет сопутствующую составляющую планирования процессов внутреннего проговаривания и процессов поддержания фонематических операций, но и может осуществлять функцию эмулятора реализуемого фонетического образа, что позволяет организовать соответствующее возбуждение корковых речевых зон и обеспечивать ощущение звукового образа в отсутствии звука. Современные исследования (Correia et al., 2020) и полученные результаты говорят о мозжечке как о моторном посреднике и блоке обработки сенсорной обратной связи при внутреннем проговаривании. Предположительно мозжечок и иные структуры могут участвовать как в торможении внешней артикуляции на уровне фонематического кодирования, так и в планировании характеристик звука во внутреннем плане. Таким образом, новизной этого исследования также является выявление и описание роли мозжечка в процессах внутреннего проговаривания фонематических единиц.

Первая гипотеза, заключающаяся в том, что реализация полимодальности как особенного свойства внутреннего проговаривания, происходит благодаря другими механизмами несмотря на то, что сохраняется активность конвенциональных речевых зон при внутреннем проговаривании, принята частично. Большая часть конвенциональных речевых зон была задействована при внутреннем проговаривании и при внешнем проговаривании, однако наблюдаются определенные пики, например, правая первичная моторная кора (BA4) была активна в серии сравнения внутреннего проговаривания фонем и слогов с внешним проговариванием, а при сравнении внешнего проговаривания фонем и слогов с их внутренним проговариванием наблюдалась билатеральная активация инсулы (BA13). Эти данные были получены при анализе речевых сетей. Также для проверки первой гипотезы в

рамках фМРТ-исследования было проведено сравнение внешнего и внутреннего проговаривания. Только при внешнем проговаривании отмечалась активация зоны передней поясной извилины (BA24), часть которой связывают с социальным познанием и положительным социальным подкреплением, что может говорить о большей социальной роли внешнего проговаривания над внутренним проговариванием (Rigney et al., 2018). Кроме того, эту область выделяют как основную в сети слуховой обратной связи во внешней речи (Christoffels et al., 2007).

При проверке первой гипотезы для выделения характеристик внутреннего проговаривания было проведено сравнение внутреннего проговаривания фонем и слогов с их восприятием. При внутреннем проговаривании фонем наблюдалась активация супрамаргинальной извилины (BA40), которая была активна в сериях сравнения внутреннего проговаривания с восприятием и с внешним проговариванием. Активацию зоны связывают с преартикуляторным фонологическим поиском (Binder et al., 2015), а также с процессами обратной связи (Shuster et al., 2005), что может говорить о роли этой зоны в прертикуляции. При внутреннем проговаривании фонем и слогов отмечалась также билатеральная активация мозжечка. При сравнении внутреннего проговаривания фонем и слогов с их восприятием наблюдалась активация лобных отделов, а именно дорсолатеральной префронтальной коры (BA9) и фронтальных глазных полей (BA8), которые, по-нашему мнению, участвуют в самоконтроле.

Кроме того, при проведении исследования было отмечено, что при внутреннем проговаривании были обнаружены активации зон верхней теменной доли (BA7) в обоих полушариях, билатерально угловой извилины (BA39), левой нижней затылочной извилины (VisAssoc) (BA19) и левой веретенообразной извилины (BA37), которые являются третичными полями коры, а часть из них входит в зону ТРО. Участие этих зон во внутреннем проговаривании может говорить о том, что во внутреннем плане происходит естественная системная связь по воссозданию единого образа

проговариваемой единицы на этапе концептуальной обработки, а также может быть связано и с более сложными образами внутреннего плана.

Для проверки четвертой гипотезы проводилась разработка алгоритма на основе машинного обучения для расшифровки ЭЭГ данных о проговаривании фонем и слогов. Разработанный алгоритм на основе машинного обучения по пространственно-временным характеристикам определял фонему или слог со следующей статистикой реконструкции: общая средняя точность – 76.44% при бинарной классификации показали сверточные нейронные сети; общая средняя точность – 62% попарной классификации показал алгоритм на основе опорных векторов. Данный результат показывает, что в характеристиках ССП содержится информация о проговариваемой единице русского языка, которую можно неинвазивно дешифровать. Таким образом, принимается четвертая гипотеза нашего диссертационного исследования.

Также результаты нашей работы показывают, что внутреннее проговаривание – более сложный процесс, чем внешнее проговаривание. В теориях онтогенетического развития процессы перехода из внешней речи во внутреннюю речь сочетаются с внутренним проговариванием внешних речевых стимулов. Так, Л.С. Выготский (1982) описывал процесс свертывания эгоцентрической речи как её переход в немую, молчаливую речь. Внутреннее проговаривание является более поздним, оттого и более сложным процессом, чем внешнее проговаривание. Для наличия внутреннего проговаривания должен сформироваться внутренний голос, а с обучением чтению и письму внутреннее проговаривание является неотъемлемой частью речевой деятельности при чтении и усвоении материала. Исходя из результатов нашего исследования была принята общая гипотеза нашей работы – внутреннее проговаривание фонем и слогов обеспечивается специальной системой мозговых механизмов, связанных с мониторингом и реализацией речевого акта, которая отличается от системы реализации и мониторинга при внешнем проговаривании или восприятии фонем и слогов.

ВЫВОДЫ

1. Условия подачи стимула, определяющие содержание речевого элемента внутреннего проговаривания, влияют на ранние характеристики кодирования процесса речепроизводства, связаны с ориентировочной артикуляционной предобработкой и отражаются в ранних компонентах связанного с событием потенциала. Разные типы инициализации не влияют на возможность реконструкции с помощью машинного обучения проговариваемых про себя фонем и слогов, что может говорить о схожести поздних компонентов.

2. Различия между внутренним проговариванием и восприятием могут быть связаны с мотивом, преартикуляцией, системами обратной связи и реализацией образа во внутреннем плане, отражаются эти отличия в активации следующих областей: лобных зон – дорсолатеральной префронтальной коры, фронтальных глазных полей, первичной соматосенсорной коры и мозжечка.

3. По данным функциональной магнитно-резонансной томографии, зоны верхней теменной доли, угловой извилины, нижней затылочной извилины и веретенообразной извилины участвуют в реализации процесса внутреннего проговаривания фонем и слогов в сравнении с внешним проговариванием. Участие этих зон во внутреннем проговаривании говорит о репрезентации образа как полимодального во внутреннем плане.

4. Обнаруживается большая вовлеченность билатеральных структур мозжечка в реализации внутреннего проговаривания фонем и слогов и в серии сравнения с восприятием, и в серии сравнения с внешней речью. Проговаривание речевых единиц различного уровня задействует разные участки мозжечка, что указывает на роль мозжечка в процессах кодирования внутреннего проговаривания.

5. По результатам функционального магнитно-резонансного томографического исследования выявлено, что область передней поясной коры активируется только во внешнем проговаривании в сравнении с

внутренним проговариванием, что говорит о её роли в сети слуховой обратной связи и о большей роли во внешнем проговаривании.

6. Наиболее эффективным инструментом для неинвазивной классификации проговаривания фонем являются сверточные нейронные сети: общая средняя точность составила 76,44% при бинарной классификации. Высокую точность в реконструкции (при бинарной классификации) можно объяснить тем, что у хорошо различаемых фонем наблюдаются сильные различия в артикуляционных характеристиках.

7. Разработана интегративная модель восприятия и внутреннего проговаривания фонем и слогов на основе модели В. Левелта и моторной концепции А.Н. Соколова. В отличие от них, предложенная модель описывает процесс речепроизводства на уровне минимальных языковых единиц – фонем и слогов – в процессе внутреннего проговаривания, а также включает важный блок артикуляционной предобработки, которая происходит произвольно и при восприятии, и при намеренном проговаривании. При дальнейшем построении программы проговаривания артикуляционная предобработка оказывает влияние на процесс фонологического кодирования.

СПИСОК СОКРАЩЕНИЙ

фМРТ– функциональная магнитно-резонансная томография

МРТ – магнитно-резонансная томография

ЭЭГ – электроэнцефалография

ЭВМ – электронная вычислительная машина

ЭМГ – электромиография

ВП – вызванный потенциал

ССП – связанный с событием потенциал

BOLD-эффект (blood-oxygenation-level dependent) – различие МР-сигнала на изображениях в зависимости от процентного содержания дезоксигемоглобина

MNI (Montreal Neurological Institute) – стандартизированная 3D-система координат человеческого мозга

ВА – поля по Бродманну

ВСИ / ИМК – Интерфейс мозг–компьютер

мс – миллисекунда

Гц – герц

FAT– лобно-косой тракт

SMA (supplementary motor area) – правая дополнительная моторная область

SLF (superior longitudinal fasciculus) – верхний продольный пучок

AF (arcuate fasciculus) – аркуатный пучок

IFOF (inferior fronto-occipital fasciculus) – нижний лобно-затылочный пучок

IPL (Inferior Parietal Lobule) – нижняя теменная доля

SVM (Support Vector Machines) – метод опорных векторов

SMG (supramarginal gyrus)– супрамаргинальная извилина

STG (superior temporal gyrus) – верхняя височная извилина

ПСК – первичная слуховая кора

FCBF (Fast Correlation Based Filter for Feature Selection) – фильтр, основанный на быстрой корреляции

AR коэффициенты (autoregressive) – авторегрессионные коэффициенты

Beam forming – формирование диаграммы направленности

СПИСОК ЛИТЕРАТУРЫ

1. Азимов, Э.Г. Новый словарь методических терминов и понятий (теория и практика обучения языкам) / Э.Г. Азимов, А.Н. Щукин // М.: Издательство ИКАР. – 2009. – С.41-44
2. Ананьев, Б.Г. К теории внутренней речи в психологии / Б.Г. Ананьев // М.: АПН РСФСР. – 1960.
3. Ахутина, Т.В. Модель порождения речи школы Л.С. Выготского: основы и верификация / Т.В. Ахутина // Ученые записки Казанского университета. Серия: Гуманитарные науки. – 2022. – Т. 164, № 1-2. – С.7-27. – DOI 10.26907/2541-7738.2022.1-2.7-27.
4. Ахутина, Т.В. Нейролингвистический анализ лексики, семантики и прагматики / Т.В. Ахутина // М.: Языки славянской культуры. – 2014.
5. Бабина, Г.В. Логопедия. Дизартрия: учебно-методическое пособие / Г.В. Бабина, Л.И. Белякова, Р.Е. Идес // М.: МПГУ. – 2016.
6. Бернштейн, Н. А. О построении движений / Н.А. Бернштейн // М.: Медгиз. – 1947.
7. Блонский, П.П. Память и мышление / П.П. Блонский // Наркомпрос, Гос. науч.-исслед. ин-т психологии // М.; Л.: Гос. соц.-экон. изд-во. – 1935.
8. Блонский, П.П. Память и мышление /Блонский П.П.// Избранные психологические произведения / П.П. Блонский; по ред. А.А. Смирнов, Б.М. Теплов, В.Н. Колбановский // М.: Просвещение. – 1964.
9. Бодуэн де Куртене, И.А. Избранные труды по общему языкознанию / И.А. Бодуэн де Куртене // М.: АН СССР. – 1963. – Т. 1. – С.351–352.
10. Вартанов, А.В. Звуковой язык эмоций человека и обезьян / А.В. Вартанов, Л.В. Терещенко, А.В. Латанов, С.А. Бурлак // Национальный психологический журнал. – 2014. – №4. – С.81-94. – DOI: 10.11621/npj.2014.0409.
11. Вартанов, А.В. Механизмы внутреннего проговаривания и восприятия при разных типах внешней инициации / А.В. Вартанов, А.Р. Суюнчева, А.О. Шевченко // Интеллектуальные системы. Теория и приложения (ранее: Интеллектуальные системы по 2014, № 2, ISSN 2075-9460). – 2021. – № 25(4). – С.302-306.
12. Вартанов, А.В. Психофизиологические механизмы внутреннего проговаривания фонем / А.В. Вартанов, А.О. Шевченко // Вестник Московского университета. Серия 14: Психология. – 2022. – №1. – С.201-220. – DOI 10.11621/vsp.2022.01.09.

13. Вартанов, А.В. Субъективное восприятие и формальный компьютерный анализ звуков русского языка / А.В. Вартанов, В.В. Швырев // Психологический журнал. – 2016. – №37 (3) – С.71-87.
14. Верани, А. Роль внутренней речи в высших психических процессах / А. Верани // Культурно-историческая психология. – 2010. – № 1. – С.7-17.
15. Воробей, И. Язык и мозг. Нейробиология раскрывает главную тайну человека / И. Воробей // М.: Автор. – 2020.
16. Выготский, Л.С. Развитие высших психических функций: из неопубликованных трудов / Л.С. Выготский // М.: Издательство Академии педагогических наук РСФСР. – 1966.
17. Выготский, Л.С. Собрание сочинений в 6 т.: Т. 2: Проблемы общей психологии / под ред. В.В. Давыдова // М.: Педагогика. – 1982. – С.322-324.
18. Гавриленко, Ю.Ю. Обзор методов распознавания внутреннего проговаривания на основе данных электроэнцефалограммы / Ю.Ю. Гавриленко, Д.Ф. Саада, А.О. Шевченко, Е.А. Ильюшин // Современные информационные технологии и ИТ-образование. – 2019. – №15 (1) – С.1-8.
19. Гальперин, П.Я. Введение в психологию / П.Я. Гальперин // М.: Университет. – 2000.
20. Ганиев, Ж.В. Современный русский язык: фонетика, графика, орфография, орфоэпия: учеб.пособие / Ж.В. Ганиев // М.: Флинта. – 2012.
21. Глухов, В.П. Основы психолингвистики: учеб.пособие для студентов педвузов / В.П. Глухов // М.: АСТ: Астрель. – 2005.
22. Горелов, И.Н. Основы психолингвистики. Учебное пособие. Третье, переработанное и дополненное издание / И.Н. Горелов, К.Ф. Седов // М.: Лабиринт. – 2001.
23. Дорот, В.Л. Толковый словарь современной компьютерной лексики / В.Л. Дорот, Ф.А. Новиков // 3-е изд. СПб: БХВ-Петербург. – 2004.
24. Дружинин, В.Н. Психология 21 века / под общ. ред. В.Н. Дружинин // М.: ПЕР СЭ – 2003.
25. Ерчак, Н.Т. Внутренняя речь и мышление / Н.Т. Ерчак // Адукацыя і выхаванне. – 2017. – № 1. – С.64-72.
26. Жинкин, Н.И. Механизмы речи / Н.И. Жинкин // М.: Академии педагогических наук. – 1958.
27. Иваницкий, А.М. Психофизиология сознания // Основы психофизиологии: Учебник / А.М. Иваницкий; Под общ. ред. Ю.И. Александрова // М.: ИНФРА-М. – 1997. – Гл. 11. – С.209-228.

28. Касаткин, Л.Л. Русский язык: Учеб. Для студ. высш. пед. учеб. заведений / Л.Л. Касаткин, Е.В. Клобуков, Л.П. Крысин; под ред. Л.Л. Касаткина // М.: Академия. – 2001.
29. Кейтер, Дж. Компьютеры – синтезаторы речи / Перевод с англ. Э.Я. Пастрона; Под ред. В.А. Усика // М.: Мир. – 1985.
30. Климов, Г.А. Фонема и морфема: к проблеме лингвистических единиц / Г.А. Климов // 20е изд., М.: Либроком. – 2009.
31. Князев, С.В. Современный русский литературный язык: Фонетика, орфоэпия, графика и орфография: Учебное пособие для вузов / С.В. Князев, С.К. Пожарицкая // 2-е изд., перераб. и доп. М.: Академический Проект; Гаудеамус. – 2011.
32. Колесникова, С.М. Современный русский язык в 3 т. Том 1. Фонетика. Орфография. Лексикология. Словообразование: учебник и практикум для вузов / Под ред. С.М. Колесниковой // 2-е изд., перераб. и доп. М.: Юрайт. – 2023.
33. Конорски, Ю. Интегративная деятельность мозга: пер. с англ. / пер. Б.А. Дашевский, И.И. Полетаева; ред. П.К. Анохин // М.: Мир. – 1970.
34. Лурия, А.Р. Высшие корковые функции человека и их нарушения при локальных поражениях мозга / А.Р. Лурия // М.: Изд-во Моск. ун-та. – 1969.
35. Лурия, А.Р. Очерки психофизиологии письма / А.Р. Лурия // М.: АПН РСФСР. – 1950.
36. Лурия, А.Р. Письмо и речь: нейролингвистические исследования: Учеб. Пособие для студ. психол. фак. высш. учеб. заведений / А.Р. Лурия // М.: Академия. – 2002.
37. Морковина, А.Д. Анализ психофизиологических механизмов и подходов в коррекции звукопроизношения / А.Д. Морковина, А.О. Шевченко, В.В. Строганова, А.В. Вартанов // Национальный психологический журнал. – 2023. – № 1. – С.77-87. – DOI 10.11621/npj.2023.0107.
38. Плоткин, В.Я. Фонологические кванты. Монография / В.Я. Плоткин // М.: Наука. – 1993.
39. Розенталь, Д.Э. Словарь-справочник лингвистических терминов / Д.Э. Розенталь, М.А. Теленкова // Изд. 2-е. М.: Просвещение. – 1976.
40. Селезнев, Л.Е. Распознавание мысленно произносимых фонем с использованием сверточных нейронных сетей на основании данных электроэнцефалографии / Л.Е. Селезнев, А.А. Чупахин, В.А. Костенко, А.В. Вартанов, А.О. Шевченко // XXIV Международная научно-техническая конференция Нейроинформатика-2022. Сборник научных трудов. МФТИ Москва. – С.454-462.

41. Селиверстов, В.И. Понятийно-терминологический словарь логопеда / Под ред. В.И. Селиверстова // М.: Гуманитарный издательский центр ВЛАДОС. – 1997.
42. Соколов, А.Н. Внутренняя речь и мышление / А.Н. Соколов // М.: Просвещение. – 1967.
43. Соколов, А.Н. Внутренняя речь и мышление / А.Н. Соколов // М.: Рус. яз. – 1984.
44. Соколов, А.Н. Психофизиологическое исследование внутренней речи как механизма мышления / А.Н. Соколов // Проблемы общей, возрастной и педагогической психологии. М.: Педагогика. – 1978.
45. Соколов, Е.Н. Теоретическая психофизиология / Е.Н. Соколов // М. – 1986.
46. Спиридонов, В.Ф. Психологические модели языковой системы билингва / В.Ф. Спиридонов, Э.В. Эзрина // Шаги/Steps. – 2015. – С.1-17.
47. Теплов, Б.М. Психология музыкальных способностей / Б.М. Теплов // М.: АПН РСФСР. – 1947.
48. Ушакова, Т.Н. Природа осмысленности слова / Т.Н. Ушакова // Москва: Институт психологии РАН. – 2019.
49. Филичева, Т.Б. Основы логопедии: Учеб. пособие для студентов пед. ин-тов по спец. «Педагогика и психология (дошк.)» / Т. Б. Филичева, Н. А. Чевелева, Г. В. Чиркина // М.: Просвещение. – 1989.
50. Хомская, Е.Д. Нейропсихология: 4-е издание / Е.Д. Хомская // СПб.: Питер. – 2005. – С.73
51. Чистович, Л.А. Физиология речи. Восприятие речи человеком / Л.А. Чистович, А.В. Венцов, М.П. Гранстрем и др. // Л.: «Наука». – 1976.
52. Шевченко, А.О. Временное отражение артикуляционных компонентов вызванного потенциала внутреннего и внешнего проговаривания / А.О. Шевченко, А.В. Вартанов // Журнал «Современная наука: актуальные проблемы теории и практики». Серия «Познание». – 2023. – №1. – С.108-111
53. Шевченко, А.О. Сравнение механизмов фонематического восприятия и внутреннего проговаривания фонем и слогов. ЭЭГ и фМРТ исследование / А.О. Шевченко, А.В. Вартанов // Российский психологический журнал. – 2022. – 19(4). – С.186–210. – DOI 10.21702/rpj.2022.4.13.
54. Энциклопедический словарь по психологии и педагогике. 2013. [Электронный ресурс] // Сайт «Академик». URL: https://psychology_pedagogy.academic.ru/11384/обратная_афферентация
55. Abrahams, S. Functional magnetic resonance imaging of verbal fluency and confrontation naming using compressed image acquisition to permit overt responses/ S. Abrahams, L.H. Goldstein, A. Simmons, M.J. Brammer, S.C.

Williams, V.P. Giampietro, C.M. Andrew, P.N. Leigh // *Hum Brain Mapp.* – 2003. – Vol. 20. – №1. – P.29-40. – DOI 10.1002/hbm.

56. Acher, A. Cortical networks involved in speech recovery after glossectomy: Preliminary results of an fMRI study / A. Acher, M. Sato, L. Lamalle, C.E. Vilain., A.Krainik et al. // *NCM 2012 — 22nd Annual Meeting of the Society for Neural Control of Movement.* – 2012. – P.7-8.

57. Ackermann, H. The contribution of the cerebellum to speech production and speech perception: Clinical and functional imaging data / H. Ackermann, K. Mathiak, A. Riecker // *The Cerebellum.* – 2007. – Vol.6 – № 3 – P.202–213. – DOI 10.1080/14734220701266742.

58. Al Saleh, M.M. Brain-computer interface technology for speech recognition: A review / M.M. Al Saleh, M. Arvaneh, H. Christensen, R.K. Moore // *Proceedings of the 2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA).* – 2016. – P.1-5. – DOI 10.1109/APSIPA.2016.7820826.

59. Alderson-Day, B. The brain's conversation with itself: neural substrates of dialogic inner speech / B. Alderson-Day, S. Weis, S. McCarthy-Jones, P. Moseley, D. Smailes, C. Fernyhough // *Social Cognitive and Affective Neuroscience.* – 2016. – Vol. 11 – №1. – P.110–120. – DOI 10.1093/scan/nsv094.

60. Angrick, M. Speech synthesis from ECoG using densely connected 3D convolutional neural networks / M. Angrick, C. Herff, E. Mugler, M.C. Tate, M.W. Slutzky, D.J. Krusienski, T. Schultz // *Journal of Neural Engineering.* – 2019. – Vol. 16 – №3 – P.1-15.

61. Arbib, M.A. Neural expectations: A possible evolutionary path from manual skills to language / M.A. Arbib, G. Rizzolatti // *Communication and Cognition.* – 1996. – Vol.29 – № 3-4. – P.393-424.

62. Ardila, A. How Localized are Language Brain Areas? A Review of Brodmann Areas Involvement in Oral Language / A. Ardila, B. Bernal, M. Rosselli // *Archives of Clinical Neuropsychology.* – 2015a. – Vol.31. – №1. – P.112–122. – DOI 10.1093/arclin/acv081.

63. Ardila, A. Language and Visual Perception Associations: Meta-Analytic Connectivity Modeling of Brodmann Area 37 / A. Ardila, B. Bernal, M. Rosselli // *Behavioural neurology.* – 2015b. – Vol.2015. – P.1-14. – DOI 10.1155/2015/565871.

64. Badre, D. Left ventrolateral prefrontal cortex and the cognitive control of memory / D. Badre, A.D. Wagner // *Neuropsychologia.* – 2007. – Vol.45 – №13. – P.2883-2901. – DOI 10.1016/j.neuropsychologia.2007.06.015.

65. Benjamini, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing / Y. Benjamini, Y. Hochberg // *Journal of the Royal Statistical Society. Series B (Methodological)*. – 1995.
66. Berent, I. Phonetic categorization relies on motor simulation, but combinatorial phonological computations are abstract / I. Berent, P.J. Fried, R.M. Theodore, D. Manning, A. Pascual-Leone // *Scientific Reports*. – 2023. – Vol.13. P.1-8. – DOI 10.1038/s41598-023-28099-w.
67. Binder, J. R. Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies / J.R. Binder, R.H. Desai, W.W. Graves, L.L. Conant // *Cerebral Cortex*. – 2009. – Vol.19. – №12. – P.2767–2796. – DOI 10.1093/cercor/bhp055.
68. Binder, J.R. The Wernicke area: Modern evidence and a reinterpretation / J.R. Binder // *Neurology*. – 2015. – Vol.85. – №24. – P.2170–2175. – DOI 10.1212/WNL.0000000000002219.
69. Bogue, R. Brain-computer interfaces: control by thought / R. Bogue // *Industrial Robot: An International Journal*. – 2010. – Vol.37. – №2. – P.126-132. – DOI 10.1108/01439911011018894.
70. Bohsali, A.A. Broca's area – Thalamic connectivity / A.A. Bohsali, W. Triplett, A. Sudhyadhom, J.M. Gullett, K. McGregor, D.B. FitzGerald, B. Crosson, T. Mareci, K. White // *Brain and Language*. – 2015. – Vol.141. – P.80–88. – DOI 10.1016/j.bandl.2014.12.001.
71. Booth, J.R. Functional anatomy of intra- and cross-modal lexical tasks / J.R. Booth, D.D. Burman, J.R. Meyer, D.R. Gitelman, T.B. Parrish, M.M. Mesulam // *NeuroImage*. – 2002. – Vol.16. – №1. – P.7–22.
72. Breznitz, Z. *Brain Research in Language* / Z. Breznitz // Springer. – 2008.
73. Brigham, K. Imagined Speech Classification with EEG Signals for Silent Communication: A Preliminary Investigation into Synthetic Telepathy / K. Brigham, B.V.K.V. Kumar. // *Proceedings of the 4th International Conference on Bioinformatics and Biomedical Engineering*. – 2010. – P.1-4. – DOI 10.1109/ICBBE.2010.5515807.
74. Causeur, D. ERP: Significance Analysis of Event-Related Potentials Data / D. Causeur, C.F. Sheu // *GRAN*. – 2019. – P.12-13.
75. Chia, X. EEG based discrimination of imagined speech phonemes / X. Chia, J.B. Hagedorn, D. Schoonover, M. D'zmura // *International Journal of Bioelectromagnetism*. – 2011. – Vol.13. – №4. – P.201-206.
76. Chou, T.-L. Developmental and skill effects on the neural correlates of semantic processing to visually presented words / T.-L. Chou, J.R. Booth, T. Bitan,

D.D. Burman, J.D. Bigio, N.E. Cone, F. Cao, D. Lu // *Human Brain Mapping*. – 2006. – Vol.27. – №11. – P.915–924. – DOI 10.1002/hbm.20231.

77. Christoffels, I.K. Neural correlates of verbal feedback processing: An fMRI study employing overt speech / I.K. Christoffels, E. Formisano, N.O. Schiller // *Human Brain Mapping*. – 2007. – Vol.28. – P.868-879. – DOI 10.1002/hbm.20315.

78. Cohen, M. Analyzing neural time series data: theory and practice / M. Cohen // MIT Press. – 2014.

79. Combaz, A. Feature Extraction and Classification of EEG Signals for Rapid P300 Mind Spelling / A. Combaz, N.V. Manyakov, N. Chumerin, J.A.K. Suyken, M.M.V. Hulle // *Proceedings of the 2009 International Conference on Machine Learning and Applications*. – 2009. – P.386-391. – DOI 10.1109/ICMLA.2009.27.

80. Correia, J.M. Phonatory and articulatory representations of speech production in cortical and subcortical fMRI responses / J.M. Correia, C. Caballero-Gaudes, S. Guediche, M. Carreiras // *Scientific Reports*. – 2020. – Vol.10. – №1. – P.1-14. – DOI 10.1038/s41598-020-61435-y.

81. Da Salla, C.S. Single-trial classification of vowel speech imagery using common spatial patterns / C.S. Da Salla, H. Kambara, M. Sato, Y. Koike // *Neural Networks*. – 2009. – Vol.22. – №9. – P.1334-1339. – DOI 10.1016/j.neunet.2009.05.008.

82. Dell, G.S. A spreading-activation theory of retrieval in sentence production / G.S. Dell // *Psychological Review*. – 1986. – Vol.93. – №3. – P.283–321. – DOI 10.1037/0033-295X.93.3.283.

83. Denby, B. Silent Speech Interfaces / B. Denby, T. Schultz, K. Honda, T. Hueber, J.M. Gilbert, J.S. Brumberg // *Speech Communication*. – 2009. – Vol.52. – №4. – P.270-287. – DOI 10.1016/j. specom.2009.08.002.

84. Dodge, R. Die motorischen Wortvorstellungen / R. Dodge // *Fisical Book*. – 1896.

85. Dufor, O. Left premotor cortex and allophonic speech perception in dyslexia: A PET study / O. Dufor, W. Serniclaes, L. Sprenger-Charolles, J.-F. Démonet // *NeuroImage*. – 2009. – Vol.46. – №1. – P.241–248. – DOI 10.1016/j.neuroimage.2009.01.035.

86. Flowers, D. Attention to single letters activates left extrastriate cortex / D. Flowers, K. Jones, K. Noble, J. VanMeter, T. Zeffiro, F. Wood, G. Eden // *NeuroImage*. – 2004. – Vol.21. – №3. – P.829–839. – DOI 10.1016/j.neuroimage.2003.10.002.

87. Ford, A.A. Broca's area and its striatal and thalamic connections: a diffusion-MRI tractography study / A.A. Ford, W. Triplett, A. Sudhyadhom, J.

Gullett, K. McGregor, D.B. FitzGerald, B. Crosson, T. Mareci, K. White // *Frontiers in Neuroanatomy*. – 2013. – Vol.7. – P.1-12. – DOI 10.3389/fnana.2013.00008.

88. Friederici, A.D. The language network / A.D. Friederici, S.M. Gierhan // *Current Opinion in Neurobiology*. – 2013. – Vol.23. – №2. – P.250–254. – DOI 10.1016/j.conb.2012.10.002.

89. Frikha, T. Source Localization of EEG Brainwaves Activities via Mother Wavelets Families for SWT Decomposition / T. Frikha, N. Abdennour, F. Chaabane, O. Ghorbel, R. Ayedi, O.R. Shahin, O. Cheikhrouhou // *Journal of healthcare engineering*. – 2021. – Vol.2021. – P.1-11. – DOI 10.1155/2021/9938646.

90. Fujii, M. Neural Basis of Language: An Overview of An Evolving Model / M. Fujii, S. Maesawa, S. Ishiai, K. Iwami, M. Futamura, K. Saito // *Neurologia Medico-Chirurgica*. – 2016. – Vol.56. - №7. – P. 379–386. – DOI 10.2176/nmc.ra.2016-0014.

91. Garn, C.L. An fMRI study of sex differences in brain activation during object naming / C.L. Garn, M.D. Allen, J.D. Larsen // *Cortex*. – 2009. – Vol.45. – №5. – P.610–618. – DOI 10.1016/j.cortex.2008.02.004.

92. Gavrilenko, Y. The electroencephalogram based classification of internally pronounced phonemes / Y. Gavrilenko, D. Saada, E. Ilyushin, A. Vartanov, A. Shevchenko // *Advances in Intelligent Systems and Computing*. – 2021. – Vol.1310. – P.335-341.

93. GIFT (Group ICA Of fMRI Toolbox) [Электронный ресурс]: *TRENDS. Translational Research in Neuroimaging & Data Science*, 2017. URL: <https://trendscenter.org/software/gift/> (Дата обращения: 12.09.2022).

94. Hartwigsen, G. Phonological decisions require both the left and right supramarginal gyri / G. Hartwigsen, A. Baumgaertner, C.J. Price, M. Koehnke, S. Ulmer, H.R. Siebner // *PNAS*. – 2010. – Vol.107. – №38. – P.16494-1649.

95. Hashim, N. Word-Based Classification of Imagined Speech Using EEG / N. Hashim, A. Ali, W.N. Mohd-Isa // *Computational Science and Technology. ICCST 2017. Lecture Notes in Electrical Engineering*. – 2018. – Vol.488. – P.195-204. – DOI 0.1007/978-981-10-8276-4_19.

96. Indefrey, P. The spatial and temporal signatures of word production components / P. Indefrey, W.J. Levelt // *Cognition*. – 2004. – Vol.92. – №1-2. – P.101–144. – DOI 10.1016/j.cognition.2002.06.001.

97. Kamalakkannan R. Imagined Speech Classification using EEG / R. Kamalakkannan, R. Rajkumar, R.M. Madan, D.S. Shenbaga // *Advances in Biomedical Science and Engineering*. – 2014. – Vol.1. – №2. – P.20-32.

98. Kemmerer, D. *Cognitive Neuroscience of Language* / D. Kemmerer // Psychology Press. – 2015.

99. Khan, M. Classification of myoelectric signal for sub-vocal Hindi phoneme speech recognition / M. Khan, M. Jahan // *Journal of Intelligent & Fuzzy Systems*. – 2018. – P.1–8. – DOI 10.3233/jifs-161067.
100. Kiyosawa, M. Functional neuroanatomy of visual object naming: a PET study / M. Kiyosawa, C. Inoue, T. Kawasaki, T. Tokoro, K. Ishii, M. Ohyama, Y. Soma, M. Senda // *Graefe's Archive for Clinical and Experimental Ophthalmology*. – 1996. – Vol.234. – №2. – P.110–115. – DOI 10.1007/BF00695250.
101. Legacy FINDLAB Data and Functional ROIs [Электронный ресурс]: Stanford. Greicius Lab, 2011. URL: <https://greiciuslab.stanford.edu/resourcesv> (Дата обращения: 12.08.2022)
102. Levelt, W.J.M. A theory of lexical access in speech production / W.J.M. Levelt, A. Roelofs, A.S. Meyer // *Behavioral and Brain Sciences*. – 1999. – Vol.22. – №1. – P.1-38. – DOI 10.1017/S0140525X99001776.
103. Levelt, W.J.M. Producing spoken language: A blueprint of the speaker / W.J.M. Levelt; Eds.: C.M. Brown, P. Hagoort // *The neurocognition of language*. – 1999. – P.83-122.
104. Levelt, W.J.M. Speaking: from intention to articulation / W.J.M Levelt // Cambridge: MIT Press. – 1989.
105. Lockwood, P.L. Ventral anterior cingulate cortex and social decision-making / P.L. Lockwood, M. K. Wittmann // *Neuroscience & Biobehavioral Reviews*. – 2018. – Vol.92. – P.187–191. – DOI 10.1016/j.neubiorev.2018.05.030.
106. Lopez-Larraz, E. Syllable-based speech recognition using EMG / E. Lopez-Larraz, O.M. Mozos, J.M. Antelis, J. Minguez // *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*. – 2010. – P.699-702. – DOI 10.1109/iembs.2010.5626426.
107. Malis, L.I. Action potentials in «motor» cortex evoked by peripheral nerve stimulation / L.I. Malis, K.H. Pribram, L. Kruger // *Journal of Neurophysiology*. – 1953. – Vol.16. – №2. – P.161-167.
108. Mandal, A.S. Brain structures and cognitive abilities important for the self-monitoring of speech errors / A.S. Mandal, M.E. Fama, L.M. Skipper-Kallal, A.T. DeMarco, E.H. Lacey, P.E. Turkeltaub // *Neurobiology of Language*. – 2020. – Vol.1. – №3. – P.319–338. – DOI 10.1162/nol_a_00015.
109. Martin, S. Decoding spectrotemporal features of overt and covert speech from the human cortex / S. Martin, P. Brunner, C. Holdgraf, H.G. Heinze, N.E. Crone, J. Rieger, G. Schalk, R.T. Knight, B.N. Pasley // *Frontiers in Neuroengineering*. – 2014. – Vol.7. – №14. – P.1-14.
110. Martin, S. Word pair classification during imagined speech using direct brain recordings / S. Martin, P. Brunner, I. Iturrate, J.R. Millán, G. Schalk, R.T.

Knight, B.N. Pasley // *Scientific Reports*. – 2016. – Vol.6. – №10. – P.1-12. – DOI 10.1038/srep25803.

111. Marvel, C.L. Functional topography of the cerebellum in verbal working memory / C.L. Marvel, J.E. Desmond // *Neuropsychology Review*. – 2010. – Vol.20. – №3. – P.271–279.

112. Matsumoto, M. Classification of silent speech using support vector machine and relevance vector machine / M. Matsumoto, J. Hori // *Applied Soft Computing*. – 2014. – Vol.20. – P.95-102. – DOI 10.1016/j.asoc.2013.10.023.

113. McDermott, K.B. A procedure for identifying regions preferentially activated by attention to semantic and phonological relations using functional magnetic resonance imaging / K.B. McDermott, S.E. Petersen, J.M. Watson, J.G. Ojemann // *Neuropsychologia*. – 2003. – Vol.41. – №3. – P.293–303. – DOI 10.1016/s0028-3932(02)00162-8.

114. McGettigan, C. Neural correlates of sublexical processing in phonological working memory / C. McGettigan, J.E. Warren, F. Eisner, C.R. Marshall, S.K. Scott, P. Shanmugalingam // *Journal of cognitive neuroscience*. – 2011. – Vol.23. – №4. – P.961—977.

115. Mesgarani, N. Influence of Context and Behavior on Stimulus Reconstruction from Neural Activity in Primary Auditory Cortex / N. Mesgarani, V. David, J.B. Fritz, S.A. Shamma // *Journal of Neurophysiology*. – 2009. – Vol.102. – №6. – P.3329-3339.

116. Min, B. Vowel Imagery Decoding toward Silent Speech BCI Using Extreme Learning Machine with Electroencephalogram / B. Min, J. Kim, H.J. Park, B. Lee // *BioMed Research International*. – 2016. – Vol.11. – P.1-11. – DOI 10.1155/2016/2618265.

117. Miyake, A. Inner speech as a retrieval aid for task goals / A. Miyake // *Acta Psychologica*. – 2004. – Vol.115. – №2-3. – P.123—142.

118. MNI [Электронный ресурс]: MNI. Talairach Converter with Brodmann Areas (1.4), 2020. URL: <https://bioimagesuiteweb.github.io/webapp/mni2tal.html> (Дата обращения: 12.06.2022)

119. Naci, L. The Brain's Silent Messenger: Using Selective Attention to Decode Human Thought for Brain-Based Communication / L. Naci, R. Cusack, V.Z. Jia, A.M. Owen // *Journal of Neuroscience*. – 2013. – Vol.33. – №22. – P.9385-9393. – DOI 10.1523/JNEUROSCI.5577-12.2013.

120. Nguyen, C.H. Inferring imagined speech using EEG signals: A new approach using Riemannian manifold features / C.H. Nguyen, G.K. Karavas, P. Artemiadis // *Journal of Neural Engineering*. – 2018. – Vol.15. – №1. – P. 1-16. – DOI 10.1088/1741-2552/aa8235.

121. Oppenheim, G.M. Inner speech slips exhibit lexical bias, but not the phonemic similarity effect / G.M. Oppenheim, G.S. Dell // *Cognition*. – 2008. – Vol.106. – №1. – P.528–537. – DOI 10.1016/j.cognition.2007.02.006.
122. Oppenheim G.M. Motor movement matters: The flexible abstractness of inner speech / G.M. Oppenheim, G.S. Dell // *Memory & Cognition*. – 2010. – Vol.38. – №8. – P.1147-1160. – DOI 10.3758/MC.38.8.1147
123. Parnin, C. Subvocalization — Toward Hearing the Inner Thoughts of Developers / C. Parnin // 2011 19th IEEE International Conference on Program Comprehension. – 2011. – P.197-200.
124. Pasley, B.N. Reconstructing Speech from Human Auditory Cortex / B.N. Pasley, S.V. David, N. Mesgarani, A. Flinker, S.A. Shamma, N.E. Crone, R.T. Knight, E.F. Chang // *PLoS Biology*. – 2012. – Vol.10. – №1. – P.1-13.
125. Patel, R.S. Determining hierarchical functional networks from auditory stimuli fMRI / R.S. Patel, F.D. Bowman, J.K. Rilling // *Human Brain Mapping*. – 2006. – Vol.27. – №5. – P.462–470. – DOI 10.1002/hbm.20245.
126. Petrides, M. Neuroanatomy of Language Regions of the Human Brain / M. Petrides // *Neuroengineering*. – 2016. – Vol.7. – P.12-13.
127. Petrosyan, A. Compact and interpretable architecture for speech decoding from stereotactic EEG / A. Petrosyan, A.Voskoboynikov, A. Ossadtchi // 2021 Third International Conference Neurotechnologies and Neurointerfaces (CNN). – 2021. – P.1-4.
128. Porbadnigk, A. EEG-based Speech Recognition - Impact of Temporal Effects / A. Porbadnigk, M. Wester, J. Calliess, T. Schultz // *Proceedings of the International Conference on Bio-inspired Systems and Signal Processing*. – 2009. – Vol.1. – P.376-381. – DOI 10.5220/0001554303760381.
129. Price, C. J. Reading and reading disturbance / C.J. Price, A. Mechelli // *Current Opinion in Neurobiology*. – 2005. – Vol.15. – №2. – P.231–238. – DOI 10.1016/j.conb.2005.03.003.
130. Qureshi, M.N.I. Multiclass Classification of Word Imagination Speech With Hybrid Connectivity Features / M.N.I. Qureshi, B. Min, H. Park, D. Cho, W. Choi, B. Lee // *IEEE Transactions on Biomedical Engineering*. – 2018. – Vol.65. – №10. – P.2168-2177. – DOI 10.1109/TBME.2017.2786251.
131. RCoreTeam.2019 [Электронный ресурс]: The R Project for Statistical Computing, 2019. URL: <https://www.r-project.org> (Дата обращения: 05.06.2019)
132. Rigney, A.E. The functional role of ventral anterior cingulate cortex in social evaluation: disentangling valence from subjectively rewarding opportunities / A.E. Rigney, J.E. Koski, J.S. Beer // *Social cognitive and affective neuroscience*. – 2018. – Vol.13. – №1. – P.14–21. – DOI 10.1093/scan/nsx132.

133. Santos, E.M. Comparison of LORETA and CSP for Brain-Computer Interface Applications / E.M. Santos, R. San-Martin, F.J. Fraga // 2021 18th International Multi-Conference on Systems, Signals & Devices (SSD). – 2021. – P.817-822. – DOI 10.1109/SSD52085.2021.9429518.

134. Sarmiento, L.C. Brain computer interface (BCI) with EEG signals for automatic vowel recognition based on articulation mode / L.C. Sarmiento, P. Lorenzana, C.J. Cortes, W.J. Arcos, J.A. Bacca, A. Tovar // Proceedings of the 5th ISSNIP-IEEE Biosignals and Biorobotics Conference (2014). – 2014. – P.1-4. – DOI 10.1109/BRC.2014.6880997.

135. Scott, M. Inner speech captures the perception of external speech / M. Scott, H.H. Yeung, B. Gick, J.F. Werker // The Journal of the Acoustical Society of America. – 2013. – Vol.133. – №4. – P.286-292. – DOI 10.1121/1.4794932.

136. Seghier, M.L. The Angular Gyrus: Multiple Functions and Multiple Subdivisions / M.L. Seghier // The Neuroscientist. – 2013. – Vol.19. – №1. – P.43-61. – DOI 10.1177/1073858412440596.

1. Seleznyev, L. Recognition of Mentally Pronounced Russian Phonemes Using Convolutional Neural Networks and Electroencephalography Data / Seleznyev L., Chupakhin A., Kostenko V., Shevchenko A., Vartanov A. // Optical Memory and Neural Networks. – 2023. – Vol. 32. – №2 – P.73-85. – DOI 10.3103/S1060992X23020066.

137. Sereshkeh, A.R. Online EEG Classification of Covert Speech for Brain-Computer Interfacing / A.R. Sereshkeh, R. Trott, A. Bricout, T. Chau // International Journal of Neural Systems. – 2017. – Vol.27. – №8. – P.2292-2300. – DOI 10.1142/S0129065717500332.

138. Shergill, S.S. A functional study of auditory verbal imagery / S.S. Shergill, E.T. Bullmore, M.J. Brammer, S.C. Williams, R.M. Murray, P.K. McGuire // Psychological medicine. – 2001. – Vol.31. – P.241-53. – DOI 10.1017/S003329170100335X.

139. Shevchenko, A. Inner speech and external speech: phonemes, syllables and words. fMRI study / A. Shevchenko, A. Suyuncheva, A. Vartanov, O. Bronov // Procedia Computer Science. – 2022. – Vol.213. – P.136-143. – DOI 10.1016/j.procs.2022.11.048.

140. Shuster, L. An fMRI investigation of covertly and overtly produced mono- and multisyllabic words / L. Shuster, S. Lemieux // Brain and Language. – 2005. – Vol.93. – №1. – P.20–31. – DOI 10.1016/j.bandl.2004.07.007.

141. Smith, J.D. The role of subvocalization in auditory imagery / J.D. Smith, M. Wilson, D. Reisberg // Neuropsychologia. – 1995. – Vol.33. – №11. – P.1433-1454.

142. Soldati, N. ICA analysis of fMRI with real-time constraints: an evaluation of fast detection performance as function of algorithms, parameters and a priori conditions / N. Soldati, V.D. Calhoun, L. Bruzzone, J. Jovicich // *Frontiers in Human Neuroscience*. – 2013. – Vol.7. – №19. – P.1-11. – DOI 10.3389/fnhum.2013.00019.

143. Sommerhoff, G. *Understanding Consciousness: Its Function and Brain Processes* / Sommerhoff G. // London; Thousand Oaks: SAGE Publications. – 2000. – P.135–156.

144. SPM12 [Электронный ресурс]: STATISTICAL PARAMETRIC MAPPING, 2020. URL: <https://www.fil.ion.ucl.ac.uk/spm/software/spm12/> (Дата обращения: 12.12.2022).

145. Starowicz-Filip, A. Cerebellar Functional Lateralization From the Perspective of Clinical Neuropsychology/ A. Starowicz-Filip, K. Prochwicz, J. Kłosowska, A.A. Chrobak, A. Myszk, B. Betkowska-Korpała, B. Kwinta // *Frontiers in Psychology*. – 2021. – Vol.12. – P.1-12. – DOI 10.3389/fpsyg.2021.775308.

146. Stoeckel, C. Supramarginal gyrus involvement in visual word recognition / C. Stoeckel, P.M. Gough, K.E. Watkins, J.T. Devlin // *Cortex*. – 2009. – Vol.45. – №9. – P.1091–1096.

147. Sun, P. Brain2Char: A Deep Architecture for Decoding Text from Brain Recordings / P. Sun, G.K. Anumanchipalli, E.F. Chang // *Journal of Neural Engineering*. – 2020. – Vol.17. – P.1-12.

148. Suppes, P. Brain Wave Recognition of Words / P. Suppes, Zh.-L. Lu, B. Han // *Proceedings of the National Academy of Sciences*. – 1997. – Vol.94. – №26. – P.14965-14969. – DOI 10.1073/pnas.94.26.14965.

149. Suyuncheva, A. Reconstruction of words, syllables and phonemes of internal speech by EEG activity / A. Suyuncheva, D. Saada, Y. Gavrilenko, A. Schevchenko, A. Vartanov, E. Ilyushin // *Advances in Intelligent Systems and Computing*. – 2021. – Vol.1358. – P.1-9.

150. Takeichi, H. Comprehension of degraded speech sounds with m-sequence modulation: An fMRI study / H. Takeichi, S. Koyama, A. Terao, F. Takeuchi, Y. Toyosawa, H. Murohashi // *NeuroImage*. – 2010. – Vol.49. – №3. – P.2697–2706. – DOI 10.1016/j.neuroimage.2009.10.063.

151. Vartanov, A.V. A new method of localizing brain activity using the scalp eegdata / Vartanov A.V. // *Procedia computer science*. – 2022. – Vol.213. – P.41–48. – DOI 10.1016/j.procs.2022.11.036.

152. Whitney, C. Task-dependent Modulations of Prefrontal and Hippocampal Activity during Intrinsic Word Production / C. Whitney, S. Weis, T.

Krings, W. Huber, M. Grossman, T. Kircher // *Journal of Cognitive Neuroscience*. – 2009. – Vol.21. – №4. – P.697–712. – DOI 10.1162/jocn.2009.21056.

153. Wildgruber, D. Differential Contributions of Motor Cortex, Basal Ganglia, and Cerebellum to Speech Motor Control: Effects of Syllable Repetition Rate Evaluated by fMRI / D. Wildgruber, H. Ackermann, W. Grodd // *NeuroImage*. – 2001. – Vol.13. – №1. – P.101–109. – DOI 10.1006/nimg.2000.0672.

154. Yamaguchi, H. Decoding Silent Speech in Japanese from Single Trial EEGs: Preliminary Results / H. Yamaguchi, T. Yamazaki, K. Yamamoto, S. Ueno, A. Yamaguchi, T. Ito, S. Hirose, K. Kamijo, H. Takayanagi, T. Yamanoi, S. Fukuzumi // *Journal of Computer Science & Systems Biology*. – 2015. – Vol.8. – №5. – P.285-291. – DOI 10.4172/jcsb.1000202.

155. Yoo, S.S. Neural substrates associated with the concurrent performance of dual working memory tasks / S.S. Yoo, G. Paralkar, L.P. Panych // *International Journal of Neuroscience*. – 2004. – Vol.114. – №6. – P.613-631. – DOI 10.1080/00207450490430561.

156. Zhong, A.J. The unique role of the frontal aslant tract in speech and language processing / A.J. Zhong, J.V. Baldo, N.F. Dronkers, M.V. Ivanova // *NeuroImage: Clinical*. – 2022. – Vol.34. – P.1-11. – DOI 10.1016/j.nicl.2022.103020.

157. Ziegler W. The Phonetic Cerebellum / W. Ziegler // *The Linguistic Cerebellum*. – 2016. – P.1–32. – DOI 10.1016/B978-0-12-801608-4.00001-3.

ПРИЛОЖЕНИЕ 1. СХЕМА РАСПОЛОЖЕНИЯ ЭЛЕКТРОДОВ ЭЛЕКТРОМИОГРАФИИ

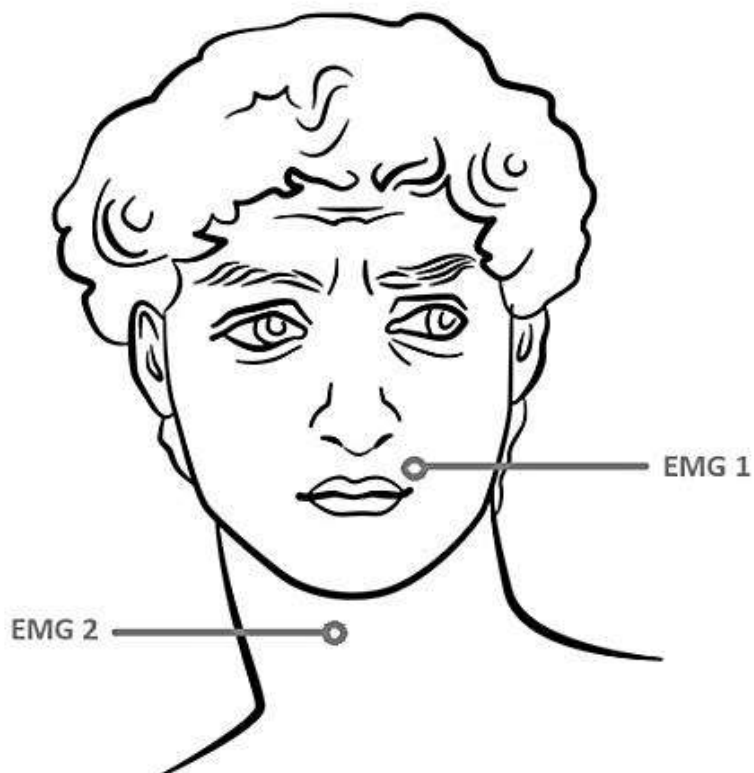


Рис. 1. Схема расположение электродов для записи ЭМГ, описанных в главе 2 пункта 2.1.

ПРИЛОЖЕНИЕ 2. ПРИМЕРЫ КЛАССИФИКАЦИИ

Таблицы точности попарной классификаций внутреннего проговаривания фонем и слогов во время звукового и визуального предъявления стимулов для каждого из испытуемых, полученные методом SVM с сигмоидным ядром, для построения классификатора, подробно описанного в главе 2 пункта 2.2.3.

Таблица 1. Пример попарной классификации фонем (аудиально) с помощью SVM, по испытуемому 1

	Б	Ф	Г	М	Р	У
А	62%	65%	59%	53%	66%	52%
Б		60%	68%	52%	69%	47%
Ф			59%	51%	70%	60%
Г				68%	72%	59%
М					72%	40%
Р						60%

Таблица 2. Пример попарной классификации фонем (аудиально) с помощью SVM, по испытуемому 2

	Б	Ф	Г	М	Р	У
А	60%	54%	70%	53%	39%	58%
Б		67%	46%	67%	52%	67%
Ф			56%	60%	63%	69%
Г				69%	75%	63%
М					65%	60%
Р						71%

Таблица 3. Пример попарной классификации фонем (визуально) с помощью SVM, по испытуемому 1

	Б	Ф	Г	М	Р	У
А	55%	60%	56%	63%	59%	56%
Б		75%	60%	58%	66%	66%
Ф			66%	65%	54%	59%
Г				69%	74%	58%
М					67%	49%
Р						56%

Таблица 7. Пример попарной классификации слогов(визуально) с помощью SVM, по испытуемому 1

	ФА	ГА	МА	РА	БУ	РУ	МУ	ФУ	ГУ
БА	77%	70%	52%	43%	68%	43%	49%	60%	56%
ФА		70%	67%	60%	60%	64%	81%	63%	67%
ГА			65%	67%	57%	51%	67%	52%	63%
МА				54%	64%	52%	65%	49%	56%
РА					57%	62%	44%	52%	69%
БУ						56%	57%	55%	63%
РУ							45%	47%	64%
МУ								52%	57%
ФУ									55%

Таблица 8. Пример попарной классификации слогов (визуально) с помощью SVM, по испытуемому 2

	ФА	ГА	МА	РА	БУ	РУ	МУ	ФУ	ГУ
БА	36%	55%	84%	65%	61%	48%	45%	55%	60%
ФА		55%	62%	55%	43%	58%	18%	55%	60%
ГА			69%	50%	90%	23%	65%	34%	79%
МА				56%	74%	67%	67%	56%	56%
РА					57%	47%	55%	45%	84%
БУ						60%	65%	67%	57%
РУ							58%	47%	53%
МУ								60%	55%
ФУ									56%

Таблица 9. Пример попарной классификации фонем на условный стимул с помощью SVM, по испытуемому 1

	Б	Ф	Г	М	Р	У
А	56%	56%	55%	54%	74%	58%
Б		62%	60%	67%	70%	83%
Ф			69%	59%	59%	60%
Г				67%	60%	59%
М					72%	61%
Р						68%

Таблица 10. Пример попарной классификации фонем на условный стимул с помощью SVM, по испытуемому 2

	Б	Ф	Г	М	Р	У
А	62%	43%	76%	67%	52%	47%
Б		53%	54%	62%	76%	67%
Ф			73%	71%	54%	51%
Г				72%	60%	66%
М					71%	62%
Р						65%

ПРИЛОЖЕНИЕ 3. ТАБЛИЦЫ СРАВНЕНИЯ фМРТ- СВЯЗНОСТИ СЕРИЙ

Таблица 1. Значимые фМРТ-различия в кластерах при сравнении внутреннего проговаривания фоном и с их внешним проговариванием по данным 30 участников.

Координаты представлены в пространстве MNI

Кластер	Объём кластера	T	X {mm}	Y {mm}	Z {mm}	L/R	Структура	BA
1	16983	15,85	-36	-70	47	L	Угловая извилина	39
		15,12	-51	-58	47	L	Угловая извилина	39
		10,78	-51	-61	32	L	Угловая извилина	39
2	12933	15,82	48	-61	44	R	Угловая извилина	39
		8,03	48	-64	29	R	Угловая извилина	39
		7,15	39	-46	59	R	Верхняя теменная долька (VisMotor)	7
3	15417	11,79	18	32	53	R	Фронтальные глазные поля	8
		11,00	-18	32	53	L	Фронтальные глазные поля	8
		10,14	-39	11	53	L	Премоторная кора + дополнительная моторная кора (SMA)	6
4	729	8,91	-6	-37	32	L	Вентральная задняя поясная извилина	23
		8,17	0	-31	41	R	Дорсальная задняя поясная кора	31
5	297	8,26	-48	23	38	L	Фронтальные глазные поля	8
6	432	7,67	30	-34	59	R	Первичная соматосенсорная кора	1
		7,42	36	-28	59	R	Первичная моторная кора	4
7	324	7,61	15	-88	-34	R	Мозжечок	
8	567	7,60	-42	-76	-40	L	Мозжечок	

Таблица 2. Значимые фМРТ-различия в кластерах при сравнении внешнего проговаривания фонем с их внутреннем проговариванием по данным 30 участников. Координаты представлены в пространстве MNI

Кластер	Объём кластера	T	X {mm}	Y {mm}	Z {mm}	L/R	Структура	BA
1	9747	15,2 4	57	20	-7	R	Орбитальная (глазничная) часть	47
		12,7 6	42	11	-1	R	Инсула	13
		7,77	39	20	-13	R	Орбитальная (глазничная) часть	47
2	9882	13,0 4	-48	11	-10	L	Височный полюс	38
		11,6 3	-57	11	-6	L	Верхняявисочнаяизвилина	22
		11,5 2	-42	11	-4	L	Инсула	13
3	4698	10,5 2	3	17	38	R	Фронтальные глазные поля	8
		9,00	-3	23	32	L	Дорсальная передняя поясная кора	32
		8,09	0	11	50	R	Премоторная кора + дополнительная моторная кора (SMA)	6
4	3456	9,34	-57	-37	26	L	Супрамаргинальная извилина	40
		9,21	-63	-28	17	L	Супрамаргинальная извилина	40
		9,20	-63	-13	8	L	Первичная слуховая кора	41
5	1458	9,15	60	-43	11	R	Верхняявисочнаяизвилина	22
		7,76	63	-37	26	R	Супрамаргинальная извилина	40
6	918	8,32	60	-19	5	R	Первичная слуховая кора	41
		7,91	60	-19	14	R	Супрамаргинальная извилина	40
7	783	8,00	-12	2	10	L	Хвостатое ядро	
8	189	7,65	12	2	10	R	Хвостатое ядро	
9	27	6,99	-24	41	26	L	Дорсолатеральная префронтальная кора (дорсальная)	9
10	54	6,95	12	5	71	R	Премоторная кора + дополнительная	6

							моторная кора (SMA)	
11	27	6,82	36	5	-19	R	Дорсальная энторинальная кора	34
12	27	6,76	-15	14	3	L	Хвостатое ядро	
13	27	6,52	3	34	11	R	Вентральная передняя поясная кора	24
14	54	6,50	24	14	-22	R	Орбитальная (глазничная) часть	47
15	27	6,46	33	-64	-19	R	Веретенообразная извилина	37
16	27	6,40	66	-25	32	R	Супрамаргинальная извилина	40

Таблица 3. Значимые фМРТ-различия в кластерах при сравнении внутреннего проговаривания фонем с их восприятия по данным 30 участников. Координаты представлены в пространстве MNI

Кластер	Объём кластера	T	X {mm}	Y {mm}	Z {mm}	L/R	Структура	BA
1	12258	16,38	-45	-43	47	L	Супрамаргинальная извилина	40
		12,57	-48	-40	57	L	Первичная соматосенсорная кора	1
		11,35	-42	-52	53	L	Супрамаргинальная извилина	40
	8181	10,93	39	-52	56	R	Верхняя теменная доля (VisMotor)	7
		10,76	48	-46	54	R	Супрамаргинальная извилина	40
		9,32	39	-43	47	R	Верхняя теменная доля (VisMotor)	7
	1647	10,30	-30	11	59	L	Премоторная кора + дополнительная моторная кора (SMA)	6
		8,58	-27	-1	59	L	Премоторная кора + дополнительная моторная кора (SMA)	6
	864	9,53	48	32	33	R	Дорсолатеральная префронтальная кора (дорсальная)	9
	702	8,70	-45	29	32	L	Дорсолатеральная префронтальная кора (дорсальная)	9

	864	7,61	24	8	62	R	Премоторная кора + дополнительная моторная кора (SMA)	6
		7,56	33	8	59	R	Премоторная кора + дополнительная моторная кора (SMA)	6
		7,12	39	8	47	R	Фронтальные глазные поля	8
	351	7,33	-45	8	35	L	Фронтальные глазные поля	8
		6,80	-51	14	35	L	Фронтальные глазные поля	8

Таблица 4. Значимые фМРТ-различия в кластерах при сравнении внутреннего проговаривания слогов с их внешним проговариванием по данным 30 участников. Координаты представлены в пространстве MNI

Кластер	Объём кластера	T	X {mm}	Y {mm}	Z {mm}	L/R	Структура	BA
	25056	16,69	-39	-64	47	L	Угловая извилина	39
		15,54	-39	-73	44	L	Угловая извилина	39
		15,49	-42	-61	35	L	Угловая извилина	39
	16308	11,36	-33	17	56	L	Фронтальные глазные поля	8
		11,05	-42	17	47	L	Фронтальные глазные поля	8
		10,89	-12	41	47	L	Фронтальные глазные поля	8
	1404	9,50	-6	-40	35	L	Вентральная задняя поясная кора	23
	1593	8,91	-45	53	-4	L	Передняя префронтальная кора	10
		8,30	-42	44	-13	L	Орбитальная (глазничная) часть	47
	1512	8,86	48	-61	35	R	Угловая извилина	39
	540	8,53	-12	56	5	L	Передняя префронтальная кора	10
		8,25	-6	59	-1	L	Передняя префронтальная кора	10
	1566	8,35	-6	-58	26	L	Дорсальная задняя поясная кора	31
		7,97	0	-61	41	L/R	Верхняя теменная долька (VisMotor)	7

		7,90	3	-61	29	R	Дорсальная задняя поясная кора	31
	189	8,29	12	-76	-28	R	Мозжечок	
	189	7,23	-63	-13	-19	L	Средняя височная извилина	21
	135	6,93	-57	-46	-10	L	Веретенообразная извилина	37
	108	6,73	45	-67	-34	R	Мозжечок	
	81	6,66	39	-73	-46	R	Мозжечок	
	54	6,61	-51	-58	-13	L	Веретенообразная извилина	37
	54	6,56	-6	-49	17	L	Вентральная задняя поясная кора	23
	27	6,53	36	-70	-43	R	Мозжечок	
	27	6,47	-66	-52	-13	L	Веретенообразная извилина	37
	27	6,41	33	14	59	R	Фронтальные глазные поля	8
	27	6,33	6	59	-4	R	Передняя префронтальная кора	10

Таблица 5. Значимые фМРТ-различия в кластерах при сравнении внешнего проговаривания слогов с их внутренним проговариванием по данным 30 участников. Координаты представлены в пространстве MNI

Кластер	Объём кластера	T	X {mm}	Y {mm}	Z {mm}	L/R	Структура	BA
1	13230	13,21 292	42	14	-4	R	Инсула	13
		13,04 693	48	20	-10	R	Орбитальная (глазничная) часть	47
		10,10 238	54	2	2	R	Премоторная кора + дополнительная моторная кора (SMA)	6
2	9045	12,96 956	-3	20	30	L	Вентральная передняя поясная кора	24
		11,70 485	3	29	26	R	Дорсальная передняя поясная кора	32
		9,697 266	6	8	41	R	Дорсальная передняя поясная кора	32
3	10584	12,07 536	-60	-1	8	L	Премоторная кора + дополнительная	6

							моторная кора (SMA)	
		11,24 146	-60	-19	11	L	Первичная соматосенсорная кора	1
		11,20 806	-63	-19	26	L	Первичная соматосенсорная кора	1
4	7182	11,35 787	60	-37	14	R	Верхняя височная извилина	22
		8,852 713	60	-40	29	R	Супрамаргинальная извилина	40
		8,242 38	51	-16	8	R	Первичная слуховая кора	41
5	864	9,321 262	54	-31	-7	R	Средняя височная извилина	21
6	2781	8,699 866	36	41	35	R	Дорсолатеральная префронтальная кора (дорсальная)	9
		8,638 404	33	53	29	R	Оперкулярная часть	44
		8,354 835	30	47	23	R	Передняя префронтальная кора	10
7	675	8,689 092	-33	47	23	L	Передняя префронтальная кора	10
8	459	8,197 294	-36	-52	-40	L	Мозжечок	
9	54	7,770 264	27	50	14	R	Передняя префронтальная кора	10
10	243	7,631 536	42	-4	-10	R	Инсула	13
11	162	7,567 318	-36	-43	-52	L	Мозжечок	
12	243	7,426 073	-30	-58	-55	L	Мозжечок	
13	270	7,395 424	3	39	2	R	Дорсальная передняя поясная кора	32
14	189	7,325 985	-30	5	-19	L	Дорсальная энторинальная кора	34
15	189	6,881 354	48	41	5	R	Дорсолатеральная префронтальная кора (вентральная)	46
16	324	6,819 528	-42	-16	5	L	Первичная слуховая кора	41

17	27	6,505 911	36	38	5	R	Дорсолатеральная префронтальная кора (вентральная)	46
18	27	6,360 427	39	-13	-4	R	Инсула	13

Таблица 6. Значимые фМРТ-различия в кластерах при сравнении внутреннего проговаривания слогов с их восприятия по данным 30 участников. Координаты представлены в пространстве MNI

Кластер	Объём кластера	T	X {mm}	Y {mm}	Z {mm}	L/R	Структура	BA
	17307	13,27	-48	-43	50	L	Супрамаргинальная извилина	40
		12,11	-39	-52	53	L	Супрамаргинальная извилина	40
		12,01	-21	-67	56	L	Верхняя теменная доля (VisMotor)	7
	2862	9,33	36	-52	53	R	Верхняя теменная доля (VisMotor)	7
		8,51	39	-46	47	R	Верхняя теменная доля (VisMotor)	7
		8,40	30	-64	47	R	Угловая извилина	39
	3186	8,79	-48	29	26	L	Дорсолатеральная префронтальная кора (дорсальная)	9
		8,78	-45	5	38	L	Премоторная кора + дополнительная моторная кора (SMA)	6
		7,34	-45	5	23	L	Оперкулярная часть	44
	972	8,25	-45	47	-1	L	Дорсолатеральная префронтальная кора (вентральная)	46
		7,39	-45	56	2	L	Передняя префронтальная кора	10
	810	8,12	-24	2	59	L	Премоторная кора + дополнительная моторная кора (SMA)	6
		6,72	-33	11	62	L	Премоторная кора + дополнительная моторная кора (SMA)	6
	297	7,11	48	-40	58	R	Супрамаргинальная извилина	40

ПРИЛОЖЕНИЕ 4. ФМРТ ИЗОБРАЖЕНИЯ АКТИВНОСТИ ГОЛОВНОГО МОЗГА В СЕРИИ ФМРТ-СВЯЗНОСТИ

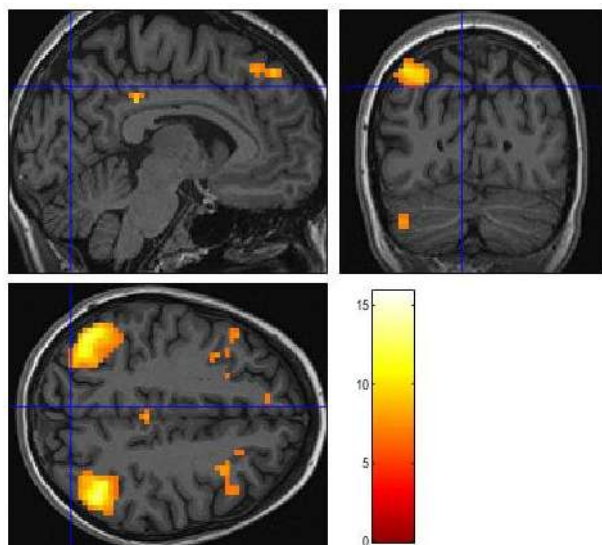


Рис. 1. ФМРТ изображение значимой разности активности зон головного мозга в серии сравнения внутреннего проговаривания фонем с их внешним проговариванием (градиентом показаны зоны, связанные с внутренним проговариванием фонем). Изображение построено по усредненным данным 30 участников

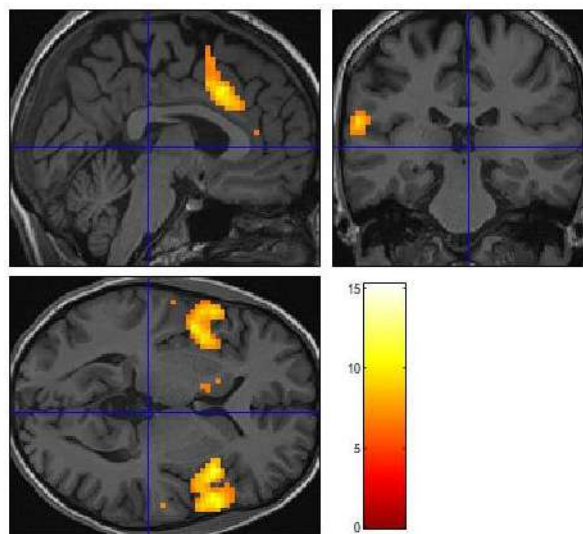


Рис. 2. ФМРТ изображение значимой разности активности зон головного мозга в серии сравнения внешнего проговаривания фонем с их внутренним проговариванием (градиентом показаны зоны, связанные с внешним проговариванием фонем). Изображение построено по усредненным данным 30 участников

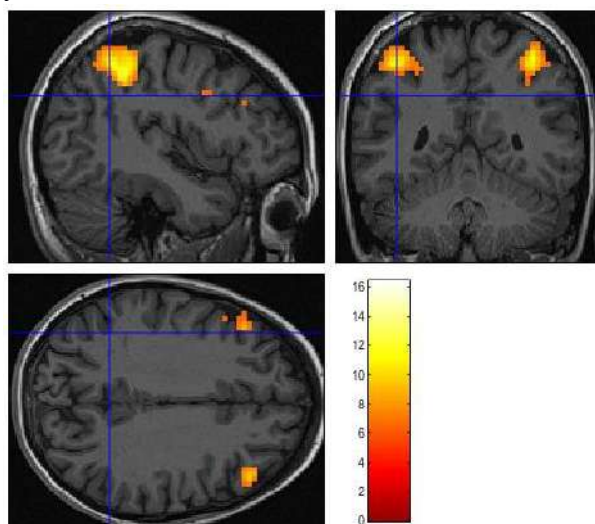


Рис. 3. ФМРТ изображение значимой разности активности зон головного мозга в серии сравнения внутреннего проговаривания фонем с их восприятием (градиентом показаны зоны, связанные с внутренним проговариванием фонем). Изображение построено по усредненным данным 30 участников

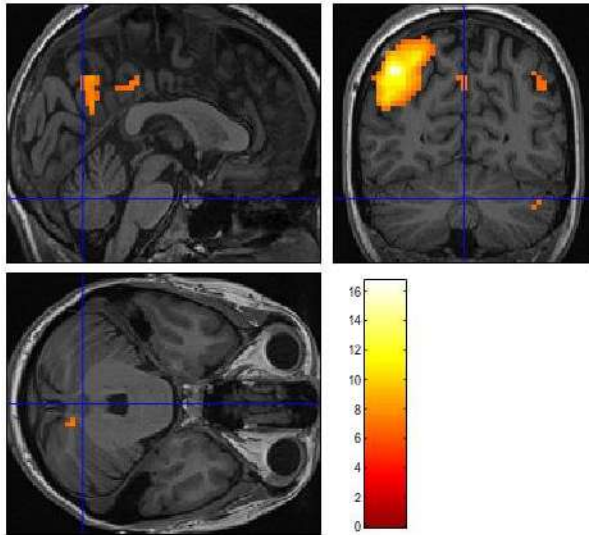


Рис. 4. ФМРТ изображение значимой разности активности зон головного мозга в серии сравнения внутреннего проговаривания слогов с их внешним проговариванием (градиентом показаны зоны, связанные с внутренним проговариванием слогов). Изображение построено по усредненным данным 30 участников

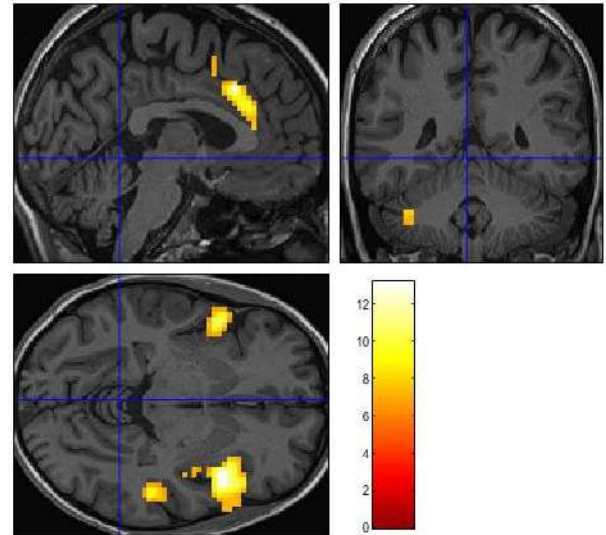


Рис. 5. ФМРТ изображение значимой разности активности зон головного мозга в серии сравнения внешнего проговаривания слогов с их внутренним проговариванием (градиентом показаны зоны, связанные с внешним проговариванием слогов). Изображение построено по усредненным данным 30 участников

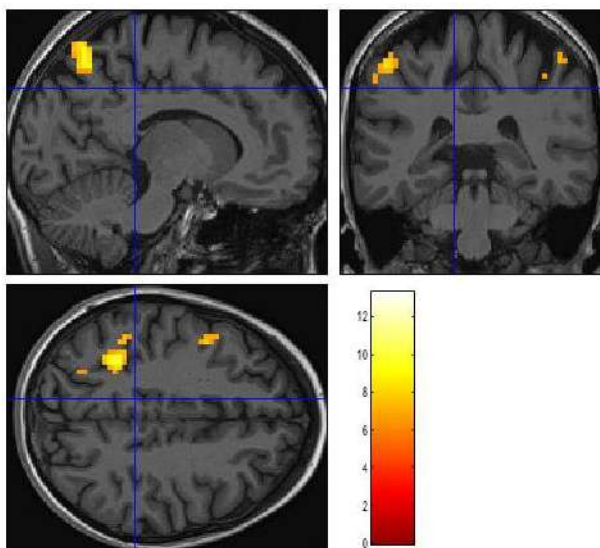


Рис. 6. ФМРТ изображение значимой разности активности зон головного мозга в серии сравнения внутреннего проговаривания слогов с их восприятием (градиентом показаны зоны, связанные с внутренним проговариванием слогов). Изображение построено по усредненным данным 30 участников