

МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г.
ЧЕРНЫШЕВСКОГО»

Кафедра общей и консультативной психологии
наименование кафедры

Психологические особенности аудиального восприятия текстов,

ОЗВУЧЕННЫХ ЕСТЕСТВЕННЫМ И СИНТЕЗИРОВАННЫМ ГОЛОСОМ

АВТОРЕФЕРАТ ВЫПУСКНОЙ КВАЛИФИКАЦИОННОЙ РАБОТЫ
МАГИСТРА

студента (ки) 3 курса 376 группы

направления (специальности)	37.04.01 «Психология» код и наименование направления (специальности)
--------------------------------	---

факультета психологии
наименование факультета, института, колледжа

Карпачевой Эльвиры Рафиковны
фамилия, имя, отчество

Научный руководитель
(руководитель)

Канд. биол. н., доцент

должность, уч. степень, уч. звание

ПОДПИСЬ, ДАТА

Е.М. Зинченко
инициалы, фамилия

Зав. кафедрой

Д. психол. н., доцент

должность, уч. степень, уч. звание

ПОДПИСЬ, дата

С.В. Фролова
инициалы, фамилия

Саратов 2025

Введение. Актуальность исследования обусловлена интенсивной интеграцией систем генерации текста и синтеза речи в повседневную жизнь при недостаточной изученности их психологического воздействия на личность. Сгенерированные искусственным интеллектом тексты (чат-боты с генеративным искусственным интеллектом) и синтезированная речь (в аудиокнигах, голосовых ассистентах, на образовательных платформах) становятся для молодежи ведущими каналами получения информации.

В связи с этим особую остроту приобретает следующий исследовательский вопрос: как происхождение текста (искусственное или антропогенное) и тип голосового сигнала (синтезированная или естественная речь) влияют на эффективность когнитивно-перцептивных процессов обработки информации (в особенности, научного характера) в образовательной среде?

Настоящее исследование направлено на изучение психологических особенностей аудиального восприятия текстов, озвученных естественным и синтезированным голосом, у основной целевой аудитории образовательных инноваций – учащихся среднего и старшего школьного возраста, студентов. Полученные данные могут составить основу для разработки программ психологической профилактики, направленных на повышение эффективности аудиального восприятия учащимися образовательного контента, а также психологических рекомендаций по оптимизации использования синтезированной речи в образовательной среде, что определяет актуальность данной темы исследования.

Объект исследования: восприятие как познавательный психический процесс.

Предмет исследования: психологические особенности аудиального восприятия текстов, озвученных естественным и синтезированным голосом.

Цель исследования: изучение психологических особенностей аудиального восприятия текстов, озвученных естественным и

синтезированным голосом, у учащихся среднего и старшего школьного возраста, студентов.

Задачи исследования:

1. Охарактеризовать возрастные особенности познавательных психических процессов.
2. Изучить категорию «семантическая связность текста» в контексте психолингвистических исследований.
3. Проанализировать отечественные и зарубежные психологические исследования особенностей аудиального восприятия текстов;
4. Рассмотреть психолингвистические критерии дифференциации «искусственных» и «естественных» текстов.
5. Разработать методический инструментарий и процедуру экспериментального исследования.
6. Проверить наличие или отсутствие значимых различий в психологических особенностях аудиального восприятия текстов, озвученных естественным и синтезированным голосом, на базе трех возрастных групп.
7. Разработать психологические рекомендации, направленные на повышение эффективности использования синтезированной речи в образовательной среде.

Гипотезы исследования:

1. Эффективность аудиального восприятия «естественных» научно-популярных текстов, воспроизводимых синтезированным голосом умной колонки, ниже у студентов по сравнению с учащимися среднего и старшего школьного возраста.
2. Эффективность аудиального восприятия по критерию дифференциации антропогенного или искусственного происхождения текста сказок ниже у учащихся среднего школьного возраста по сравнению с другими группами испытуемых.

Методы исследования:

– психологический эксперимент, направленный на оценивание психологических особенностей аудиального восприятия текстов, озвученных естественным (экспериментатор) и «синтезированным» голосом (умная колонка «Яндекс Станция Лайт 2» с голосовым помощником Алиса), у разных возрастных групп;

– анкетирование (вопросы открытого и закрытого типов) как методы сбора данных;

– качественные и количественные методы обработки данных (качественный контент-анализ ответов на вопросы анкеты открытого типа, математические методы статистической обработки данных).

К математическим методам статистической обработки данных относятся: критерий Шапиро–Уилка, апостериорный критерий Геймса-Хоуэлла с поправкой Холма.

Математико-статистические расчеты осуществлялись в лицензированной статистической программе JASP, v. 0.95.4.

Эмпирическая база исследования. Исследование проводилось в очном формате на базах ФГБОУ ВО «Саратовский национальный исследовательский государственный университет имени Н.Г. Чернышевского» (набор группы студентов) и МБОУ СОШ №52 г. Пенза (набор групп учащихся 5–8 классов, 9–11 классов).

Совокупную выборку исследования составили 45 испытуемых трех возрастных групп: 15 учащихся среднего школьного возраста (10–14 лет, 5–8 классы; женский пол – 7 испытуемых, мужской пол – 8 испытуемых), 15 учащихся старшего школьного возраста (15–17 лет, 9–11 классы; женский пол – 7 испытуемых, мужской пол – 8 испытуемых), 15 студентов-бакалавров факультета психологии «СГУ им. Н.Г. Чернышевского» (18–24 лет, 1–4 курсы бакалавриата; женский пол – 7 испытуемых, мужской пол – 8 испытуемых).

Теоретико-методологическую основу исследования составили фундаментальные труды по возрастной и когнитивной психологии Л.С. Выготского, Д.Б. Эльконина, Ж. Пиаже, Б.М. Величковского,

П.Я. Гальперина; классические и современные работы по психолингвистике и теории речевой деятельности А.А. Леонтьева, А.Р. Лурии, Н.И. Жинкина, И.А. Зимней, А.А. Залевской, Л.П. Доблаева; эмпирические исследования А.Ф. Пантелеева, О.А. Колпинской, Е.Г. Санниковой, К. Драгер, Т. Хогана, К. Насс и др., посвященные сравнительной оценке когнитивных особенностей личности при восприятии текстов или иных продуктов деятельности антропогенного и искусственного происхождения.

Теоретическая значимость и научная новизна исследования состоят в разработке системы критериев оценивания эффективности аудиального восприятия текстов, выявлении значимых различий в психологических особенностях аудиального восприятия текстов, озвученных естественным и синтезированным голосом, учащимися среднего и старшего школьного возраста, студентами.

Практическая значимость исследования заключается в возможности использования полученных результатов в целях повышения эффективности аудиального восприятия учащимися образовательного контента, дифференциации ими текстов по критерию антропогенного или искусственного происхождения, оптимизации применения технологий синтезированной речи в образовательной среде с учетом возрастных особенностей аудиального восприятия целевой аудитории.

Опыт апробации результатов исследования. По результатам исследования, проведенного в рамках подготовки выпускной квалификационной работы магистра, была опубликована научная статья:

– Карпочева Э.Р. Сравнительный анализ особенностей аудиального восприятия нарративов, озвученных человеком и искусственным интеллектом: когнитивные аспекты восприятия у детей и студентов // Сборник материалов XV ежегодной научно-практической конференции молодых ученых факультета психологии «Проблемы современной психологии: теория, практика, эксперимент» (21 февраля 2025 г.) – Саратов, 2025. С.52-54.

Основная часть. По результатам проведенного нами анализа отечественной и зарубежной научной литературы, включая классические труды и публикации, изданные за последние десять лет, можно сформулировать следующие выводы:

1. Возраст выступает ключевым фактором, опосредующим базовые познавательные возможности личности, которые лежат в основе эффективного аудиального восприятия как естественных, так и синтезированных текстов.

Теоретический анализ, опирающийся на периодизацию психического развития Д.Б. Эльконина и исследования в области возрастной психологии, показал, что для младших подростков характерна незавершенность перехода к формально-логическим операциям, относительно низкий уровень произвольности внимания и смыслового запоминания, что делает их восприятие более зависимым от внешних, поверхностных характеристик вербального сообщения.

Старшие подростки обладают уже сформированным абстрактным мышлением и развитой рабочей памятью, что позволяет им активно конструировать смысл и выстраивать внутренний диалог с текстом. Когнитивное развитие студентов достигает своего «акме», обеспечивая им высокий уровень метакогнитивного контроля, избирательности и критичности в восприятии информации.

2. Семантическая связность текста является ведущим психологическим условием его успешного понимания, обеспечиваемым комплексом когнитивных операций по интеграции смысловых элементов в единую ментальную репрезентацию, причем эффективность этих операций имеет выраженную возрастную детерминацию.

Анализ отечественных психолингвистических концепций (Л.П. Доблаев, И.А. Зимняя, А.А. Леонтьев) свидетельствует о том, что связность не является имманентным свойством текста, а возникает в процессе его смыслового восприятия как результат деятельности читателя или слушателя

по установлению логико-смысловых отношений между пропозициями. Эта деятельность опирается на операции рабочей памяти, антиципации и смыслового обобщения.

Для младших подростков установление глубинных связей представляет большую когнитивную сложность, тогда как старшая аудитория способна реконструировать имплицитные связи и оценивать целостность замысла. В контексте восприятия «искусственных» текстов связность становится противоречивым параметром: несмотря на высокую формальную когезию, в них могут наблюдаться смысловые разрывы или шаблонность переходов, которые могут быть выявлены на когнитивном уровне подготовленными слушателями.

3. Современные эмпирические исследования, преимущественно зарубежные, констатируют быстрое сближение объективных акустических параметров синтезированной и естественной речи, но одновременно подтверждают устойчивость субъективно-психологических различий в их восприятии, касающихся доверия к источнику информации, воспринимаемой естественности и социальной привлекательности голоса, а также глубины переработки контента, сгенерированного искусственным интеллектом.

Обзор эмпирических работ последних лет позволил выявить парадоксальную тенденцию: несмотря на то, что системы синтеза речи (TTS) проходят «слепые» тесты на естественность, методы нейровизуализации фиксируют иную активность мозга в ответ на синтезированные стимулы, а психологические эксперименты показывают стойкое предпочтение и большее доверие к человеческим голосам и текстам.

Психологическими маркерами искусственного происхождения контента выступают, прежде всего, субъективно переживаемые ощущения безличности текста, шаблонности, эмоциональной обедненности и прагматической неадекватности. Однако остается неисследованным комплексное влияние фактора происхождения текста (человек / искусственный интеллект) и типа его озвучивания (естественный / синтезированный голос) на эти субъективные

оценки в разных возрастных группах, что определяет научную новизну настоящего исследования.

4. Дифференциация текстов антропогенного и искусственного происхождения представляет собой сложный психолингвистический процесс, в котором формальные лингвистические маркеры (просодические аномалии, лексико-синтаксическая шаблонность) взаимодействуют с психологическими критериями, основанными на особенностях социального познания, метакогнитивного чувства подлинности и эмоционального резонанса.

Теоретический анализ современных публикаций позволил выделить два взаимосвязанных уровня критериев. На лингвистическом уровне это статистические отклонения в просодии и предсказуемость лексико-грамматических структур, выявленные в современных компьютерно-лингвистических исследованиях. На психологическом уровне ключевую роль играют феномены атрибуции авторства, оценки доверия и компетентности источника (следовательно, и автора) текста, а также сензитивность к тонким социально-эмоциональным сигналам в голосе.

Способность к такой дифференциации не является врожденной; она развивается с возрастом и когнитивным опытом, опираясь на сформированность критического мышления, объем фоновых знаний и социального интеллекта.

По результатам проведенного нами эмпирического исследования можно сформулировать следующие выводы:

1. Успешно разработана и апробирована комплексная система критериев для оценки эффективности аудиального восприятия «искусственных» и «антропогенных» текстов (озвученных синтезированным и естественным голосом), включающая многоуровневые показатели (от глобального и детального понимания воспринимаемого текста до метакогнитивного распознавания его авторства).

Разработанный методический инструментарий (процедура эксперимента, стимульный материал, оценочная шкала) позволил

дифференцированно оценить вклад различных когнитивных процессов в эффективность аудиального восприятия текста и выявить статистически значимые различия между возрастными группами.

2. Первая выдвинутая нами гипотеза исследования подтвердилась в полном объеме. Установлено, что эффективность аудиального восприятия научно-популярных текстов, озвученных синтезированным голосом, ниже у студентов по сравнению с учащимися среднего и старшего школьного возраста.

Так, значимые различия между возрастными группами обнаружены по следующим критериям эффективности аудиального восприятия текста: глобальному, детальному и терминологическому пониманию, а также по выраженности показателя субъективного оценивания доступности текста для понимания.

Это свидетельствует о наличии возрастнo-специфических когнитивных стратегий обработки «обедненного» синтезированного речевого сигнала при восприятии фактологического контента (научно-популярной направленности).

3. Вторая гипотеза нашего исследования подтвердилась частично: учащиеся среднего школьного возраста (5–8 классы) значительно уступают студентам в способности дифференцировать тексты, созданные человеком и нейросетью, однако данная тенденция характерна и для группы учащихся 9–11 классов. При этом между учащимися среднего и старшего школьного звена значимые различия по данному показателю аудиального восприятия отсутствуют.

Дополнительно мы обнаружили, что репродуктивная память (которую мы исследовали на основе анализа воспроизведения испытуемыми ключевых тематических блоков нарративных текстов) развита в большей степени у студентов, чем у обеих групп школьников.

Полученные результаты указывают на то, что способность к критическому анализу дискурса и метакогнитивной оценке происхождения

информации формируется и достигает оптимального уровня развития лишь к юношескому возрасту, оставаясь недостаточной в подростковом.

4. На основе результатов эмпирического исследования был разработан комплекс дифференцированных психолого-педагогических рекомендаций.

Рекомендации носят адресный характер и направлены на:

- а) оптимизацию разработки учебных материалов, содержащих синтезированную речь, с учетом «когнитивных профилей» разных возрастов;
- б) целенаправленное развитие у учащихся навыков активного смыслового слушания, метакогнитивного мониторинга воспринимаемых текстов и критического анализа цифрового дискурса.

Практическая ценность рекомендаций заключается в возможности их использования для повышения эффективности образовательного процесса и медиаграмотности учащихся в условиях цифровой трансформации обучения.

Заключение. Настоящее исследование было направлено на изучение психологических особенностей аудиального восприятия текстов, озвученных естественным и синтезированным голосом, у учащихся среднего и старшего школьного возраста, студентов.

В рамках теоретического исследования были решены следующие задачи и получены соответствующие выводы:

1. Проанализированы возрастные особенности познавательных психических процессов, что позволило обосновать выбор репрезентативных возрастных групп для эмпирического исследования и сформулировать гипотезы об ожидаемых различиях в особенностях аудиального восприятия, основанные на закономерностях когнитивного развития личности.

2. Изучена категория семантической связности текста, позволившая раскрыть психолингвистические механизмы понимания текста и определить связность как ключевой параметр, чувствительный к происхождению текста (искусственному или антропогенному).

3. Проведен обзор современных отечественных и зарубежных работ по проблеме исследования, по результатам которого можно сделать вывод о

существовании различий в восприятии естественной и синтезированной речи, несмотря на технологическое сближение их акустических параметров благодаря научно-техническому прогрессу.

4. Систематизированы психолингвистические критерии дифференциации искусственных и естественных текстов, что легло в основу разработки критерия эффективности аудиального восприятия текста – метакогнитивного распознавания его авторства (антропогенного или искусственного происхождения).

По результатам эмпирического исследования можно сделать следующие выводы:

1. Был успешно разработан оригинальный методический инструментарий, включающий процедуру эксперимента, стимульный материал и систему многоуровневых критериев оценивания эффективности аудиального восприятия текста (от декодирования до метапознания).

2. Полностью подтверждена первая гипотеза исследования: установлено, что эффективность восприятия научно-популярных текстов, озвученных синтезированным голосом, ниже у студентов по сравнению с группами школьников, что указывает на нелинейный характер влияния возрастного фактора на эффективность аудиального восприятия, возможно, связанного с разными когнитивными стратегиями обработки «обедненного» искусственного речевого сигнала.

3. Вторая гипотеза исследования подтвердилась частично: выявлено, что способность к распознаванию авторства текстов (искусственное или антропогенное происхождение) значимо ниже у учащихся школьного возраста, что свидетельствует о более позднем формировании комплекса метакогнитивных и критических навыков, необходимых для анализа происхождения дискурса.

Дополнительно мы обнаружили, что репродуктивная память (которую мы исследовали на основе анализа воспроизведения испытуемыми ключевых

тематических блоков нарративных текстов) развита в большей степени у студентов, чем у обеих групп школьников.

4. На основе полученных результатов был разработан комплекс дифференцированных психолого-педагогических рекомендаций, направленных на оптимизацию использования синтезированной речи в образовательной среде и развитие у учащихся навыков активного слушания, критической оценки искусственного контента с учетом их возрастных когнитивных профилей.

К основным ограничениям работы относится малый объем выборки, что требует осторожности при экстраполяции выводов исследования на генеральную совокупность. Также в рамках экспериментального исследования нами использовался ограниченный набор жанров стимульного материала (сказки и научно-популярные статьи), оставляющий открытым вопрос о восприятии синтезированной речи в иных коммуникативных контекстах (например, лекции, инструкции, диалоги).

К перспективам развития нашего исследования можно отнести:

1. Изучение личностных и когнитивных предикторов выявленных возрастных различий аудиального восприятия синтезированной речи, таких как когнитивные стили, уровень цифровой грамотности, отношение к технологиям.

2. Расширение методического инструментария исследования за счет включения в него психофизиологических методов (например, ЭЭГ, фНIRS, айтрекинг) для объективной регистрации когнитивной нагрузки и эмоционального отклика субъекта в процессе восприятия стимулов естественного и искусственного происхождения, что позволит построить более целостную нейрокогнитивную модель воздействия синтезированной речи на человека.