

НЕЙРОСЕТЕВОЙ АУДИОВИЗУАЛЬНЫЙ ПЕРЕВОД

NEURAL NETWORK
AUDIOVISUAL TRANSLATION**O. Danchuk
A. Serebrova**

Summary: In this article the peculiarities of neural network usage in linguistics in audiovisual translation, is considered. The video translation from English to Russian with Yandex Browser is analyzed. During the research the main mistakes made by Artificial Intelligence on the verbal and paraverbal levels are identified and classified. There are grammatical, syntactical and lexical-semantic mistakes and the violation of kinesic and articulatory synchrony and isochronism.

Keywords: machine translation, neural network, audiovisual translation, Yandex.

Данчук Ольга Васильевна

Кандидат культурологии, доцент, Санкт-Петербургский
Гуманитарный университет профсоюзов
odanchuk@yandex.ru

Сереброва Анастасия Вячеславовна

Санкт-Петербургский Гуманитарный университет
профсоюзов
anastasiya.20102004@gmail.com

Аннотация: В данной статье рассматриваются особенности использования нейронных сетей в лингвистике, в частности в аудиовизуальном переводе. Анализируется перевод видео, выполненный Яндекс Браузером с английского языка на русский. В ходе исследования выявлены и классифицированы основные ошибки, допущенные искусственным интеллектом на вербальном и паравербальном уровнях. Среди них выделены грамматические, синтаксические и лексико-семантические неточности, а также нарушение кинесической и артикуляционной синхронности и изохронии.

Ключевые слова: машинный перевод, нейросеть, аудиовизуальный перевод, Яндекс.

Введение

По данным исследовательской компании Mediascope в 2024 году в структуре интернет потребления россиян преобладал видеоконтент (20% от общего количества), среднее количество человек, посещавших видеохостинг YouTube в 2024 составило в среднем 45% от общего населения России. Приведённые данные показывают, что аудиовизуальный формат сообщения становится преобладающим в структуре потребления пользователей. В связи с развитием международных контактов и появлением общедоступных площадок, таких как видеохостинг YouTube, возникает необходимость перевода большого количества видеоматериала, что стало возможным благодаря появлению нейросетевого перевода.

Цель исследования – провести сравнительно-сопоставительный анализ видео и его перевода, выполненный нейросетевыми технологиями Яндекс Браузера версии 25.2.0.2123.

Для достижения цели необходимо решить ряд задач: проанализировать особенности использования нейронных сетей в лингвистике, определить особенности войсовера как одного из видов аудиовизуального перевода (АВП), выявить и классифицировать ошибки, допущенные при переводе Яндекс Браузером.

В исследовании использовались следующие методы: сравнительно-сопоставительный анализ, метод сплошной выборки (для анализа видео и транскрипта его пере-

вода), контекстный анализ, элементы количественного анализа, классификация.

Материал исследования – англоязычное выступление австралийского предпринимателя Скотта Болланда (Scott Bolland) «Neuroscience, AI and the Future of Education» на конференции TEDx [15] и его войсовер, выполненный Яндекс Браузером. Общий хронометраж – 15 минут 36 секунд. Видеозапись размещена на канале Tedx Talks на видеохостинге YouTube (более 42 млн подписчиков) и на момент исследования его посмотрели более 170 тысяч раз.

Теоретическую базу исследования составили работы по особенностям нейронных сетей и их использованию в лингвистике [2, 10, 13], об особенностях АВП и войсовера как его отдельного вида [6, 7, 16], работы, посвященные исследованию переводческих трансформаций и переводческих ошибок [1, 9].

Обсуждение и результаты

Начало развития нейронным сетям положено в 1943 году американским нейрофизиологом Уорреном МакКаллоком (Warren McCulloch) и его коллегой математиком Уолтером Питтсом (Walter Pitts). Они разработали математическую модель искусственных нейронов и предложили метод их объединения в искусственные нейронные сети [13, с. 29–30]. Постепенно совершенствуется архитектура, алгоритмы и методы обучения нейронных сетей, для чего создаются программные системы/библиотеки.

Сфера применения нейросетей в лингвистике очень обширна, так как они могут выполнять морфологический (определение частеречной принадлежности слова и его грамматических категорий), синтаксический (определение структуры предложения) и семантический анализ (выявление семантических связей между словами и предложениями), распознавать и генерировать текст и т.д.

В связи с тем, что в настоящее время активно расширяются области применения нейросетевых технологий, появляется значительное количество работ, посвященных возможностям использования искусственного интеллекта в лингвистике. В целом, эти работы можно разделить на две большие категории: исследования о потенциальных возможностях и перспективах использования нейросетей в лингвистических исследованиях [2, 14, 17] и исследования прикладного характера, изучающие перевод разных типов и стилей текстов с помощью инструментов искусственного интеллекта [5, 8, 10, 11, 12].

Большинство представленных на рынке моделей ориентированы на работу с текстом и картинками (GPT, Grok, BERT, T5, XLNet, BigGAN, Midjourney и т.д.). Существует гораздо меньше решений, способных быстро обрабатывать видеоролик и предоставлять пользователю качественный перевод дубляж или полудубляж/войсовер, а не субтитры [3].

Аудиовизуальный перевод является особой сферой переводческой деятельности. При переводе полимодального текста, которым является аудиовизуальное произведение, необходимо учитывать холистические факторы (визуальные, референтные). Перевод для закадрового звучания, при котором исходная звуковая дорожка приглушается, является самым простым видом [7, с. 376]. Его основными особенностями являются меньшая, по сравнению с письменным переводом, точность в передаче отдельных значений, синхронизация длины текста перевода с исходной длительностью речи, отставание звуковой дорожки перевода от оригинальной для создания эффекта аутентичности исходного текста [16, р. 27], «синхронизация вербальной информации с компонентами оригинального видеоряда» [6, с. 42]. В последнем случае имеется в виду кинесическая синхронность, когда текст перевода совпадает со средствами невербальной коммуникации. Важно отметить, что в разных видах АВП степень точности кинесической синхронности может варьироваться – от максимальной в дубляже до условной при субтитрировании [4, с.11].

В 2021 году пользователям Яндекс Браузера стал доступен прототип переводчика видеоконтента, который работал с ограниченным числом роликов на платформе YouTube. Чуть позже нейросетевой закадровый перевод стал доступен для всех YouTube-каналов, Vimeo, VK и на платформе Coursera. В настоящее время (на момент

написания статьи) перевод осуществляется с восьми языков: английского, французского, испанского, немецкого, итальянского, китайского, японского и корейского и доступен только в Яндекс Браузере или приложении Яндекс для пользователей, использующих операционную систему Windows или Android/iOS. Перевод односторонний, то есть осуществляется с иностранного языка на русский (с русского есть перевод только на казахский). Процесс перевода занимает некоторое время.

По утверждению компании, модель переводчика не только распознает речь с учетом тайминга и переводит ее, но и учитывает информацию о гендерной принадлежности говорящего, синхронизирует два потока речи за счет сокращения пауз между словами и фразами [3].

Для анализа было выбрано выступление австралийского исследователя искусственного интеллекта (ИИ) и нейробиологии Скотта Болланда на TEDx – это конференция в формате TED (Technology, Entertainment, Design), где спикеры – специалисты в своей предметной области, рассказывают об актуальных тенденциях и перспективах развития науки. Выступление рассчитано на максимально разнообразную аудиторию, поэтому главная задача спикера – понятно донести идею до реципиента. Выступления относятся к научно-популярному стилю, где наряду с большим количеством научных и общеупотребительных терминов используется эмоционально-окрашенная лексика, неформальные и устойчивые выражения, игра слов, тропы – что является факторами, осложняющими перевод.

В ходе анализа рассмотрены вербальные и паравербальные средства коммуникации. Первые составляют особый интерес при переводе, так как ведут к определенным трудностям, связанными со структурными и культурными различиями языков. Учитывается и паравербальный аспект, влияющий на качество восприятия информации.

В данной работе внимание уделено ошибкам, полученным в ходе перевода. Они выявлены в результате сопоставления оригинала и нейросетевого перевода на трех уровнях языка: грамматическом, синтаксическом и лексико-семантическом.

На грамматическом уровне допущены следующие ошибки:

1. Замена глагола прошедшего времени глаголом настоящего времени. В предложении «...I loved going to class and I did well» [15, 7:24] обе части сложного предложения относятся к прошедшему времени, в то время как в переводе это предложение выглядит так: «...мне нравится ходить на занятия, и я хорошо учусь» [там же, 7:24]. Такой перевод не является уместным, так как меняет

временную принадлежность и ведет к искажению смысла.

2. Замена повелительного наклонения глагола на изъявительное. Происходит смещение с одного субъекта, выполняющего действие, на другой. Например, фраза «...pretend the correct terminology is "expert"» [там же, 7:05] переведена нейросетью как «...я буду называть себя экспертом» [там же, 7:05]. В данном предложении изменилось наклонение глагола и лицо подлежащего.

При рассмотрении предложения в контексте нет серьезного искажения смысла, но тон повествования меняется и лишается изначальной экспрессивности. В оригинальном предложении слушатель становится оппонентом спикера, благодаря использованию повелительного наклонения. Замена этой категории ведет к потере выстроенной взаимосвязи между говорящим и слушателем.

3. Замена условного наклонения глагола на изъявительное наклонение. Например, в предложении «...when you put your hand in front of your face, you'd see this white blob» [там же, 4:12] условное наклонение предназначено для указания гипотетически возможной ситуации. Однако перевод, выполненный следующим образом: «...когда вы подносили руку к лицу, вы видели белое пятно» [там же, 4:12], предполагает, что говорящий и слушатели имеют совместный опыт, а действие совершалось кем-то из аудитории, что приводит к нарушению эквивалентности перевода.
4. Согласование имени числительного с именем существительным. В анализируемом видеофрагменте спикер несколько раз прибегает к использованию статистических данных. Однако в двух случаях их перевод имеет неверную грамматическую форму – имя числительное не согласуется с именем существительным, что затрудняет понимание услышанного. Так, например, выражение «one of the 63%» [там же, 6:34] переведено на русский язык как «одним из тех шестьдесят три процента» [там же, 6:34], а выражение «fill a classroom with 30 students» [там же, 5:46] как «заполним класс тридцать учениками» [там же, 5:46]. Можно заметить, что при явном нарушении обязательных норм языка перевода сбивается и ритм речи, и восприятие информации.

На синтаксическом уровне допущены следующие ошибки:

1. Буквальный перевод местоимения «it». В одном из фрагментов видеоматериала, состоящем из 21 предложения, 11 раз использовано слово «это», чего можно было бы избежать путём объединения нескольких предложений в одно, замены тема-рематического членения, а также перечислений однородных членов предложения. Это обусловлено

тавтологией в оригинале видео, однако перевод можно сделать более лаконичным. Например, предложение «как только они это поняли, им это наскучило, и они перестали вести себя так» [там же, 4:05] можно преобразовать следующим образом: «они перестали себя так вести, как только поняли, что оно им наскучило».

2. Опущение риторического вопроса в переводе. Рассказывая об эксперименте, спикер задал риторический вопрос «And what've happened?» [там же, 3:52], который никак не был отражён в переводе, выполненном нейросетями. Данное предложение не несёт особой смысловой нагрузки, однако является немаловажной частью публичного выступления, отражающей взаимодействие спикера с аудиторией. А потому должно быть сохранено и в переводе.

На лексико-семантическом уровне можно выделить следующие единицы, в которых в процессе перевода допущены ошибки:

1. Стилистически-окрашенная лексика. В оригинале найдено 35 примеров употребления стилистически-окрашенной лексики. Среди них 12 не являются адекватными в переводе на русский язык. В процессе перевода данного видео нейросетью многие стилистически-окрашенные слова приобретали нейтральное значение. Это может быть связано с особенностью восприятия экспрессивной лексики в русской культуре. Однако иногда подобная замена ведёт к «сухому» и однообразному переводу. Например, вопрос «Why is it that so many kids struggle at school?» [там же, 2:16] в русском переводе звучит так: «Почему так много детей испытывают трудности в школе?» [там же, 2:16]. Глагол «struggle» при переводе потерял негативную коннотацию и был заменён нейтральным словосочетанием. В качестве альтернативного варианта может быть использовано слово «мучаются».

В переводе также встречается 3 случая, когда стилистически нейтральное на английском языке слово в переводе не приобретает окраску. Например, предложение «none of us, if given the choice, would stare blankly at a wall» [там же, 5:24] переведено как «Никто из нас, если бы у него был выбор, не стал бы тупо смотреть в стену» [там же, 5:24]. Слово «тупо» имеет негативную коннотацию, отсутствующую в оригинале. Вместо него, можно использовать слово «бестолково».

2. Идиомы, устойчивые выражения и фразовые глаголы. В оригинале обнаружено 16 случаев употребления подобной лексики, 9 из них переведены на русский язык стилистически нейтральным словом, не являющимся устойчивым выражением. Например, английская идиома «I was just <...> bored out

of my mind» [там же, 6:34] переведена как «я был одним из тех <...>, которым было скучно» [там же, 6:34]. Эмоционально-окрашенное устойчивое выражение заменено на стилистически нейтральное «скучно», не передающее в равной степени вызванные у спикера эмоции. В качестве эквивалентного перевода можно использовать фразу «Я с ума сходил от скуки».

3. Терминология. Ошибки в переводе терминологии крайне редки, что обусловлено их однозначностью в конкретном контексте. Однако в видео найдено 2 примера некорректного перевода терминов. В одном из них «they release endorphins with an inverted U-shape with respect to familiarity» [там же, 2:28] переведено с опущением термина, означающего перевернутую U-образную кривую: «уровень эндорфина повышается по мере того, как мы узнаем что-то новое» [там же, 2:28]. Во втором случае термин «exponential rate» [там же, 14:17] переведён наречием «постепенно» [там же, 14:17]. Употребленное в русском языке слово является примером генерализации, что нельзя назвать адекватным при переводе точных математических терминов.
4. Стилистические фигуры. Несмотря на наличие в видео научных терминов, спикер использует стилистические фигуры, чтобы сделать выступление более интересным и запоминающимся. Однако из 16 примеров использования фигур речи в 5 допущены ошибки при переводе. Например, метафора «learning journey» [там же, 6:30], символизирующая долгий и сложный путь, пройденный спикером в процессе обучения, передана в переводе словосочетанием «опыт обучения» [там же, 6:30]. Данная конструкция не обладает стилистической окраской и не передаёт заложенное значение, поэтому нейтрализацию здесь нельзя считать адекватной трансформацией. Ее можно заменить на метафору «познавательное путешествие».
5. Неформальная лексика. В анализируемом видеофрагменте разговорная лексика используется крайне редко, однако отмечено 4 случая её употребления. Перевод двух из них сложно назвать эквивалентным, так как в одном случае слово «uni» [там же, 0:58] передано в переводе через слово «университет» [там же, 0:58], что делает речь менее живой и неформальной. В качестве альтернативного варианта может быть использовано сокращение «универ». Во втором случае слово «geek» [там же, 7:03] переведено на русский язык с помощью транскрипции, заимствованной в русский язык и относящейся к жаргонной лексике. Оно редко используется носителями и, может быть, не понято некоторыми слушателями. Вполне адекватной заменой является слово «ботан», также относящееся к нефор-

мальной лексике, но известное более широкому кругу людей.

Важным аспектом аудиовизуального перевода являются паравербальные средства коммуникации, сохранение которых обеспечивает динамическую эквивалентность. В связи с этим следует учитывать три типа синхронности: кинесическую, артикуляционную и изохронию.

1. Кинесическая синхронность требует соответствия между вербальной и невербальной коммуникацией, то есть синхронизации слов с жестами говорящего. В анализируемом видео не наблюдается явных и многократных нарушений, так как жесты спикера, по большей части, не имеют конкретного значения. Однако, из-за длинных конструкций на русском языке, иногда происходит сдвиг перевода от оригинального момента речи, в связи с чем, жест как бы «запаздывает» на несколько слов. Так, например, в оригинале видео спикер, говоря слово «brain-damaged», иллюстрирует руками кавычки. Однако при переводе данный жест выпадает на слово «ученика», и восприятие русского слушателя сбивается, вызывая недопонимание. Помимо этого, при произнесении некоторых риторических вопросов, спикер делает акцент на вопросительном слове, активно используя мимику, которая в переведенном фрагменте отстаёт и выпадает почти на конец фразы.
2. Артикуляционная синхронность обеспечивает совпадение речи с артикуляционным движением. Такой вид синхронности характерен для дублированного перевода фильмов, однако крайне редко может быть замечен в нейросетевом переводе. В нашем случае артикуляция спикера и озвученный перевод совпадают только при окончании фраз или предложений, и то не всегда. Из-за характерного для перевода с английского языка на русский увеличения длины предложения, перевод либо произносится быстрее, либо начинается раньше, чем в оригинале.
3. Изохрония представляет собой соблюдение длины реплик и паузации оригинального видеофрагмента в его конечном переводе. Поскольку длина предложения в русском и английском языках отличается, соблюдения изохронии можно добиться с помощью переводческих трансформаций на разных уровнях. Однако такое преобразование не всегда является эквивалентным. Например, в оригинале видео в предложении «...I was slow to walk, slow to talk and slow to learn» [там же, 0:41] использование риторического повтора позволяет усилить эмоциональное воздействие. Но данная стилистическая фигура опускается при переводе, чтобы сократить предложение и попасть в тайминг: «...я поздно начал ходить, говорить и учиться» [там же, 0:41].

Нейросетью выбрана и иная тактика, позволяющая завершить озвучивание перевода синхронно с оригинальным голосом. Для этого голос нейросети начинает фразу раньше, чем к ней приступает спикер, создавая возможность завершить ее в нужный момент. Однако эта стратегия приводит к непрерывному потоку речи, без пауз и замедлений, что затрудняет восприятие информации слушателем. Необходимые паузы сохраняются лишь между короткими риторическими вопросами или простыми предложениями, которые в переводе на русский язык практически полностью сохраняют свою длину.

Заключение

Таким образом, сравнительно-сопоставительный анализ показал:

- на вербальном уровне выявлена 41 ошибка при переводе: 5 на грамматическом уровне, 2 на синтаксическом уровне и 34 на лексико-семантическом уровне;
- большинство ошибок на грамматическом уровне связано с глагольными категориями, что приводит к искажению времени повествования или смене тона повествования;
- ошибки на синтаксическом уровне допущены, главным образом, из-за темпа и качества речи говорящего в оригинальном видео. Они вызваны от-

сутствием синтаксических преобразований;

- ошибки при переводе на лексическо-семантическом уровне связаны, в основном, со стремлением сделать текст на русском языке нейтральным, формальным и тактичным. Однако такая стратегия, во многом, ведёт к потере искренности и выразительности спикера.
- на паравербальном уровне кинесическая синхронность, как правило, соблюдается с небольшими отклонениями при смещении слова внутри предложения или быстром воспроизведении речи. Артикуляционная синхронность практически отсутствует, что связано с разницей в фонетическом строе исходного и переводящего языков. Изохрония нарушена, что обусловлено различной длиной предложений в русском и английском языках, а также недостаточным количеством проведения трансформаций, способствующих «сжатию» текста.

Технология Яндекс Браузера достойно справляется с поставленной задачей и способна подбирать подходящие для культуры переводящего языка единицы, однако, не может учитывать некоторые стилистические особенности оригинальной речи и грамматических свойств языка перевода, а также преобразовывать текст таким образом, чтобы он чётко следовал таймингу оригинала видео.

ЛИТЕРАТУРА

1. Бархударов Л.С. Язык и перевод. — М.: Международные отношения, 1975. — 239 с.
2. Бурнашев Р.Ф., Анварова Л.А. Применение нейронных сетей в автоматическом переводе и обработке естественного языка // *Universum: технические науки*. — 2024. — № 4 (121). — С. 39–43.
3. Гаскаров Т. Как Яндекс помогает преодолеть языковой барьер: нейросетевой перевод видео, картинок и текста (Блог компании Яндекс) // Хабр. — URL: <https://habr.com/ru/companies/yandex/articles/576438/> (дата обращения: 05.02.2025).
4. Демидова А.С., Мамонова А.Ю. О кинесической синхронности при переводе под закадровое озвучивание (на примере французской научно-популярной передачи *C'est pas sorcier*) // *Didactica translatologica*. — 2024. — № 3. — С. 10–13.
5. Иванченко Т.А. Ошибки в машинном переводе с немецкого языка на русский (на материале статей немецкоязычных СМИ и текстов их переводов) // *Ученые записки Санкт-Петербургского университета технологий управления и экономики*. — 2021. — Т. 22. — № 4. — С. 30–41.
6. Калинин А.Ю. Прагматически детерминированные приемы аудиовизуального медиаперевода // *Вестник Московского университета. Серия 19. Лингвистика и межкультурная коммуникация*. — 2020. — № 2. — С. 41–50.
7. Козуляев А.В. Аудиовизуальный полисемантический перевод как особая форма переводческой деятельности и особенности обучения данному виду перевода // *Царскосельские чтения*. — 2013. — № XVII. — Т. I. — С. 374–381.
8. Козьякова М.В., Корнеева И.В. Творческий потенциал ChatGPT в современной практике перевода корейских поэтических текстов на примере жанра сиджо // *Язык как структура и дискурсивная практика*. — 2023. — № 2. — С. 18–25.
9. Комиссаров В.Н. Современное переводоведение. Учебное пособие. — М.: ЭТС, 2002. — 424 с.
10. Митин П.С. Типичные ошибки машинного перевода общеполитической лексики (на примере перевода прессы Арабской республики Египет) // *Вестник Московского университета. Серия 22. Теория перевода*. — 2021. — С. 17–33.
11. Норец М.В., Рейнова А.В. Использование искусственного интеллекта при переводе художественного текста (на примере романа Дж. Р.Р. Мартина «Пламя и кровь») // *Ученые записки Крымского федерального университета имени В.И. Вернадского. Филологические науки*. — 2024. — № 10 (76). — № 3. — С. 149–158.
12. Смирнова В.Н. Технология DeepL Translate в переводе пословиц и поговорок (на материале английского языка) // *Вестник Челябинского государственного университета*. — 2023. — № 9 (479). — С. 26–131.
13. Созыкин А.В. Обзор методов обучения глубоких нейронных сетей // *Вестник Южно-Уральского государственного университета. Серия: Вычислительная математика и информатика*. — 2017. — № 3. — Т.6. — С. 28–59.

14. Alaqlobi O., Alduais A., Qasem F., & Alasmari M. Artificial intelligence in applied (linguistics): a content analysis and future prospects // Cogent Arts & Humanities. – 2024. – Vol. 11. – No 1. – URL: <https://www.tandfonline.com/doi/epdf/10.1080/23311983.2024.2382422?needAccess=true> (дата обращения: 14.02.2025).
 15. Bolland S. Neuroscience, AI, and the Future of Education [видеозапись выступления] // YouTube. – 30.06.2016 – URL: https://www.youtube.com/watch?v=_cYlvfS-knA (дата обращения: 09.02.2025).
 16. Franco E., Matamala A., Orero P. Voice-over translation: an overview. – Bern, Peter Lang, 2010. – 248 p.
 17. Vajjala S. Generative artificial intelligence and applied linguistics // JALT Journal. – 2024. – Vol. 46. – No 1. – P. 55-76.
-

© Данчук Ольга Васильевна (odanchuk@yandex.ru), Сереброва Анастасия Вячеславовна (anastasiya.20102004@gmail.com).

Журнал «Современная наука: актуальные проблемы теории и практики»