

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное учреждение науки
Институт физиологии им. И.П. Павлова Российской академии наук

Лабутина Ольга Васильевна

НАУЧНЫЙ ДОКЛАД
ОБ ОСНОВНЫХ РЕЗУЛЬТАТАХ
НАУЧНО-КВАЛИФИКАЦИОННОЙ РАБОТЫ

**СЛУХОВОЙ АНАЛИЗ СЦЕН, МОДЕЛИРУЮЩИХ УСЛОВИЯ
КОММУНИКАЦИИ ПРИ РАЗНОМ ПОЛОЖЕНИИ ИСТОЧНИКОВ
РЕЧЕВЫХ И НЕРЕЧЕВЫХ СИГНАЛОВ**

Программа подготовки научно-педагогических кадров в аспирантуре по
направлению подготовки кадров высшей квалификации
06.06.01 Биологические науки
Специальность 03.03.01 – физиология

Научный руководитель: **Огородникова Елена Александровна,**
кандидат биологических наук, заведующий лабораторией психофизиологии речи
ФГБУН Институт физиологии им. И.П. Павлова РАН

Санкт-Петербург
2022

СОДЕРЖАНИЕ

Актуальность темы исследования	3
Цель и задачи исследования	5
Научная новизна	6
Положения, выносимые на защиту	7
Теоретическая и практическая значимость	8
Материалы и методы	9
Основные результаты и обсуждения	15
Заключение	25
Выводы	26
Список публикаций по теме исследования	27
Апробация работы	28
Список литературы.....	29

Актуальность темы исследования

Проблема выделения и распознавания полезного (целевого) сигнала в условиях сложной многокомпонентной акустической среды является одной из фундаментальных проблем в области физиологии слуха и изучения механизмов обеспечения эффективной работы естественных и искусственных сенсорно-когнитивных систем. При ее решении в контексте слухового восприятия на первый план выступает задача изучения и моделирования закономерностей анализа акустических сцен, предполагающих одновременную обработку пространственной информации и характеристик речи в процессе коммуникации. Взаимосвязь этих аспектов слухового восприятия в наибольшей степени отражена в так называемой «проблеме вечеринки» («cocktail-party problem»), впервые описанной Черри еще в 1953 году (Cherry, J. *Acoust. Soc. Am.*, 1953). Исследования, затрагивающие процессы пространственной избирательности слуха, энергетической и информационной маскировки, перцептивной группировки, сохраняют свою актуальность, теоретическую и практическую значимость до настоящего времени (Bronkhorst, *Atten. Percept. Psychophys.*, 2015; *The Auditory System at the Cocktail Party*. Eds: Middlebrooks, Simon, Popper, Fay, 2017).

С точки зрения изучения закономерностей слухового восприятия речи в реальных условиях коммуникации, а также развития систем ее интеллектуального анализа с поддержкой принятия решения, одним из ключевых вопросов выступает обеспечение надежного выделения речевого сигнала на фоне других звуковых сигналов, фонового шума и речи других людей (Bronkhorst, *Atten. Percept. Psychophys.*, 2015; *The Auditory System at the Cocktail Party*. Eds: Middlebrooks, Simon, Popper, Fay, 2017). Определение факторов, способствующих выделению целевого речевого сигнала (ЦРС) на фоне помехи с близкими спектральными характеристиками (речеподобный шум или шум многоголосия), в диапазоне расстояний коммуникативного взаимодействия человека будет способствовать пониманию закономерностей пространственной избирательности речевого слуха, алгоритмов слухового анализа сложной акустической сцены (конкурирующие речевые сигналы) и принятия решения в условиях перцептивной неопределенности.

Одним из центральных аспектов разрабатываемой проблемы выступают особенности обнаружения и восприятия речевых сигналов при различном

пространственном расположении их источников и приемников. При этом наиболее изученным направлением являются пространственные схемы с разделением целевого источника речи (диктора) и источника помехи (маскера) по азимутальной координате. На основе результатов этих исследований разработан ряд моделей, описывающих закономерности процесса выделения и распознавания речевого сигнала (Durlach, Schweiz. Rundsch. Med. Prax., 1972; Wan, Durlach, Colburn, J. Acoust. Soc. Am., 2010; Beutelmann et al., J. Acoust. Soc. Am., 2010; Lavandier et al., J. Acoust. Soc. Am., 2012, Audio Source Separation, Ed: Makino, 2018). Подтверждение эффективности бинауральных механизмов в решении задачи пространственной настройки на целевой сигнал привела, в частности, к развитию технологий слухопротезирования и разработке слухопротезирующих систем нового поколения, ориентированных на реализацию механизмов бинаурального взаимодействия, а также к обоснованию необходимости проведения двухсторонней кохлеарной имплантации (Marrone et al., Trends of Amplification, 2008; Dietz, McAlpine, Trends in Hearing, 2015; Moberly, et al. Otol Neurotol, 2016). Это направление востребовано в практике реабилитации пациентов с хронической сенсоневральной тугоухостью и возрастными нарушениями слуха (Бобошко, Голованова, Жилинская, Огородникова. Успехи геронтологии, 2017).

Однако для пространственных конфигураций сцен, при которых источники речи и помехи расположены под одинаковым азимутальным углом, но имеют разную удаленность от слушателя, вопрос о пространственной избирательности слуха изучен в значительно меньшей степени, как для условий свободного поля, так и для закрытых помещений с реверберацией (Brungart, Simpson, J. Acoust. Soc. Am., 2007; Westermann, Buchholz, J. Acoust. Soc. Am., 2015).

В то же время считается, что и в случае локализации источников звука с разной фронтальной удаленностью от приемника, бинауральные механизмы слуха играют существенную роль (Coleman, J. Acoust. Soc. Am., 1968; Zahorik et al., 2005; Kolarik et al., 2016). Проявления подобной зависимости при выполнении задачи локализации источников звука на разном расстоянии (до 4-5 м) показана в ряде работ (Altman, Andreeva, Int. J. Audiol., 2004; Андреева, Сенсорные системы, 2004; Малинина, Андреева, ЖЭБФ, 2013). Имеющиеся данные позволили обосновать **гипотезу** о том, что пространственное разделение источников речи и помехи по расстоянию в диапазоне до 4 м будет способствовать выделению целевого речевого сигнала, в

первую очередь, его обнаружению на фоне даже близкого по спектральным характеристикам речеподобного шума. Экспериментальная проверка гипотезы в серии психофизических экспериментов позволит получить новые данные о процессах слухового анализа сложных акустических сцен, разработать новые методические подходы к оценке состояния пространственного и речевого слуха в медицинской аудиологической практике и при реабилитации пациентов с тугоухостью. Результаты исследования также могут быть использованы для модельного описания процессов выделения целевого сигнала и алгоритмов принятия решения в условиях анализа сложных коммуникативных сцен, ориентированных на повышение эффективности технологий автоматического распознавания акустических сигналов и речи, а также систем искусственного интеллекта.

Цель и задачи исследования

Цель исследования – получение новых данных по пространственной избирательности речевого слуха и закономерностям выделения целевого сигнала в условиях сложной акустической сцены с разделением источников речи и помехи по расстоянию в диапазоне дистанций коммуникативного взаимодействия (1-4 м).

Задачи:

1. Разработка методики формирования сложных коммуникативных сцен, моделирующих воздействие локальной и пространственно-распределенной помехи на базе анэхоидной камеры большого объема (62,5 м³).
- 1.1. Запись монофайлов речевого материала и их предварительная обработка (удаление помех, контроль темпа речи и качества произношения, выравнивание уровня речевого сигнала и т.д.); генерация на основе речевого материала разных дикторов речеподобных шумовых стимулов.
- 1.2. Создание акустической установки для работы в большой анэхоидной звукоизолированной камере, акустические измерения в месте расположения головы слушателя при различных конфигурациях экспериментальных акустических сцен (контроль параметров)
- 1.3. Предварительное обследование группы испытуемых (пороги слышимости, показатели временного разрешения слуха, дихотическое тестирование с речевыми сигналами, оценка когнитивного стиля и уровня тревожности)

1.4. Проведение серий психофизических экспериментов с разной удаленностью источника речи (1, 2, 4 м) от испытуемого и разным соотношением сигнал/шум.

1.5. Обработка полученных результатов по всей группы испытуемых; сравнительная оценка вероятности обнаружения речевых сигналов при разной удаленности источника речи на фоне пространственно-распределенного речеподобного шума различной интенсивности - соотношения сигнала/шума (Signal-to-Noise ratios) SNR= -5, -8, -11 и -14 Дб.

2. Анализ взаимосвязи сенсорных и несенсорных факторов с вероятностью обнаружения речевого сигнала в экспериментальной пространственной сцене.

3. Формирование коммуникативных сцен при отсутствии пространственной составляющей (стимуляция в головных телефонах): восприятие речи в условиях голосовой конкуренции (выделение речевых сигналов целевого диктора при одновременном произнесении слов мужским и женским голосом) без предъявления шумовой помехи; различение голоса диктора (мужской/женский голос) при прослушивании речевых сигналов на фоне речеподобного шума (соотношение сигнал/шум 1:1).

3.1. Получение и анализ экспериментальных данных в группах испытуемых разного возраста и пола, при норме слуха и нарушениях слуха или речи. Оценка влияния голосовых характеристик диктора на восприятие целевых речевых сигналов и выделение целевого диктора в группах сравнения.

4. Подведение итогов исследования с формулированием его выводов.

Научная новизна исследования

В результате исследования получены **новые данные** о пространственной избирательности речевого слуха и о слуховом анализе сложной акустической сцены с разделением конкурирующих источников по расстоянию (пространственная сцена) и по голосу (непространственная сцена). Важно отметить, что для условий свободного поля (безанехоидное пространство) экспериментальные работы при удаленности целевых источников более 1 м в научной литературе практически отсутствуют. Вместе с тем, диапазон расстояний от 1 до 4 м является значимым для коммуникативных контактов и приближен к условиям речевого общения в окружающей человека

повседневной акустической среде. Таким образом, в результате исследования получены новые данные о пространственной избирательности речевого слуха в этом диапазоне эгоцентрических расстояний (1-4 м), а также о влиянии пространственного статуса помехи (воздействие локального и пространственно-распределенного речеподобного шума) на процессы освобождения от маскировки при обнаружении речевой цели и ее распознавании в непространственной сцене в условиях голосовой конкуренции.

Кроме того, получены новые данные о влиянии индивидуальных особенностей испытуемых на вероятность обнаружения ЦРС в сложной сцене, таких как пол, возраст, острота слуха на речевых частотах и выраженность междушной разницы, выраженность латерального предпочтения по восприятию речи, а также уровень нейротизма, экстра-/интроверсии и когнитивный стиль. Проведена оценка действия фактора голоса на выборках взрослых и детей с нормой слуха и нарушением слухоречевой функции, направленная на прикладное применение данных в программах тренинга в рамках коррекционных и реабилитационных мероприятий.

Положения, выносимые на защиту

1. При восприятии в условиях сложной акустической сцены пространственное разделение источников речи и распределенного речеподобного шума является значимым фактором для обнаружения речевого сигнала при низких соотношениях сигнал/шум (SNR): увеличение расстояния до речевого источника с 1 до 2 м приводит к достоверному повышению вероятности обнаружения речи при SNR -14 дБ, а увеличение с 1 до 4 м значительно улучшает обнаружение речи при SNR -11 и -14 дБ.
2. На процессы обнаружения речевого сигнала в условиях сложной пространственной сцены влияют индивидуальные особенности испытуемых, которые можно отнести к сенсорным (междушная разница остроты слуха и его «симметричность», латеральное предпочтение при восприятии речи) и несенсорным (экстра/интроверсия, уровень нейротизма, когнитивный стиль, пол, возраст) факторам.
3. Частота основного тона и гендерные характеристики голоса влияют на выделение целевого речевого сигнала и целевого диктора на фоне речеподобного шума в пространственной и непространственной сценах, а также при распознавании речи в

условиях голосовой конкуренции. При этом выявлены особенности слухового восприятия испытуемых, связанные с состоянием его слухоречевой функции (норма слуха и речи, нарушения слуха, нарушения речи).

Теоретическая и практическая значимость

В результате проведенной работы получены новые данные, связанные с закономерностями слухового анализа сложной пространственной и непространственной коммуникативной сцены и помехоустойчивостью речевого слуха. Были выявлены зависимости обнаружения речевого сигнала в условиях увеличения интенсивности шума и расстояния от слушателя до источника речи: с увеличением SNR ухудшается выделение целевого речевого сигнала, с увеличением расстояния (разделение источников речи и шума в пространстве) – оно улучшается. При этом оказалось, что испытуемые мужского пола демонстрируют тенденцию к большей помехоустойчивости слухового восприятия по сравнению с испытуемыми женского пола в наиболее сложных условиях: минимальное расстояние до источника ЦРС (1м) и максимальная интенсивность шума. Также обнаружено влияние выраженного латерального предпочтения на помехоустойчивость восприятия речи: у амбивалентных по восприятию речи испытуемых она оказалась ниже, чем у испытуемых с выраженным латеральным предпочтением по восприятию речи. Кроме того, показана взаимосвязь ряда других индивидуальных сенсорных и несенсорных характеристик испытуемых с вероятностью обнаружения ЦРС в условиях различных экспериментальных сцен и соотношений SNR - возраст, острота слуха на речевых частотах, уровень экстра-/интроверсии, когнитивный стиль и др.

Полученные данные могут быть использованы при разработке новых методик оценки пространственной избирательности речевого слуха в норме и при нарушениях слухоречевой функции, а также в процессе реабилитации пациентов с тугоухостью (слуховые аппараты, кохлеарная имплантация). Еще одна область практического применения – развитие системы слуховых тренировок, направленных на профилактику центральных слуховых нарушений в пожилом возрасте и ускорение процессов адаптации к новым условиям «слушания» после первичного слухопротезирования.

Кроме того, результаты исследования дают новые возможности организации сенсорной стимуляции, включая не только речевые, но и музыкальные сигналы, для развития речи и коррекции речевых расстройств у детей с нарушениями развития.

Материалы и методы

Основная экспериментальная серия. В рамках основной экспериментальной серии акцент был сделан на оценку вероятности обнаружения речевых сигналов в условиях пространственной сцены с разной удаленностью источника речи от слушателя и воздействия пространственно-распределенного речеподобного шума разной интенсивности.

Стимуляция. Целевой речевой сигнал подавался из динамика, расположенного фронтально перед испытуемым на расстоянии 1, 2 и 4 м; речеподобный masker звучал из динамиков, равноудаленных от головы испытуемого на расстоянии в пределах 1 м (Рис. 1), создавая пространственно-распределенный шум.

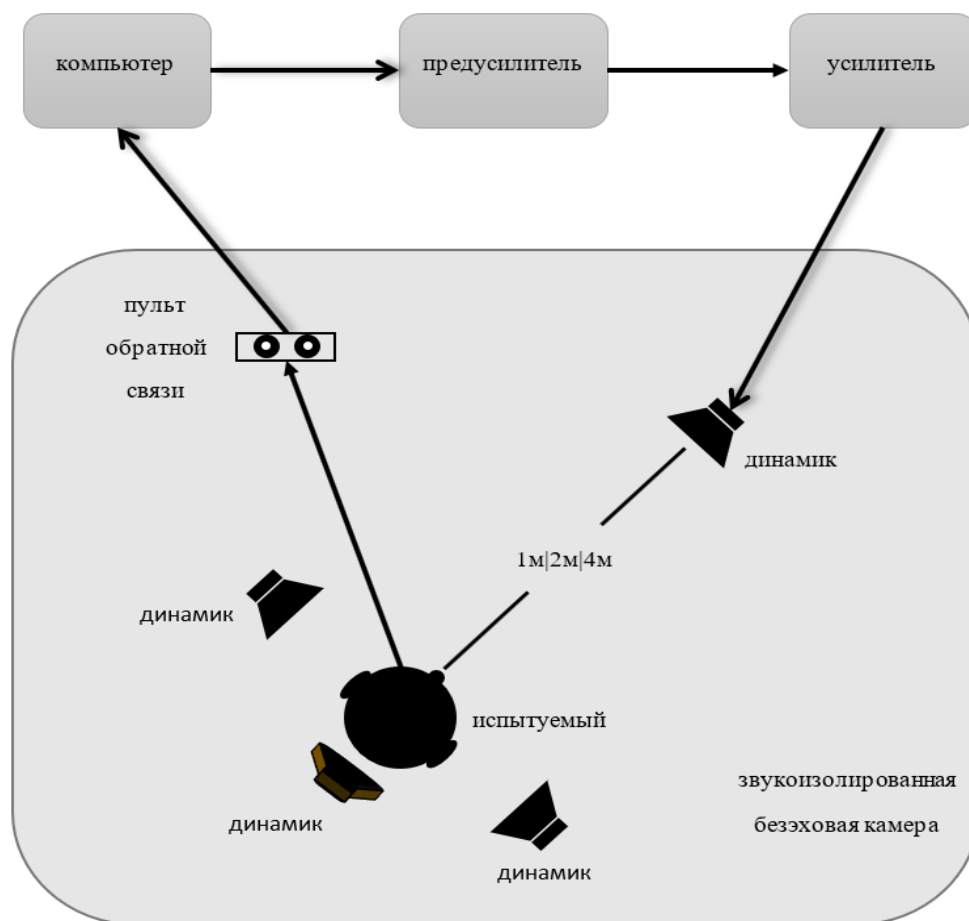


Рис.1. Схема экспериментальной установки.

Речевые стимулы представляли собой 8 двусложных слов: "почва", "строка", "рубеж", "плата", "выпуск", "кредит", "ручка", "набор" с близкими частотами употребления [Национальный корпус русского языка, 2009]. Слова произносили

четыре диктора - двое мужчин (средние значения частоты основного тона голоса 117 и 139 Гц) и две женщины (основные частоты голоса - 208 и 234 Гц). Все дикторы были носителями русского языка и не имели речевых патологий.

Речеподобный шум (маскер) был получен в результате микширования 32 аудиофайла с записями слов (8 слов*4 диктора) - последовательно, без пауз. Его предъявление создавало у испытуемого слуховые ощущения, подобные тем, которые люди испытывают, находясь в многолюдной среде.

В результате выравнивания среднеквадратичной мощности всех речевых стимулов в месте прослушивания их уровень составлял 50 дБ УЗД (SPL) и был одинаковым во всех условиях измерений. Уровень интенсивности речевого маскира в месте прослушивания изменялся с шагом 3 дБ и был равен 55, 58, 61 или 64 дБ (SPL). Это соответствовало вариантам соотношения сигнал/шум при SNR=-5, -8, -11 и -14 дБ. В целом, за исключением характеристик пространственного распределения шума, методические условия серии были близки к условиям в работе [Andreeva et al., 2019].

Временная схема предъявления стимулов в эксперименте отражена на рисунке 2 (Рис. 2).



Рис.2. Временная схема предъявления звуковых сигналов.

Последовательно подавались два отрезка речеподобного шума длительностью по 2 с. На фоне одного из них мог предъявляться целевой речевой стимул. Прослушав каждую пару звучащих отрезков, испытуемый должен был ответить на вопрос: «В течение какого периода шума, первого или второго, появился речевой сигнал?». Ответом выступало нажатие одной из двух кнопок на панели управления, которое фиксировалось автоматически.

Испытуемые участвовали в трех сериях эксперимента, различавшихся расстоянием до источника целевого речевого сигнала (1, 2 и 4 м). Каждая серия представляла собой последовательность из 40 предъявлений пар стимулов.

При обработке вероятности правильных ответов оценивались для каждого испытуемого и для трех пространственных конфигураций (удаленность источника речи) и 4-х соотношений SNR (-5, -8, -11 и -14 дБ). Достоверность различий в обнаружении ЦРС рассчитывалась с помощью непараметрического статистического критерия Уилкоксона.

Испытуемые. В экспериментах основной серии приняли участие 30 испытуемых в возрасте 20-45 лет (15 мужчин и 15 женщин). Все испытуемые прошли предварительное обследование состояния слуха и дополнительное психологическое тестирование.

Состояние слуха оценивали с помощью тональной пороговой аудиометрии (Аудиометр АА-02) и теста обнаружения паузы (GDT – Gap Detection Test). Кроме того, по данным аудиограмм фиксировали междушную разницу в остроте слуха на «речевых частотах» (500-4000 Гц).

Дополнительное тестирование было направлено на получение следующих показателей:

- латеральное предпочтение по восприятию речи (дихотический тест);
- уровень экстра-/интроверсии; уровень нейротизма (тест Айзенка);
- когнитивный стиль (полезависимость / полenezависимость по методике «Фигуры Готтшальдта»).

По этим показателям, а также по возрасту и полу, были выделены подгруппы испытуемых, между которыми проводилось сравнение результатов обнаружения ЦРС (критерий Манна-Уитни, Краскела-Уоллиса). Дополнительно был сделан корреляционный анализ с использованием метода ранговых корреляций Спирмена. Минимальная подгруппа сравнения соответствовала 5 испытуемым.

Помещение. Психофизические измерения проводили в условиях свободного поля - в звукоизолированной безэховой камере объемом 62,5 м³ с затуханием внешнего шума в диапазоне частот 0,5–16 кГц не менее 40 дБ.

Дополнительная экспериментальная серия была организована с целью исследования влияния голосовых характеристик дикторов (мужской/женский голос) на восприятие сложной акустической сцены при отсутствии пространственного компонента (стимуляция через головные телефоны).

Психофизические измерения проводили с применением методики, имитирующей условия «речевого коктейля» при одновременном произнесении разных слов диктором-мужчиной и диктором-женщиной (условия голосовой конкуренции) и при выделении целевого голоса (мужской/женский) на фоне речеподобного шума. Для организации стимуляции использовали программу «Учись слушать» - совместную разработку специалистов Института физиологии им. И.П. Павлова РАН и Санкт-Петербургского Института уха, горла, носа и речи Минздрава РФ, направленную на формирование и развитие базовых навыков слухоречевого восприятия у пациентов после кохлеарной имплантации (Ogorodnikova et al., 2009; Люблинская и др., 2009; Королева и др., 2013; Огородникова и др., 2014).

При восприятии *непространственной сцены с фоновым речеподобным шумом* речевыми сигналами выступали те же 8 слов, произнесенных 4 дикторами (2 мужчины и 2 женщины), которые предъявлялись в основной серии в условиях свободного поля. Рабочее соотношение «сигнал/шум» соответствовало SNR=0 (1:1). Задача испытуемого состояла в различении голосов диктора и в их гендерной оценке - мужской/женский. Экспериментальная таблица состояла из 32 стимулов (8 слов * 4 диктора). Медианные показатели частоты основного тона (ЧОТ) для голосов дикторов и используемых сегментов речеподобного шума приведены на рисунке 3.

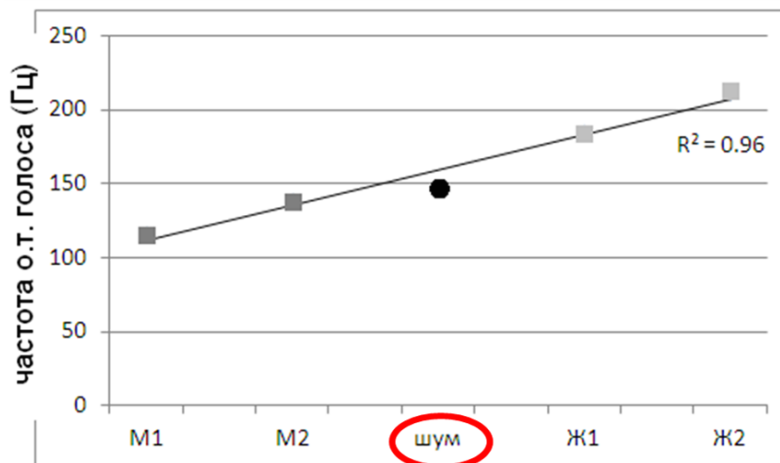


Рис. 3. Характеристики голосов дикторов и сегментов речеподобного шума. По оси ординат – условные обозначения 2-х дикторов-мужчин (M1, M2); 2-х дикторов-женщин (Ж1, Ж2) и шумовой помехи («шум»). По оси абсцисс – частота основного тона голоса (ЧОТ, в Гц)

В процессе тестирования стимулы предъявлялись в случайном порядке на комфортном уровне стимуляции (60-65 дБ), через головные телефоны с близкими амплитудно-частотными характеристиками.

В измерениях приняли участие 2 экспериментальные группы взрослых испытуемых с нормальным слухом: 1 – 14 человек разного пола и возраста (7 мужчин и 7 женщин, средний возраст 37 лет, диапазон от 26 до 57 лет); 2 – 28 студентов в возрасте от 17 до 22 лет (6 мужчин, 22 женщины, средний возраст 18.4 ± 1.1).

При восприятии *непространственной сцены с голосовой конкуренцией* задачей являлось выделение и распознавание речевых сигналов целевого диктора (мужской/женский голос).

Психофизические измерения проводили на базе методики имитации «речевого коктейля» путем одновременного произнесения разных слов диктором-мужчиной и диктором-женщиной (Ogorodnikova et al., 2009; Люблинская и др., 2009; Королева и др., 2013; Огородникова и др., 2014).

Речевой материал был представлен записями изолированных трехсложных слов 2-х дикторов - мужчины (средняя частота основного тона голоса ЧОТ=108 Гц) и женщины (ЧОТ=185 Гц) с нормативным для носителей русского языка произношением и с нейтральной интонацией. Цифровые реализации записей выравнивались по интенсивности и микшировались, чтобы суммарный тестовый стимул представлял собой смесь разных слов, произнесенных мужским (М) и женским (Ж) голосом с синхронизацией времени начала их звучания (рис. 4).

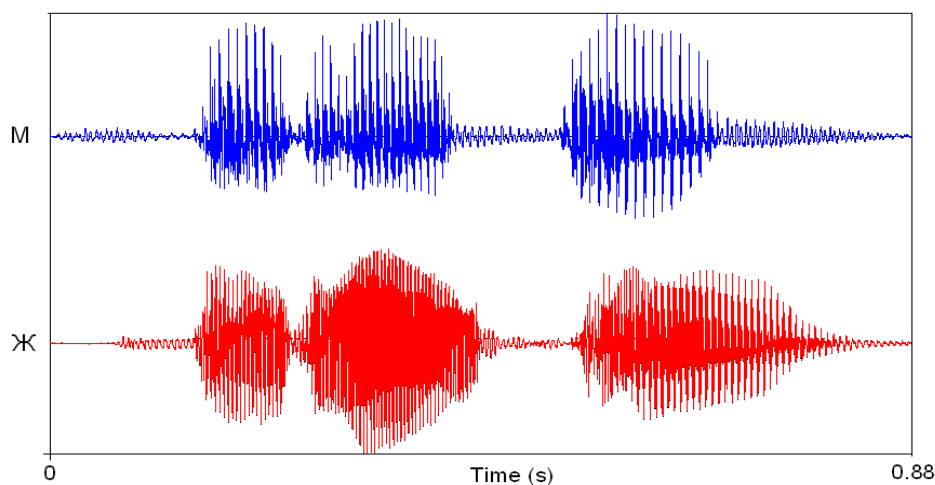


Рис.4: Примеры записи 2-х трехсложных слов, произнесенных диктором-мужчиной (М, «барабан») и диктором-женщиной (Ж, «борода»), до их микширования.

Всего в тестовый набор стимулов были включены 6 пар конкурирующих слов: /Ягода+Армия/, /Борода+Барабан/, /Воробей+Берёза/, /Бумага+Дерево/; /Диваны+Лимоны/, /Ягода+Выгода/ - прямым шрифтом отмечены слова диктора М, курсивом – диктора Ж. Пример осциллограммы одной из них приведен на рис. 5.

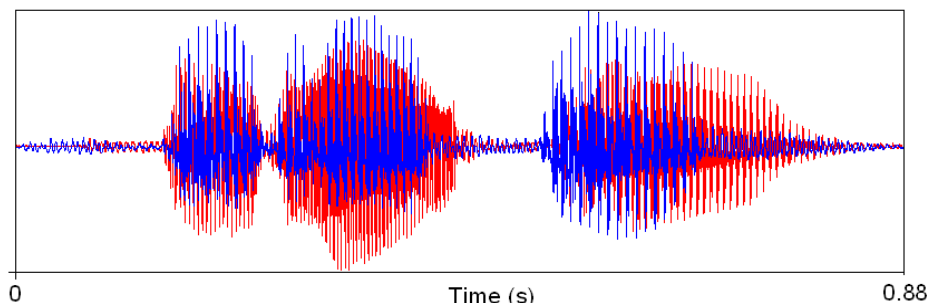


Рис.5: Пример осциллограммы тестовой пары конкурирующих слов /Борода+Барабан/ после микширования исходных записей М и Ж-дикторов (см. рис. 4).

В процессе тестирования конкурирующие пары предъявлялась в случайном порядке, на комфортном уровне стимуляции (от 60 до 70 дБ). Для выбора ответа на экран монитора выводилась таблица с текстовыми записями всех слов, входящих в тестовые пары (рис. 6).

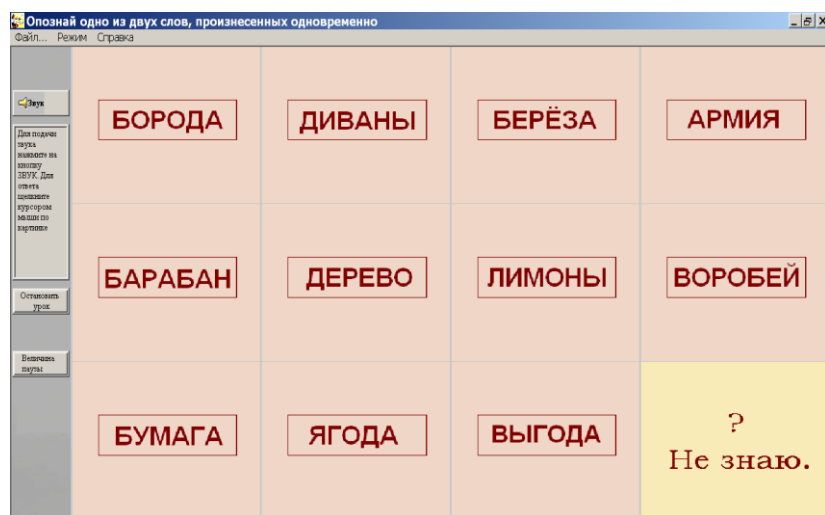


Рис. 6: Отображение на экране вариантов ответа при прослушивании конкурирующих речевых стимулов.

Испытуемый реагировал щелчком «мышки» по опознанному слову. Выбранный ответ и время реакции фиксировались в таблицах EXCEL:

Согласно инструкции испытуемые должны были выделить и распознать слова, которые произнесены мужским (тест 1) или женским (тест 2) голосом.

В опытах приняли участие испытуемые разного возраста и слухоречевого статуса: взрослые испытуемые в возрасте от 18 до 24 лет с нормой слуха ($n=27$) и нарушениями слуха (3 степень сенсоневральной тугоухости, кохлеарная имплантация; $n=17$), школьники в возрасте 10-12 лет с типичным развитием ($n=12$) и с речевыми нарушениями - общее недоразвитие речи (ОНР), проявления дислексии и дисграфии ($n=12$).

Статистический анализ. Для сравнения парных данных и оценки достоверности различий результатов, полученных в различных экспериментальных условиях применяли непараметрические критерии Уилкоксона и Манна-Уитни. При сравнении данных нескольких групп использовался Н-критерий Краскела-Уоллиса.

С целью выявления связей между отдельными выделенными признаками и их возможного влияния на переменные значения применялся коэффициент ранговой корреляции Спирмена.

Статистические расчеты производились в программе STATISTICA и, в некоторых случаях, в программе SPSS.

Основные результаты и обсуждения

Результаты основной экспериментальной серии с разделением источников целевого речевого сигнала и маскера по расстоянию показали, что вероятность обнаружения речевого сигнала на фоне пространственно-распределенного речеподобного шума закономерно снижается с уменьшением SNR для всех исследуемых расстояний от источника речи до слушателя.

Наибольшие изменения этого показателя наблюдаются при расстоянии 1 м от целевого речевого стимула, наименьшие – для расстояния 4 м. Это означает, что в рамках пространственной экспериментальной сцены, речевые стимулы из удаленных от слушателя источников с большей вероятностью обнаруживаются при максимальном уровне маскировки пространственно-распределенным шумом с близкими к речи спектральными характеристиками.

Зависимость вероятности обнаружения сигнала при разных SNR и тестируемых расстояниях показана на рис. 7. Видно, что она наиболее выражена при удалении целевого речевого стимула на расстояние в 1 м, наименее – при расстоянии в 4 м.

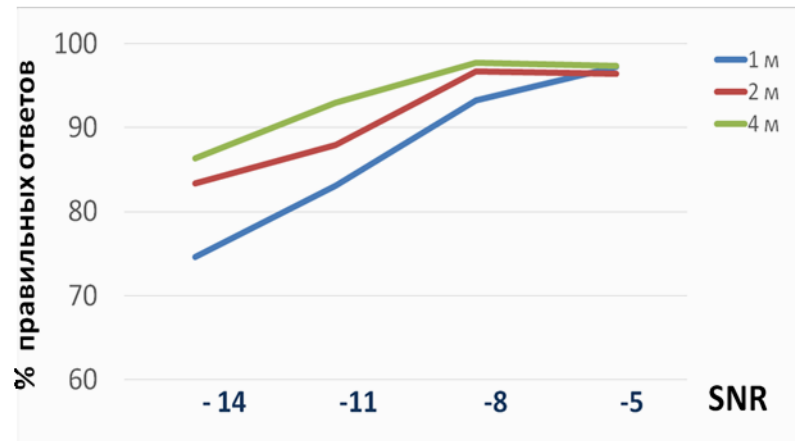


Рис.7. Базовая зависимость средних значений вероятности обнаружения ЦРС (правильные ответы, %) при различных расстояниях источника речи от слушателя и SNR (значения по горизонтальной оси).

Сравнение усредненных по группе испытуемых данных для разных соотношений сигнал/шум демонстрирует, что вероятность обнаружения ЦРС значительно повышается с увеличением расстояния до источника речи и разделения пространственного положения источников речи и речеподобного шума (рис. 8). Вместе с тем, маскер слабой интенсивности практически не влияет на вероятность распознавания целевого сигнала, она оказывается приближенной к 100% на всех расстояниях.

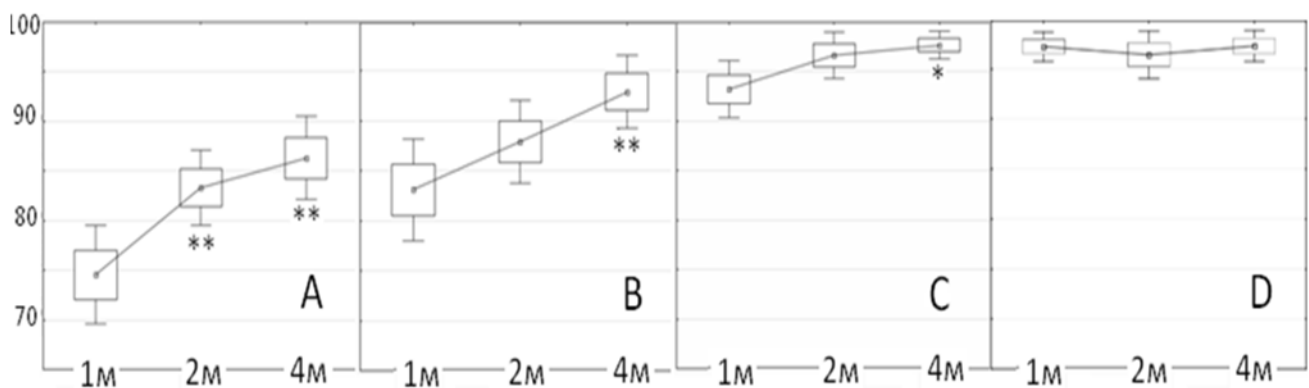


Рис. 8. Средняя по группе вероятность обнаружения ЦРС при разных расстояниях от слушателя до источника сигнала при различных соотношениях С/Ш (А – -14 дБ, В – -11 дБ, С – -8 дБ, D – -5 дБ). N = 272. По оси абсцисс – расстояние от слушателя до источника ЦРС, м, по оси ординат – вероятность обнаружения сигнала, %. Указаны стандартная ошибка среднего и удвоенная стандартная ошибка среднего.

Звездочками отмечены случаи достоверных отличий (** при $p < 0.01$ и * при $p < 0.05$) показателей вероятности обнаружения при 4 м и 2 м от соответствующего показателя при 1 м.

Таким образом, полученные результаты показали, что расстояние от источника речи до источников пространственно-распределенного шума является значимым фактором пространственной избирательности, как и разделение в пространстве локализованных источников речи и речеподобного маскиера в работе [Andreeva et al., 2019].

При этом проявились гендерные особенности восприятия исследуемой пространственной сцены (рис.9), которые были особенно выражены в условиях максимальной сложности при маскировке ЦРС – удаленность 1 м, SNR= -14 дБ.

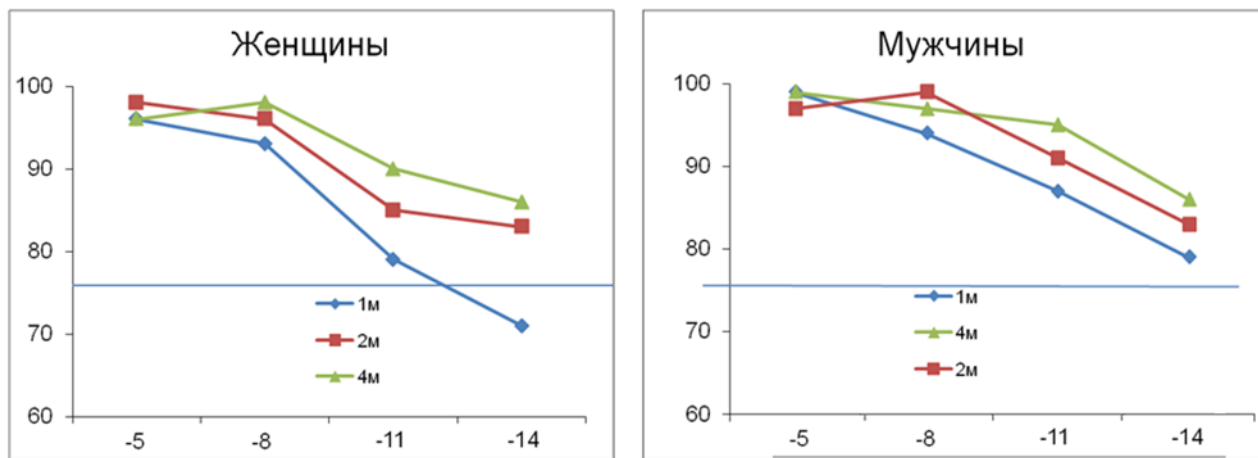


Рис.9. Зависимости вероятности обнаружения ЦРС от SNR при различных расстояниях источника речи от слушателя (данные испытуемых-женщин и испытуемых-мужчин).

По вертикали – оценки вероятности обнаружения ЦРС (%), по горизонтали – соотношение сигнал/шум (SNR).

Продолжением исследования в условиях многокомпонентной пространственной сцены стало получение и анализ данных о сенсорных и несенсорных факторах, влияющих на вероятность обнаружения целевого речевого сигнала (ЦРС) на фоне речеподобного пространственно-распределенного шума. Ряд выделенных характеристик (пол, возраст, латеральное предпочтение по восприятию речи) продемонстрировали взаимосвязь с вероятностью обнаружения ЦРС при определенных условиях восприятия экспериментальной акустической сцены (рис. 10). В наиболее сложной для испытуемых ситуации (максимальный шум SNR= -14 дБ, источник речи на расстоянии 1 м от слушателя – без разделения в пространстве источников шума и речи) вероятность обнаружения РС значимо коррелировала с междушной разницей остроты слуха. В более простой (минимум шума при SNR= -5 дБ) - с экстра/интраверсией (1 м) и с возрастом (2 м).

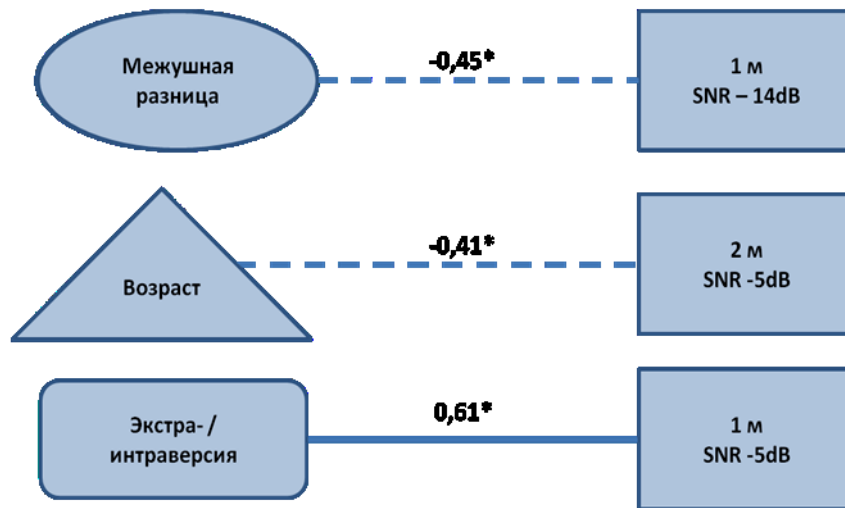


Рис. 10. Взаимосвязь вероятности обнаружения речевого сигнала для различных расстояний и SNR с характеристиками испытуемых (межушная разница в остроте слуха, возраст, уровень экстра-/интраверсии). Данные по всей выборке испытуемых ($n=30$). Обозначение * - $p<0.05$.

Для групп испытуемых с разным латеральным предпочтением по восприятию речи значимые факторы различались (рис. 11):

- при амбивалентности - это были экстра/интраверсия, уровень нейротизма и когнитивный стиль ($r=0.80-0.87$; 1 и 2 м; -5, -8 и -11 дБ);
- при проявлении право- или левосторонней асимметрии – возраст ($r= -0.60$ и -0.67 ; 4 м; -11 и -14 дБ)

Амбивалентность

Латеральное предпочтение

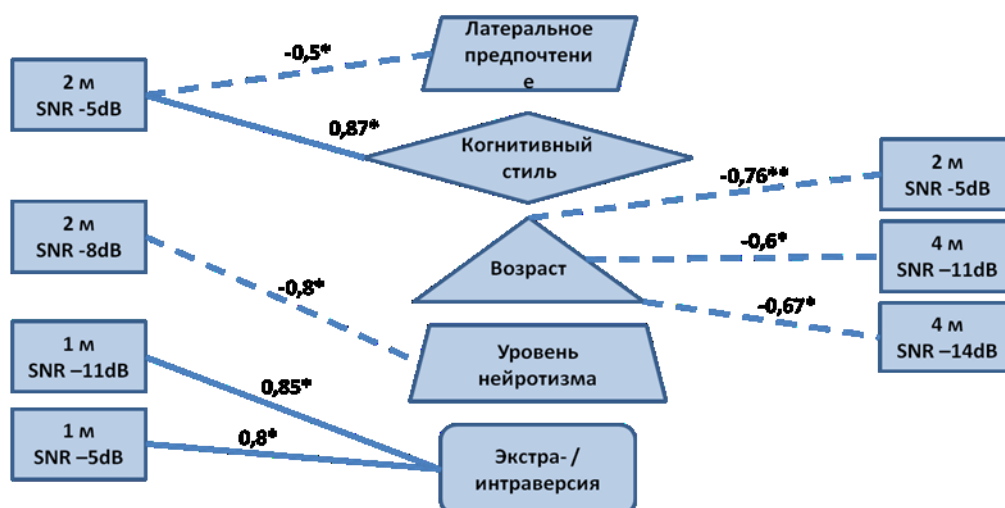


Рисунок 11. Взаимосвязь вероятности обнаружения речевого сигнала при различных расстояниях и SNR для групп испытуемых с разным латеральным предпочтением при восприятии речи.

В наиболее сложных условиях восприятия (*минимум разделения – расстояние 1 м от слушателя; максимум уровня шума SNR= -14 дБ*) получены достоверные различия в вероятности обнаружения РС для испытуемых с латеральным предпочтением по восприятию речи и амбивалентностью ($p < 0,01$). Их отражают данные, приведенные на рисунке 12.

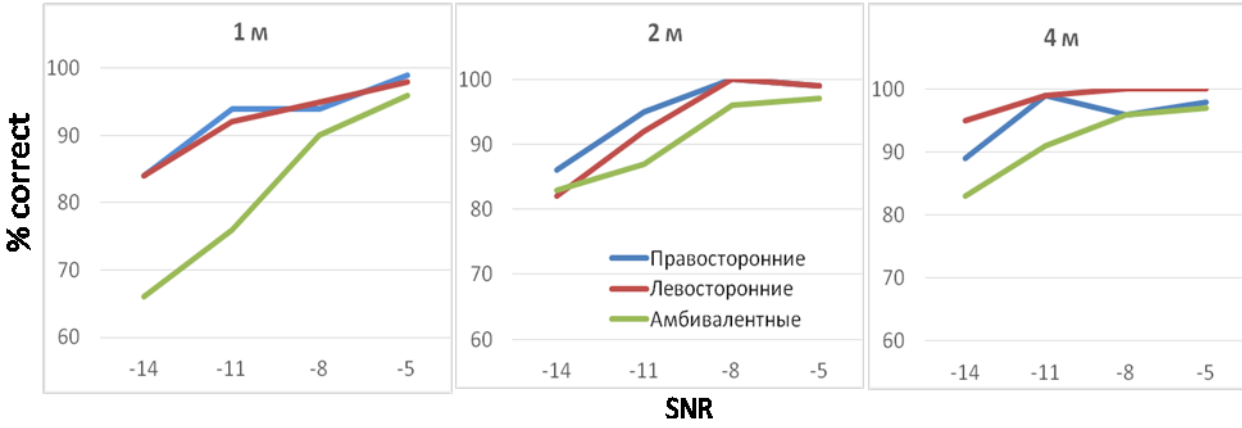


Рис. 12. Данные испытуемых с различными латеральными предпочтениями по восприятию речи при обнаружении речевого сигнала в условиях разных пространственных сцен (удаленность источника речи, SNR)

В отношении когнитивного стиля (полезависимость) значимых различий по всей группе испытуемых обнаружено не было, но отмечена тенденция к их проявлению, особенно в условиях восприятия «средней» сложности (2 м удаленности) (рис. 13).

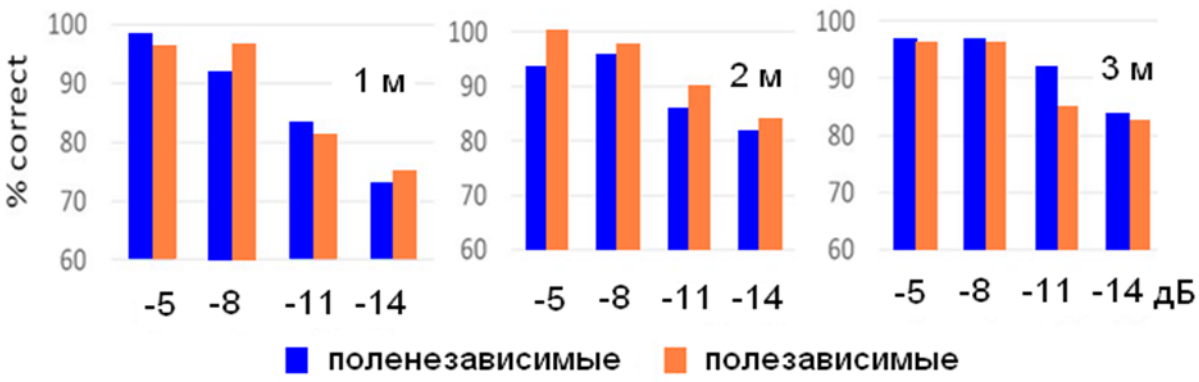


Рис. 13. Данные для испытуемых с разным когнитивным стилем по показателям полезависимости/полнезависимости.

При этом значимая связь ($r=0.8$) с когнитивным стилем была получена у испытуемых младше 35 лет. У испытуемых старше 35 лет проявилась сильная корреляция с уровнем «экстра/интраверсии» ($r=0.92$).

Результаты дополнительного исследования. При восприятии непространственной сцены с различием голосовых характеристик дикторов (мужской/женский голос) на фоне речеподобного шума в экспериментальной группе 1 (n=14) было показано, что успешно (критерий – 70% правильных ответов) эту решают задачу 57.1% испытуемых, неуспешно - 42.9%.

Средние показатели по группе составили: правильных ответов (определение пола диктора) $71 \pm 2.6\%$, ошибок - $28.1 \pm 2.5\%$, пропусков - 0.9%, среднее время реакции - 2.99 ± 0.2 сек. Были также показаны значимые различия ($p < 0.05$) для возрастных подгрупп - до и после 40 лет – средний возраст 28 и 47 лет, соответственно (рис. 14).

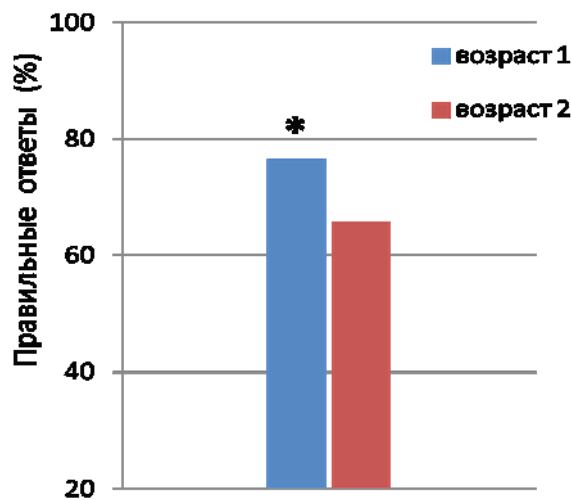


Рис. 14. Данные для 2-х возрастных групп испытуемых: возраст 1 - 7 человек <40 лет (средний возраст 28 лет); возраст 2 - 7 человек ≥ 40 лет (средний возраст 47 лет). По вертикали - % правильных опознаваний голоса диктора (М, Ж голос). * - достоверность различий на уровне $p < 0.05$ (критерий Манна-Уитни).

При этом различий для испытуемых мужского и женского пола обнаружено не было. В то же время выраженные различия наблюдались при различении М и Ж-голосов ($p < 0.01$) - ошибок для Ж-голоса оказалось существенно больше как в общей группе, так и во всех подгруппах сравнения (возраст, пол испытуемых). Для подтверждения зависимости была проведена дополнительная серия экспериментов с более однородной по возрасту группой испытуемых.

Измерения, проведенные с участием экспериментальной группы 2 из 28 студентов (средний возраст 18 ± 1.1 год), показали, что с определением гендерного соответствия голоса диктора на фоне речеподобного шума (SNR=1:1) в условиях акустической сцены без пространственной составляющей (прослушивание через

головные телефоны), успешно (более 70% правильных ответов) справились только 46% испытуемых.

В то же время уровень различения М и Ж голосов менее 60% правильных ответов наблюдался лишь у 4 человек или в 14% случаев по всей группе (Рис. 15).

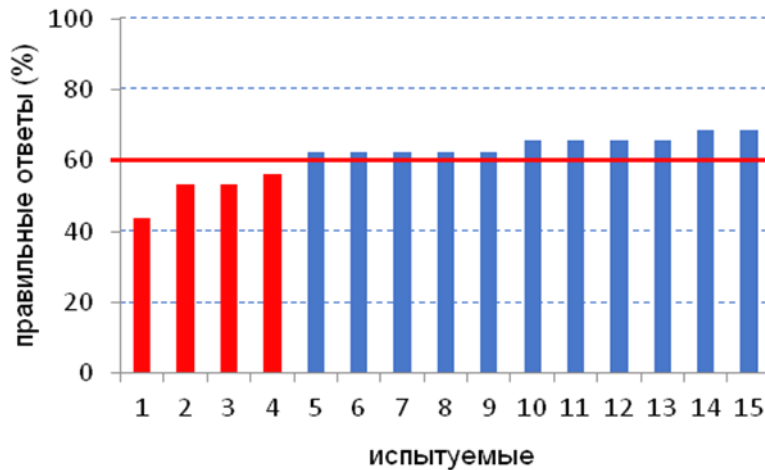


Рис. 15. Данные «неуспешных» испытуемых, показавших низкий уровень правильных ответов при различении голосовых характеристик диктора на фоне речеподобного шума (менее 70%). Цветом выделены результаты с показателем менее 60% правильных ответов (14% случаев в экспериментальной группе 2).

Средние показатели по экспериментальной группе 2 составили: для правильных распознаваний гендерной принадлежности голоса диктора 68% ($m \pm 1.8$); для ошибок – 30% ($m \pm 1.9$), для времени реакции – 3.2 сек ($m \pm 0.1$). Сравнение этих данных с данными, полученными в экспериментальной группе 1, свидетельствовало о проявлении значимых различий ($p < 0.05$) между подгруппами «возраст 1» (средний – 18 лет) и «возраст 2» (средний – 28 лет), т.е. также как ранее для подгрупп «возраст 3» (средний 47 лет) и «возраст 2» (Рис. 16).

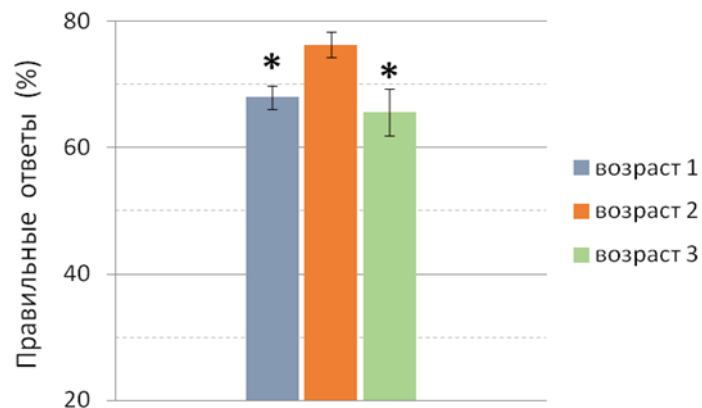


Рис. 16. Данные для 3-х возрастных подгрупп испытуемых: возраст 1 (средний - 18 лет); возраст 2 (средний - 28 лет); возраст 3 (средний - 47 лет). Обозначение * - уровень значимости различий $p < 0.05$ (критерий Манна-Уитни).

Можно предположить, что такая закономерность определяется выделением возрастного периода наилучшего сенсорно-когнитивного функционирования (25-35 лет) с оптимальным соотношением уровня концентрации и избирательности рабочего внимания и характеристик слуха.

В отношении восприятия М и Ж-голосов значимых различий в экспериментальной группе 2 выявлено не было. Среднее число ошибок при их опознавании составило $30 \pm 2.8\%$ (М-голоса) и $29 \pm 2.7\%$ (Ж-голоса). При этом максимальный процент ошибок наблюдался для дикторов М2 и Ж1 (Рис. 17), относящихся к среднему диапазону ЧОТ близкому к характеристикам речеподобного шума (Рис. 17), что, вероятно, свидетельствовало о максимальном проявлении эффектов маскировки.

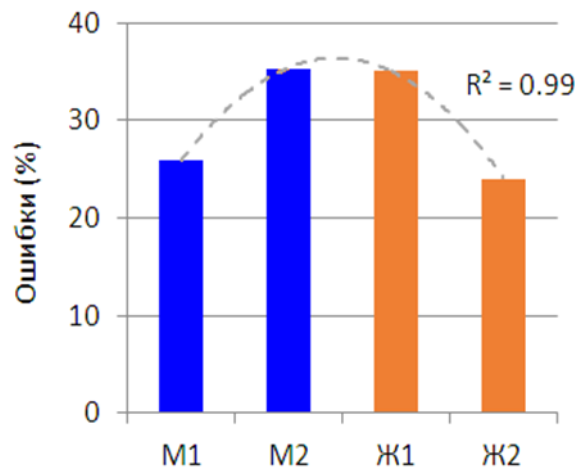


Рис. 17. Доля ошибок при определении гендерной принадлежности голосов дикторов. Обозначения: М1, М2 – 2 диктора-мужчины; Ж1, Ж2 – 2 диктора-женщины. Пунктир – полиномиальная аппроксимация данных и коэффициент ее достоверности (R^2).

При этом зависимость правильных ответов от характеристик голоса диктора оказалась очень близкой (Рис. 18) к зависимости обнаружения речевого сигнала в пространственной сцене при разделении локализованных источников речи и речеподобной помехи (3 м) с максимальным уровнем зашумления – $SNR = -14$ дБ, полученной в работе Андреевой с соавторами (2019).

Данные рассматривались как подтверждение адекватности выбранных характеристик экспериментальных непространственных сцен с включением речеподобной помехи.

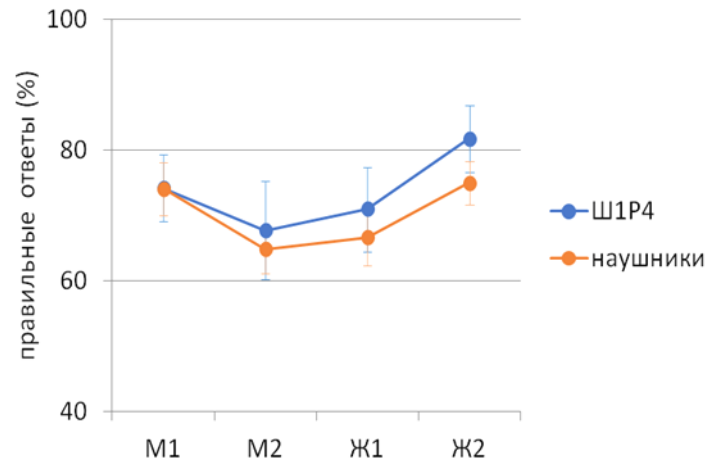


Рис. 18. Зависимости обнаружения речевого сигнала в пространственной сцене с разделением источников шума (1 м от слушателя) и речи (4 м от слушателя) при SNR= -14 дБ и правильного определения гендерной принадлежности голоса диктора при прослушивании речевого сигнала на фоне речеподобного шума (SNR 1:1) через наушники. Разброс соответствует показателю стандартной ошибке среднего (m).

Результаты, полученные с участием испытуемых с нормой слуха и слухоречевыми нарушениями, продемонстрировали значительные различия в показателях выделения и распознавания целевых слов в условиях непространственной сцены с голосовой конкуренцией (табл. 1).

Таблица 1: результаты измерений в условиях голосовой конкуренции

группы сравнения	взрослые испытуемые		группы сравнения	школьники	
	N (%)	T реакции (с)		N (%)	T реакции (с)
норма слуха	92.3±1.2	2.3±0.07	типичное развитие	84±2.6	2.9±0.2
нарушения слуха	53.7±2.6	4.7±0.2	нарушения речи	67.2±4.2	3.2±0.4

Можно видеть, что успешное выделение и распознавание целевых слов при голосовой конкуренции наблюдается у взрослых аудиторов с нормой слуха и у школьников с типичным развитием (норма слуха и речи), у которых средний уровень правильных ответов превышает уровень надежного распознавания в 75% и

клинический критерий успешности – 70% правильных ответов. В этих же группах зафиксированы и более низкие показатели времени реакции.

Наибольшие трудности при выполнении задания испытывали аудиторы с нарушениями слуха и, в первую очередь, пациенты после кохлеарной имплантации на начальных этапах реабилитации, которые проходили тестирование без предварительного обучения. Не достигли уровня надежного распознавания и школьники с нарушениями речи. У них также наблюдался большой разброс индивидуальных данных, а средние показатели времени реакции существенно превышали соответствующие значения у взрослых и школьников с нормой слухоречевой функции.

Кроме того, в группах с нормой слуха (взрослые испытуемые и школьники) отмечалась тенденция к лучшему выделению слов, произнесенных диктором-женщиной, в среднем, на 5%. В то же время в группе с нарушениями слуха, ситуация была обратная – показатели распознавания слов, произнесенных мужским голосом, были выше на 16.8% (у пациентов с КИ) и на 7.9% (у пациентов с тугоухостью).

Результаты измерений в условиях голосовой конкуренции свидетельствуют об ухудшении способности выделять целевые речевые потоки и распознавать речь в условиях сложной акустической среды при нарушениях слухоречевой функции. Такое ухудшение может определяться не только сенсорным дефицитом (снижение слуховой чувствительности), но и недостаточной сформированностью центральных механизмов слухового анализа и связанных с ними когнитивных процессов, в частности селективного слухового внимания при недостатке сенсорного опыта. Это, в первую очередь, относится к пациентам после кохлеарной имплантации на начальных стадиях реабилитации. Сходные проявления наблюдаются и у пациентов с тугоухостью 3-й степени. У этих групп пациентов в силу особенностей перцептивного опыта с опорой на остаточный слух в области низких частот преобладает надежность распознавания слов, произнесенных мужским голосом с более низким значением частоты основного тона голоса.

У школьников с нарушениями речи низкие показатели распознавания в условиях конкуренции речевых сигналов также могут отражать проблемы с центральной обработкой акустической информации, что хорошо согласуется с данными исследования процессов слухового сегментного анализа при нарушениях

речи и письма [Огородникова, Балякова, 2014]. При этом у них тоже проявились различия в отношении гендерных предпочтений по голосу, которое ближе к испытуемым из группы нормы – лучше выделяются слова, сказанные женским голосом (в среднем, на 13%). Эти эффекты в норме и при нарушениях речи могут быть связаны с особенностями развития речи в период раннего детства (диады «мать-дитя») и с социальными условиями - преобладанием воспитателей и педагогов женского пола в дошкольных и школьных образовательных учреждениях, особенно в области дефектологии и при проведении коррекционной работы [Еркудов и др., 2019].

Заключение

Проведенное исследование предоставило экспериментальные подтверждения начальной гипотезы и положений, выдвинутых на защиту. В результате анализа результатов основной экспериментальной серии были получены новые данные в отношении закономерностей обнаружения речевого сигнала в сложной пространственной сцене с нелокализованной речеподобной помехой. При этом вероятность обнаружения сигнала уменьшается пропорционально уменьшению SNR при всех тестируемых расстояниях от слушателя до источника речи. Наиболее заметное снижение наблюдается для дистанции 1 м, наименее заметное - 4 м.

Данные свидетельствуют, что, как и в случае с локализованными источниками речи и шума, расстояние и разделение их в пространстве является решающим фактором в отношении обнаружения речевого сигнала при низких SNR: увеличение расстояния с 1 до 2 м приводит к значимому улучшению обнаружения речи при SNR – 14 дБ, а увеличение с 1 до 4 м для двух уровней шума при –11 и –14 дБ SNR. Таким образом, бинауральные механизмы пространственного слуха, ответственные за восприятие удаленности, способствуют проявлению эффектов освобождения от маскировки и обнаружению целевого речевого стимула на фоне наиболее сильной помехи – близкого по спектральным характеристикам речеподобного шума с размытыми параметрами локализации.

Показано также, что на вероятность обнаружения речевого сигнала в сложной пространственной сцене влияет ряд дополнительных сенсорных и несенсорных факторов. Так, обнаружена значимая связь с показателями латерального предпочтения по восприятию речи, остроты слуха и междушумных различий в порогах на «речевых»

частотах, уровнем экстра/интроверсии и нейротизма, а также когнитивного стиля, пола и возраста, которая проявляется при разных параметрах акустической сцены (расстояние между источниками речи и шума, соотношение сигнал/шум).

В условиях непространственной сцены (прослушивание через головные телефоны) при определении гендерной принадлежности голоса диктора на фоне речеподобного шума проявилась зависимость от возраста испытуемых. При восприятии речи в условиях голосовой конкуренции показаны различия в группах разного возраста и состояния слухоречевой функции. Таким образом, дополнительные серии подтвердили, что частота основного тона является значимым перцептивным фактором восприятия сложной коммуникативной сцены как при норме слуха и речи, так и в случае их нарушений.

Полученные данные, помимо теоретического значения в отношении расширения представлений о пространственной избирательности речевого слуха и анализа сложных акустических сцен, имеют прикладную важность с точки зрения развития технологий слухопротезирования и систем автоматического распознавания речи, а также подготовки программ слухоречевого тренинга для ускорения адаптации пациентов к использованию слуховых аппаратов и повышения эффективности курсов реабилитации после операций кохлеарной имплантации.

Выводы

1. Показано, что при восприятии в условиях сложной акустической сцены пространственное разделение источников речи и распределенного речеподобного шума является значимым фактором для повышения вероятности обнаружения речевого сигнала при низких соотношениях сигнал/шум (SNR): увеличение расстояния до речевого источника с 1 до 2 м приводит к достоверному повышению вероятности обнаружения речи при SNR -14 дБ, а увеличение с 1 до 4 м значительно улучшает обнаружение речи при SNR -11 и -14 дБ.

2. Обнаружено, что на процессы обнаружения речевого сигнала в условиях сложной пространственной сцены влияют индивидуальные особенности испытуемых, которые относятся как к сенсорным факторам - междушная разница остроты слуха и его «симметричность», латеральное предпочтение при восприятии речи, так и к

несенсорным факторам – уровень экстра/интроверсии, уровень нейротизма, когнитивный стиль, пол и возраст испытуемых.

3. Подтверждена значимость гендерных характеристик голоса и частоты основного тона для вероятности выделения целевого речевого сигнала и целевого диктора на фоне речеподобного шума в пространственной и непространственной сценах, а также при распознавании речи в условиях голосовой конкуренции. Выявлены особенности слухового восприятия сложных коммуникативных сцен при различном состоянии слухоречевой функции (норма слуха и речи, нарушения слуха, нарушения речи) и возраста.

Список публикаций по теме исследования

1. Огородникова Е.А., Лабутина О.В., Пак С.П. Имитация сложной акустической сцены при стимуляции через головные телефоны. «Вестник психофизиологии». 2022. №2. С.140-146. DOI: 10.34985/o0640-6924-4290-f.

2. Огородникова Е.А., Пак С.П., Столярова Э.И., Лабутина О.В. Особенности восприятия конкурирующих речевых сигналов в норме и при нарушениях слухоречевой функции. Труды III Всероссийской акустической конференции. Раздел «Акустика речи». СПб: Политех-Пресс, 2020. С.381-386. ISBN 978-5-7422-7029-4 (РИНЦ, eLIBRARY ID: 44052314).

3. Smirnova V. A., Labutina O. V., Gvozdeva A. P. Chapter 9: Speech detection in spatially distributed speech-like noise. In: Neural Networks and Neurotechnologies (eds: Yu. Shelepin, E. Ogorodnikova, N. Solovyev, E. Yakimova). SPb, Publish by VVM, 2019. P.52-60. ISBN 978-5-9651-1259-3.

4. Огородникова Е.А., Андреева И.Г., Гвоздева А.П., Лабутина О.В., Пак С.П. Фактор голоса в пространственной избирательности речевого слуха. Интегративная физиология: Всероссийская конференция с международным участием, посвящённая 95-летию Института физиологии им. И.П. Павлова РАН (СПб, 9-11 декабря 2020). Тезисы докладов. СПб.: Ин-т физиологии им. И.П. Павлова РАН, 2020. С.66-67. ISBN 978-5-6045715-0-7

5. Огородникова Е.А., Лабутина О.В., Андреева И.Г., Гвоздева А.П., Баулин Ю.А.. Фактор просодики в восприятии коммуникативной сцены с пространственным разделением источников речи и речеподобной помехи. Тезисы в сборнике по итогам Международной конференции «Лингвистический форум 2020: Язык и искусственный

интеллект». 12–14 ноября 2020 г. Институт языкознания РАН, Москва: Москва: Тезисы докладов / Под ред. А.А. Кибрика, В.Ю. Гусева, Д.А. Залманова. М.: Институт языкознания РАН, 2020. С.127-128.

6. Лабутина О.В., Огородникова Е.А., Гвоздева А.П., Андреева И.Г. Влияние несенсорных факторов на обнаружение речевого сигнала в условиях пространственно распределенного шума. Интегративная физиология: Всероссийская конференция с международным участием (Санкт-Петербург, 8-10 декабря 2021 г.). Тезисы докладов. СПб.: Ин-т физиологии им. И.П. Павлова РАН, 2021. С.71.

Апробация научной работы

1. IEEE International Conference «Video and audio signal processing in the context of Neurotechnologies – SPCN-2019» (27 - 31 мая 2019, Санкт-Петербург). Smirnova V., Labutina O., Gvozdeva A. Speech detection in spatially distributed speech-like noise. Стендовый доклад.

2. Международная конференция «Лингвистический форум 2020: Язык и искусственный интеллект» (12–14 ноября 2020 г. Институт языкознания РАН, Москва). Е.А. Огородникова, О.В. Лабутина, И.Г. Андреева, А.П. Гвоздева, Ю.А. Баулин. Фактор просодики в восприятии коммуникативной сцены с пространственным разделением источников речи и речеподобной помехи. Устный доклад.

3. III Всероссийская акустическая конференция (25–29 мая 2020 г., Санкт-Петербург). Огородникова Е.А., Пак С.П., Столярова Э.И., Лабутина О.В. Особенности восприятия конкурирующих речевых сигналов в норме и при нарушениях слухоречевой функции. Устный доклад.

4. Всероссийская конференция с международным участием «Интегративная физиология», посвящённая 95-летию Института физиологии им. И.П. Павлова РАН (СПб, 9-11 декабря 2020). Огородникова Е.А., Андреева И.Г., Гвоздева А.П., Лабутина О.В., Пак С.П. Фактор голоса в пространственной избирательности речевого слуха. Устный доклад.

5. Всероссийская конференция с международным участием «Интегративная физиология» (8-10 декабря 2021 года Санкт-Петербург). Институт физиологии им. И.П. Павлова РАН. О.В. Лабутина, Е.А. Огородникова, А.П. Гвоздева, И.Г. Андреева. Влияние несенсорных факторов на обнаружение речевого сигнала в условиях пространственно распределенного шума. Стендовый доклад.

Список литературы

1. Андреева И.Г. Пороговая длительность сигналов при восприятии человеком радиального движения звуковых образов различного спектрального состава // Сенсорные системы. 2004. Т. 18. № 3. С. 233.
2. Андреева И.Г. Пространственная избирательность слуха при распознавании речи на фоне речевого маскера. Физиология человека. 2018. №2. С.124.
3. Бобошко М.Ю., Голованова Л.Е., Жилинская Е.В., Огородникова Е.А. Разборчивость речи при тугоухости у лиц пожилого и старческого возраста // Успехи геронтологии. 2016. Т.29. №4. С.663-669.
4. Еркудов В.О., Огородникова Е.А., Пуговкин А.П., Сергеев И.В., Сляпцова Т.Н., Кундрат Е.О., Лупанова П.А., Пахомова М.А., Денисенко М.Д., Балашов А.Л. Выделение голоса целевого диктора в условиях речевой конкуренции у школьников с различным психоэмоциональным статусом. Педиатр. 2019. Т.10. №4. С.51–59.
5. Королева И.В., Огородникова Е.А., Пак С.П., Левин С.В., Балякова А.А., Шапорова А.В. Методические подходы к оценке динамики развития процессов слухоречевого восприятия у детей с кохлеарными имплантами. Российская оториноларингология. 2013. №3. С.75-85.
6. Люблинская В.В., Огородникова Е.А., Королева И.В., Пак С.П., Рыбаков М.В. Опыт использования компьютера при исследовании и тренировке слухоречевого восприятия у пациентов после кохлеарной имплантации. Речевые технологии. 2009. №1. С.43-53.
7. Малинина Е.С., Андреева И.Г. Слуховое последствие приближения и удаления источника звука: зависимость от траектории и области предъявления адаптирующих стимулов // Журн. эволюц. биохимии и физиол. 2013 Т. 49 № 3 С. 211-223.
8. Огородникова Е.А., Балякова А.А. Слуховой сегментный анализ при нарушении слуха, речи и письма. LAP-LAMBERT Academic Publishing, 2014.
9. Огородникова Е.А., Октябрьский В.П., Пак С.П., Столярова Э.И., Балякова А.А., Рыбаков М.В., Охарева Н.Г., Кузьмина Т.В. Использование программных средств для обучения слабослышащих и инофонон. Сенсорные системы. 2014. N4. С.22.
10. Огородникова Е.А., Андреева И.Г., Гвоздева А.П., Лабутина О.В., Пак С.П. Фактор голоса в пространственной избирательности речевого слуха. Тезисы докладов

Всероссийской конференции с международным участием «Интегративная физиология». СПб.: Ин-т физиологии им. И.П. Павлова РАН, 2020. С.66-67.

11. Altman J.A., Andreeva I.G. Monaural perception and binaural perception of approaching and withdrawing auditory images in humans // *Int. J. Audiol.* 2004. Apr. V. 43. № 4. P. 227.

12. Andreeva I.G., Dymnikowa M., Gvozdeva A.P., Ogorodnikova E.A., Pak S.P. Spatial separation benefit for speech detection in multi-talker babble-noise with different egocentric distances. *Acta Acustica united with Acustica.* 2019. V.105. N3. P.484.

13. Bee M.A., Micheyl C. The «Cocktail party problem»: What is it? How can it be solved? And why should animal behaviorists study it? *J. Comp. Psychol.* 2008. V.122. №3. P.235–251.

14. Beutelmann R., Brand T., Kollmeier B. Revision, extension, and evaluation of a binaural speech intelligibility model // *J. Acoust. Soc. Am.* 2010. V. 127. № 4. P. 2479.

15. Bidet-Caulet A., Bertrand O. Neurophysiological mechanisms involved in auditory perceptual organization. *Frontiers in Neuroscience.* 2009. V.3. P.182-191.

16. Bregman A.S. Auditory scene analysis: the perceptual organization of sound. Cambridge: MIT Press. 1990.

17. Bronkhost W. The Cocktail Party Phenomenon: A Review of Research on Speech Intelligibility in Multiple-Talker Condition. *Acustica.* 2000. V.86. P.117–128.

18. Bronkhorst A.W. The cocktail-party problem revisited: early processing and selection of multi-talker speech. *Atten. Percept. Psychophys.* 2015. V.77. N5. P.1465.

19. Effect of target-masker similarity on across-ear interference in a dichotic cocktail-party listening task. *J Acoust Soc Am* 122 (3) (2007) P. 1724.

20. Cherry E.C. Some experiments on the recognition of speech, with one and with two ears. *J. Acoust. Soc. Am.* 1953. V.25. N5. P.975.

21. Coleman P.D. Dual role of frequency spectrum in determination of auditory distance // *J. Acoust. Soc. Amer.* 1968. V. 44. P. 631.

22. Dietz M., McAlpine D. *Advancing Binaural Cochlear Implant Technology* // Trends Hear. UCL Ear Institute, London, UK, 2015.

23. Durlach N.I. Binaural signal detection: Equalization and cancellation theory // Ed. J.V. Tobias. *Foundations of modern auditory theory.* New York: Academic Press. 1972. P. 369.

24. Gutschalk A., Dykstra A.R. Functional imaging of auditory scene analysis. *Hear. Res.* 2014. V.307. P.98.
25. Kolarik A.J., Moore B.C.J., Zahorik P. et al. Auditory distance perception in humans: a review of cues, development, neuronal bases, and effects of sensory loss // *Atten. Percept. Psychophys.* 2016. V. 78. P. 373.
26. Lavandier M., Culling J.F. Speech segregation in rooms: monaural, binaural, and interacting effects of reverberation on target and interferer // *J. Acoust. Soc. Am.* 2008. V. 123. P. 2237.
27. Lavandier M., Jelfs S., Culling J.F. et al. Binaural prediction of speech intelligibility in reverberant rooms with multiple noise sources // *J. Acoust. Soc. Am.* 2012. V. 131. № 1. P. 218. *Audio Source Separation*, Ed: Makino, 2018.
28. Marrone N., Mason C.R., Kidd G. Jr. Tuning in the spatial dimension: evidence from a masked speech identification task // *J. Acoust. Soc. Am.* 2008. V.124. P. 1146.
29. Ogorodnikova E.A., Koroleva I.V., Lublinskaja V.V., Pak S.P., Stoljarova E.I., Baljakova A.A. Computer in rehabilitation of patients with cochlear implants. *Proc. of 13-th International Conference «Speech and Computer-SPECOM'2009»*. SPb: SPIIRAS, 2009. P.483-486.
30. Popper A.N., Fay R.R. (Eds). *Perspectives on auditory research / Springer handbook of auditory research.* 2014.
31. Shamma S.A., Elhilali M., Micheyl C. Temporal coherence and attention in auditory scene analysis. *Trends Neurosci.* 2011. V.34. P.114.
32. *The Auditory System at the Cocktail Party*. Eds: Middlebrooks, Simon, Popper, Fay, 2017.
33. Wan R., Durlach N., Colburn S., J. Application of an extended equalization-cancellation model to speech intelligibility with spatially distributed maskers. *Acoust. Soc. Am.*, 2010. 128 (6). P.3678.
34. Westermann A., Buchholz J. The influence of informational masking in reverberant, multi-talker environment. *J Acoust Soc Am* 2015. 138 (2). P.584.
35. Zahorik P., Brungart D.S., Bronkhorst A.W. Auditory distance perception in humans: A summary of past and present research // *Acta Acustica united with Acustica.* 2005. 91(3). P.409.