

ЯЗЫКОВОЙ ОБРАЗ В МОДЕЛИРОВАНИИ «ЦИФРОВЫХ ДВОЙНИКОВ»

LINGUISTIC IMAGE IN DIGITAL TWINS MODELLING

**A. Gorodishchev
S. Uskova
E. Khodenkova**

Summary: This paper deals with the tasks of modelling the linguistic image of artificial intelligence (AI) in the digital twin's creation. The authors raise the issue that in order to create the speech of a digital twin, it is necessary not only to fill its vocabulary, but the words must be pronounced in the right voice with the right intonation and speed, accompanied by pictures, inscriptions, etc. The totality of these factors will form the perception of the message model as a whole. The high role of phonological theories as a basis for creating realistic speech simulators and speech recognition systems is outlined.

Keywords: artificial intelligence, linguistic image, phonetics, semantics, digital twin, sound, phonology.

Городищев Алексей Викторович

Кандидат филологических наук, доцент, Университет
науки и технологий МИСИС (г. Москва)
agorodischev@gmail.com

Ускова Светлана Викторовна

Кандидат филологических наук, доцент, Сибирский
государственный университет науки и технологий
имени академика М.Ф. Решетнева (г. Красноярск)
s_us78@mail.ru

Ходенкова Эльга Владимировна

Кандидат философских наук, доцент, Сибирский
государственный университет науки и технологий
имени академика М.Ф. Решетнева (г. Красноярск)
ellie4ka@mail.ru

Аннотация: В данной статье рассматриваются вопросы моделирования языкового образа искусственного интеллекта (далее ИИ) при создании цифровых двойников. Авторы поднимают вопрос о том, что для создания речи цифрового двойника необходимо не только наполнить его лексикон, но слова должны быть произнесены правильным голосом с нужной интонацией и скоростью, сопровождаться картинками, надписями и др. Совокупность данных факторов будет формировать восприятие модели сообщения в целом. Обозначена высокая роль фонологических теорий как основы для создания реалистичных речевых симуляторов, систем распознавания речи.

Ключевые слова: искусственный интеллект, языковой образ, фонетика, семантика, цифровой двойник, звук, фонология.

Появление новых цифровых технологий все чаще заставляет пересматривать отношение к исследованию языка. Если цифровые аватары и цифровые двойники человека, созданные с помощью оцифровки физических параметров человека, имеют практически идентичные визуальные характеристики живого человека, то отличить речь робота от речи живого человека может даже обычный человек.

В рамках восприятия концепции цифровых двойников исследователи (Zhihan, 2022) предлагают пересмотреть теорию с точки зрения не только материальной вселенной, но и результатов интеллектуального труда и творческой деятельности. В частности, в качестве экстралингвистических особенностей речи рассматривается воздействие звука. Например, если принять позицию, что песня как материальное выражение чувств композитора, способна описать связь между разными эмоциональными явлениями через определенные гармонические сочетания, то и любовь, страсть, ненависть и прочее возможно описать по отдельным языковым схемам, которые соответствуют звуковым схемам традиционной музыки. Отметим, что подобные эксперименты проводи-

лись и в нецифровую эпоху. Например, дадаисты пытались «выйти за пределы сложившегося понимания мира через фонетические, грамматические и семантические механизмы человеческой речи» (Григорьев, 2015), что позволяло им провоцировать публику, заставляя искать новое понимание смысла произведений искусства, даже если материал не меняется (Григорьев, 2015). При таком подходе в современной науке появилась теория цифрового искусства, как постепенного перехода от реального объекта к его изображению и последующему словесному описанию. Виртуальная среда же принимается как арт-среда, используемая художниками для создания артефактов цифрового искусства (Григорьев, 2015).

Моделирование цифровых двойников возможно не только в рамках теорий цифрового искусства, но и в рамках волновой теории, где звук как волна воздействует на человека и вызывает нужные чувства. При данном подходе применяются технологии, разработанные и используемые в маркетинге продаж в больших центрах и кафе. Моделируют с помощью звуковых волн и «когнитивный дурман», когда «виртуальные технологии разлагают мелодию на разные звуковые диапазоны, после чего зву-

ковой сигнал подается через стереонаушники на правое и левое ухо с бинауральными различиями по частоте, т.е. асинхронно... навязываемая биполярная звуковая атака приводит к тому, что мозговая нейросистема под воздействием фонового дисбаланса приходит в состояние когнитивной турбулентности», что, по мнению Р.Я. Подоль, способно вызвать психофизиологическую эйфорию, подобную наркотическому опьянению (Подоль, 2023). Подобная технология моделирования служит для подбора наиболее эффективного акустического воздействия при стимуляции тех или иных эмоциональных порывов. Звуковые волны могут подавлять когнитивные способности, снижает память, нарушать рефлексивную функцию мышления, что использовалось как в прошлом (технологии Геббельса в пропаганде), так и в современных деструктивных «колумбайн-сообществах», когда для усиления воздействия на аудиторию использовались резкие звуки, призывы, психоделическая музыка (Баева, 2019). Для воздействия на большую аудиторию выбирались наиболее близкие к нужной волновой модели словосочетания, синонимы в речи, и т.п.

В современном мире не обязательно, чтобы выбранную речь произносил человек. Например, стартап Air.ai (<https://www.air.ai/>) продемонстрировал, что человек зачастую не может отличить хорошо синтезированный голос и адаптированную речь искусственного интеллекта от обычного человека, хотя ей порой не хватает эмоциональности. Для создания цифрового двойника или аватара человека недостаточно «научить» его произносить слова, они должны быть произнесены с правильной интонацией, правильным голосом и даже с нужной скоростью, так как это будет формировать модель восприятия сообщения в целом.

Подготовка к нормализации восприятия «цифровых аватаров» велась с начала эры компьютерных игр. Если до этого энтузиасты реконструировали исторические события, то теперь появилась возможность, как с героями комиксов, переносить части цифровых моделей в реальный мир. Появилась практика экспериментов с костюмами, подражания моделям поведения, общения и даже образа законов реконструируемого общества. Особый вид практики – маркетинговые исследования, которые используют специалисты для формирования лояльности бренда. Маркетологи знают, что мужской голос воспринимается с большим «доверием» (Дударева), он должен быть ещё и достаточно низким, внушая картинку «скалы», твердой опоры и т.д. А вот если наша задача «расслабить» аудиторию и сделать её более лояльной на долгое время лучше использовать женские «паттерны». Чувство страха является наиболее острым, но при этом, наиболее недолговечным, поэтому звуки, соответствующие запретным темам культуры, также используются.

Многие вопросы применения звука как фактора воз-

действия на аудиторию, можно решить, используя разработки лингвистов, которые в рамках фонологических/фонетических исследований текстов и речи выявили, что не изменение звука, а изменение функций звука отличают одну эпоху от другой (Курулёнок, 2022). Анализ фонетических изменений, основанных на отношениях фонем в разные периоды истории, дают целостное понимание истории языка, и такое понимание может дать основу и для понимания программирования языка цифровой персоны.

В работах по исторической фонетике Р.И. Аванесова рассматриваются связи между вокализмом, консонантизмом, редуцированных звуков, носовых гласных, переход <e> в <o> и проч. Через выстраивание данных связей показано, что звуковое явление должно рассматриваться в рамках существующей фонетической системы и историей ее эволюции (Курулёнок, 2022). Лингвистические методы, как в синхронном, так и в диахронном подходах к истолкованию языковых явлений, могут быть полезными при создании алгоритма генерации речи. Так сосюрковский синхронный подход позволяет выделить системность языка и речи, тем самым определить отношения фонем в системе «набор её дифференциальных признаков, устанавливаются корреляции между фонемами, выясняется поведение фонемы, условия её возможных изменений в разных позициях» (Курулёнок, 2022). Л.В. Щерба выделяет фонемы, «которые имеют смысловоразличительное значение», и оттенки (аллофоны, комбинаторного варианты фонемы), «не имеющие его» (Попов, 2015).

Исследования в области ораторского искусства и воздействия на публику показывают, что человеческий голос подвержен многим факторам настроению, внутреннему и внешнему состоянию организма, социальному окружению и т.д. В результате физиологических изменений происходит перестройка системы фонем – единиц, способных дифференцировать слова и их формы, участвовать в образовании звукового облика значимой единицы языка морфемы, слова (Ленинградская фонологическая школа). Реализуясь в звуке, фонемы отражают определенные смыслы слова или выражения, которые свойственны той или иной культурной эпохе, настроению или физиологическому состоянию человека. Так, по частотности фонем можно определять авторство или принадлежность человека к определенной языковой группе, определять некоторые заболевания (Чойнзонов, 2022). И, на наш взгляд, именно фонологические теории могут стать основой для создания реалистичных речевых симуляторов, систем распознавания речи, потому что в искусственной системе задать корреляции между фонемами проще, чем заставить следовать правилам живого человека.

Искусственный интеллект, в отличие от человека, в

состоянии применять практически «идеальную» модель голосового синтеза для достижения нужного результата. Вычислив необходимый стиль речи для того или иного разговора, ИИ может самостоятельно подстроить волновую модель. А при учёте возможности аналитики огромного пласта материалов выдающихся политиков и ораторов, подобрать, в том числе и необходимые синонимы с большей скоростью и под нужный культурный код. При проработке любого языкового фреймворка можно задавать как тональность высказывания, так и учитывать особенности звукового фона собеседника, когда фонема – «как общественно признаваемое в рамках той или иной языковой общности (национального языка или (даже) его варианта) представление о (единичном, элементарном – опять же, в пределах общности) звуке» (Курдюмов, 2023). Например, если цифровой аватар создается для коммуникации с носителем русского языка, то необходимо учитывать звуковой строй русского языка, акустически понимаемый как «правильный». А для собеседников, понимающих русский язык, но не являющихся его носителями, звуки могут быть изменены, в соответствие с акустическим идеалом их родного языка (если в родном языке отсутствуют некоторые звуки, то для речи аватара надо создать алгоритм подбора синонимов без этих звуков). Но это только предположение, так как в некоторых языках, например в китайском, фонема осознаётся как составная часть (заполнитель позиции) морфемы и «варианты звука осознаются как соотносимые с той или иной морфемой именно по единству слога (слогоморфемы), в то время как русском языке фонемы, произнесённые с (иностраным, региональным) акцентом, отождествляются с идеалом звукотипа» (Курдюмов, 2023). Следовательно, выбирая информативные параметры для создания цифрового аватара, придется учитывать и фонетические особенности языка.

В случае с цифровыми аватарами или цифровыми двойниками важными являются не только вербальные коммуникации через голос, но и картинка, надписи, выделение определенных слов на заднем плане и даже передача слов и посланий через определенное сочетание предметов. Прекрасная картинка и генерация заднего фона и окружения может осуществляться «на лету», что, в свою очередь, даёт огромное пространство для вшивания так называемых «скрытых посланий» текстом, «25 кадром», внутри которых возможно разместить всё что угодно, то, что закрыто от концентрированного внимания и призвано нести дополнительное послание на более глубинном уровне (Arunesh Mathu, 2021). Здесь применяется прием «теневых паттернов» (Dark Pattern), когда буквы и слова не воспринимаются в чистом виде, но в сознании активируются ассоциативные связи с образом букв, которые позволяют «читать» текст. Такой «текст» прочитывается условно, потому что основные смыслы не читаются, а достраиваются сознанием на основании предварительного знания или маркеров, пре-

доставленных авторами текста. Согласно создаваемой сообществом HCI (Human-interface interaction) архитектуре проблема включения технологии «теневых паттернов» возникла не сразу, но с развитием технологии цифровой коммуникации начала приобретать всё более и более распространенный характер.

Практически это классические приемы работы с медиатекстами. Креализованные медиатексты тексты с картинками и звуком, в большинстве случаев, подчиняются принятым в данное время в обществе социальным моделям, и поэтому легко формализуются и поддаются программированию. Однако простое решение при высокой информационной культуре пользователей приводит к интересным последствиям. Пользователи начинают самостоятельно выделять, какие слова, действия, фреймы и т.д. вмешиваются в их цепочку принятия решений или взаимодействия с программами (существуют целые корпуса, собранные пользователями (Colin M. Gray, 2018). То есть пользователи замечают все «зашитые» разработчиками и маркетологами технологии, что показывает несовершенство цифровой технологии и отталкивает их от взаимодействия с цифровым аватаром.

В сегодняшней ситуации данные приемы реализуются в форме обратной связи через пользовательский чат или беседу в аккаунте социальной сети (мошеннические технологии работы с текстами в тему данной работы не входят), где используется технология чат-ботов (см., цифровой аватар). Они могут быть обезличены – просто сообщения на экране, или персонифицированы изображением, могут быть текстовыми или голосовыми помощниками. «Творцом сценария» во всех случаях является человек. Когда к разговору подключается ИИ, то ответы на вопросы, заданные пользователем, будут из тех же источников, которые использует человек – сайт магазина, поисковый сервис, база данных и т.п. ИИ ищет в тех же поисковых системах, где проблема с «чистотой данных» и релевантностью решается не качеством контента и достоверностью, а частотой упоминания и количеством обращений. Тест, генерируемый ботом с ИИ, будет порождать высказывания, которые в итоге станут вирусными, внушающими доверие толпе, но абсолютно недостоверными. Тем не менее, за счёт огромной силы влиять на простые умы, эти слова и фразы будут использоваться в той самой концепции «теневых паттернов». Следовательно, при программировании ботов также необходимы как отдельные социальные модели, так и отдельные лингвистические исследования.

Так как цифровые аватары (цифровые двойники живых инфлюенсеров) являются образом собирательным и с динамически изменяемыми характеристиками, схема, по которой выстраивается речевое (устное и графическое) общение с пользователем также носит собирательный характер. Так, система не может самостоятельно

использовать синоним из другой области знаний, при условии, что у неё заданы критерии точного соответствия. ИИ очень сильно ограничен законами, которые закладываются в него при проектировании и, в отличие от человека, при условии отсутствия сценария, который он может использовать для основы, создать «собираемый» сценарий на основе отдельных компонентов той или иной части языковой ситуации он, на текущий момент, не способен.

ИИ должен также воспринимать и этические законы и нормы, которые разрабатываются во многих странах мира, но пока еще не везде приняты. В России еще в 2021 году был подписан «Кодекс этики искусственного интеллекта» (авторы кодекса - Альянс в сфере искусственного интеллекта РФ, Аналитический центр при правительстве РФ и Минэкономразвития), который определил, что одной из ключевых задач технологий ИИ должно стать «содействие культурному и языковому многообразию» (Кодекс этики в сфере искусственного интеллекта), которую невозможно решить без применения теорий и методов разработки стилистики языка, этимологии и семиотики.

В этом направлении возможны сценарии развития технологии ИИ. Первый, определение «специализации» для цифрового аватара и построение матрицы опорных

языковых значений по профилю (вокабуляр и терминологическая область, в которой ИИ будет искать синонимы). Данный аватар может применяться в любом информационном проекте «по профилю». Второй сценарий предполагает выбор социальной сети и построение матрицы опорных языковых значений цифрового аватара под профиль сети. Например, для VK (Социальная сеть «ВКонтакте») и Грустнограм (Социальная сеть «Грустнограм») настройки коммуникации «лидеров обсуждения» первичных модераторов живой беседы, которые и формируют большую часть рейтинговости тех или иных кусков контента, стимулируя обсуждения, будут сильно отличаться. Настройки такого аватара должны соответствовать параметрам т.н. «цифровых активистов» потому что именно ветки обсуждений-комментариев модерлируемые такими личностями часто становятся более активными, чем изначальный материал.

Подводя итоги, подчеркнем, что моделирование адекватного языкового образа цифрового двойника (аватара), как достаточно сложной системы параметров и регуляции, потребует подключение к работе не только лингвистов, филологов, философов, но и тех, кто будет переводить с языка гуманитарных наук на технически понятный, а это уже дизайнеры цифровых систем совсем с другим уровнем организации мышления.

ЛИТЕРАТУРА

1. Баева Л.В. «Группы смерти» и «колумбайн-сообщества» в онлайн-культуре и реальном социуме // Информационное общество. – 2019. – № 3. – С. 33–42.
2. Григорьев А.Д. Развитие концепций цифрового искусства / А.Д. Григорьев, Т.А. Захарченко // Вестник ОГУ. 2015. №5 (180). URL: <https://cyberleninka.ru/article/n/razvitie-kontseptsii-tsifrovogo-iskusstva>.
3. Дударева А. Влияние гендерных стереотипов на восприятие голоса диктора в рекламе URL: <http://www.advlab.ru/articles/article172.htm>
4. Кодекс этики в сфере искусственного интеллекта URL: <https://goo.su/jax0B>
5. Курдюмов В.А. Понятия фонемы и аллофона применительно к китайскому языку (с учётом опыта преподавания в языковой среде гоюй) THE CONCEPTS OF PHONEME AND ALLOPHONE IN THE CHINESE LANGUAGE (based on the experience of teaching in the Guoyu language environment) // ОБУЧЕНИЕ КИТАЙСКОМУ ЯЗЫКУ И КУЛЬТУРЕ: ПРОШЛОЕ, НАСТОЯЩЕЕ, БУДУЩЕЕ: сборник научных статей II международной конференции / Редколлегия: М.А. Смолова, А.В. Попова, О.И. Короленко, В.Ю. Вашквичус, А.А. Казанцева, О.В. Павлова, О.А. Малых, Д.А. Коробко. – М.: Языки Народов Мира, 2023. – 79–88 с
6. Курулёнок А.А. К вопросу фонологического описания фонетической системы русского языка на разных этапах его развития // Русский язык: прошлое, настоящее, будущее (аннотированные доклады, презентации, статьи) / отв. ред. С.В. Лесников. СПб., 2022. ISBN 978-5-9903886-2-8944. ISSN 2713-2013. №09 (сентябрь). 2022. С.141–144.
7. Ленинградская фонологическая школа. URL: <https://fonetica.philol.msu.ru/bibl/lfsh.htm>
8. Подоль Р.Я. Когнитивный инфантилизм в рефлексии ноосферного сознания // Вестник ИвГУ. Серия: Гуманитарные науки. 2023. №2. URL: <https://cyberleninka.ru/article/n/kognitivnyy-infantilizm-v-refleksii-noosfernogo-soznaniya>.
9. Попов М.Б. ИА Бодуэн де Куртенэ, ЛВ Щерба и возникновение фонологии // ИА Бодуэн де Куртенэ и мировая лингвистика. Международная конференция (V Бодуэновские чтения): Труды и материалы. Том 1. Под общей редакцией К.Р. Галиуллина, Е.А. Горобец, Г.А. Николаева. 2015. Издательство: Казанский (Приволжский) федеральный университет (Казань) – 2015. – С. 276–278.
10. Социальная сеть «ВКонтакте». URL: <https://vk.com/>
11. Социальная сеть «Грустнограм». URL: <https://grustnogram.ru/>
12. Чойнзонов Е.Л. Валидация акустического индекса качества голоса для русского языка / Е.Л. Чойнзонов, Л.А. Кононова, Е.А. Красавина, Д.Е. Кульбакин, В.И. Мухина // Вестник оториноларингологии. 2022;87(5):34–38. <https://doi.org/10.17116/otorino20228705134>
13. Arunesh Mathur, Mihir Kshirsagar, and Jonathan Mayer. 2021. What Makes a Dark Pattern... Dark? Design Attributes, Normative Considerations, and Measurement Methods. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21). Association for Computing Machinery, New York, NY, USA, Article 360, 1–18. <https://doi.org/10.1145/3411764.3445610>

14. Colin M. Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L. Toombs. 2018. The Dark (Patterns) Side of UX Design. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (Montreal QC, Canada) (CHI '18). ACM, New York, NY, USA, Article 534, 14 pages. <https://doi.org/10.1145/3173574.3174108>
15. Zhihan LV1, Shuxuan XIE2, Yuxi LI2, M. Shamim HOSSAIN3, Abdulmotaieb El SADDIK4 Building the metaverse using digital twins at all scales, states, and relations - December 2022 Vol. 4 No. 6: 459–470 10.1016/j.vrih.2022.06.005

© Городищев Алексей Викторович (agorodischev@gmail.com), Ускова Светлана Викторовна (s_us78@mail.ru),
Ходенкова Эльга Владимировна (ellie4ka@mail.ru).

Журнал «Современная наука: актуальные проблемы теории и практики»

