

# Обзор технологий синтеза речи

## Тема №18

Синтез речи (Text-to-Speech, TTS) — это процесс преобразования текста в звучащую речь. В XXI веке синтез речи стал одним из ключевых направлений искусственного интеллекта (ИИ), так как он обеспечивает удобство взаимодействия человека и компьютера, делает цифровую среду более доступной, а также открывает новые возможности в образовании и науке.

Сфера применения:

- голосовые ассистенты (Siri, Алиса, Alexa, Google Assistant);
- системы навигации;
- доступность для людей с нарушениями зрения;
- автоматизация озвучивания лекций, аудиокниг, мультимедиа;
- образовательные сервисы, тренажёры и игровые приложения.

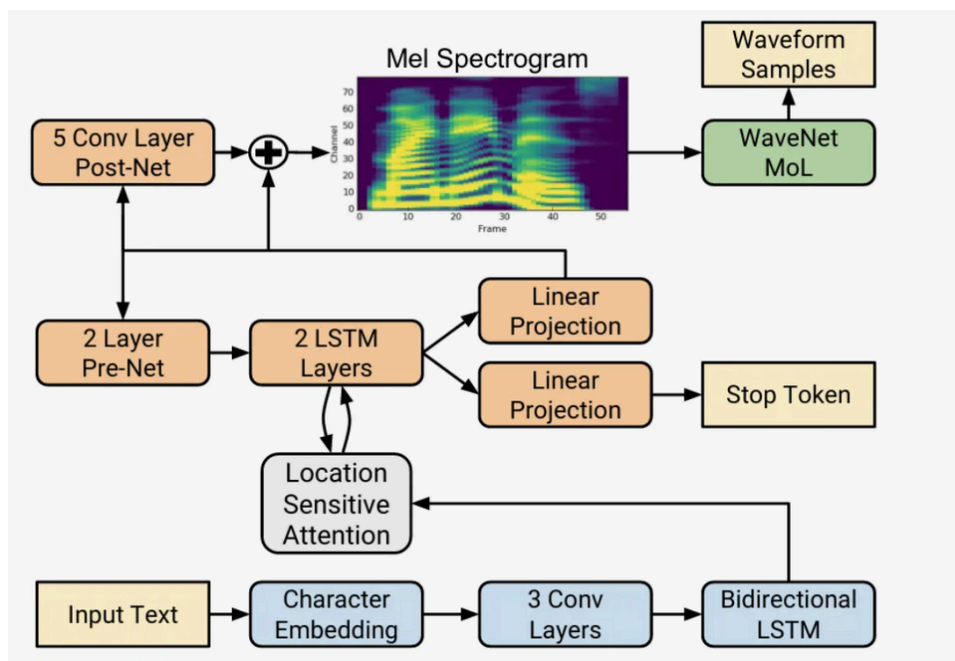
## История развития технологии

Годы	События
1770–1791 гг.	Вольфганг фон Кемпелен создал «говорящую машину», которая воспроизводила отдельные гласные и согласные.
XIX в	Чарльз Уитстоун и Герман фон Гельмгольц продолжили исследования в области акустики и механического синтеза речи
1939 г.	Bell Labs представили <b>VODER</b> на Всемирной выставке в Нью-Йорке. Это был первый электронный синтезатор речи.
1970–1990-е	Методика «склеивания» заранее записанных фрагментов речи. Недостатки: ограниченный словарь, невозможность выразительной интонации.
с 2010-х гг.	Появление глубокого обучения и нейросетевых моделей:

## Архитектура современных TTS-систем

Современные системы синтеза речи обычно состоят из двух компонентов:

- Модуль преобразования текста в спектрограмму (Text-to-Mel)**
  - примеры: Tacotron, FastSpeech, Glow-TTS;
  - определяет произношение, длительность звуков, интонацию.
- Вокодер (Mel-to-Wave)**
  - преобразует спектрограмму в звуковую волну;
  - примеры: WaveNet, HiFi-GAN, Parallel WaveGAN.



## Современные сервисы синтеза речи

### Google Cloud Text-to-Speech

- более 220 голосов, поддержка SSML;
- WaveNet для естественного звучания.

### Yandex SpeechKit

- поддержка русского языка, разные стили озвучки;
- используется в «Алисе».

### Microsoft Azure Speech Service

- Custom Neural Voice (создание уникальных голосов);
- гибкие настройки SSML.

### OpenAI TTS (2024+)


- интеграция с GPT-моделями;
- мультимодальные возможности.

## Применение в образовании


- Автоматизация озвучивания лекций и конспектов.
- Создание адаптивных учебников (озвучка + текст).
- Поддержка студентов с ОВЗ.
- Создание учебных подкастов и аудиолекций.
- Интерактивные задания с голосовым сопровождением.

## Бесплатные онлайн-сервисы для синтеза речи


## 1. TTSMP3.com

-  <https://ttsmp3.com>
- Поддержка русского языка, выбор мужских и женских голосов.
- Можно изменять скорость речи, добавлять паузы через **SSML**.
- Ограничение: до 3000 символов в день бесплатно.
- Удобно для подготовки небольших озвучек к уроку.


## 2. Natural Readers (онлайн-версия)

-  <https://www.naturalreaders.com/online/>
- Бесплатный режим доступен сразу, без регистрации.
- Поддержка русского языка.
- Возможность выбора разных голосов (ограниченно в бесплатной версии).
- Удобен для подготовки озвучек учебных текстов.

## 3. Readloud.net

-  <https://readloud.net>
- Бесплатный сервис без регистрации.
- Поддерживает русский язык.
- Генерирует MP3-файл с синтезированной речью.
- Минус: качество голосов ниже, чем у современных нейросетевых TTS.

## 4. ttsfree.com

-  <https://ttsfree.com>
- Бесплатный TTS с возможностью выбора стиля речи.
- Работает быстро, поддерживает русский язык.
- Генерирует MP3 для скачивания.

## Работа с сервисом TTSMP3.com

**TTSMP3.com** — простой онлайн-сервис: вставляете текст → выбираете голос → слушаете и скачиваете **MP3**. Поддерживает русский язык и кириллицу.



## Free Text-To-Speech and Text-to-MP3 for US English

Easily convert your **US English** text into professional speech for free. Perfect for e-learning, presentations, YouTube videos and increasing the accessibility of your website. Our voices pronounce your texts in their own language using a specific accent. Plus, these texts can be downloaded as MP3. In some languages, multiple speakers are available. [Text conversion service](#) [Natural sounding voices](#)

Regular Voices AI Voices

your text here

× [Language-Inspired Home Decor](#)

Бесплатный лимит: **≈ 375 слов или 3 000 символов в день** (для всех языков суммарно).

Движок голосов — **Amazon Polly**; сервис ссылается на теги **SSML** от Amazon (вставка пауз, изменение темпа/тона и т. д.). На сайте указано, что **созданные MP3 можно свободно использовать** (в т. ч. в YouTube/электронном обучении), независимо от того, платная или бесплатная версия использовалась.

Сайт предоставляет 2 основных режима работы “Regular Voices” и “AI Voices”. “Regular Voices” представляет стандартную озвучку похожую на переводчик, особенность такого подхода заключается в том, что отсутствует корректная постановка ударений и чтение сложных слов

Режим “AI Voices” применяет технологии искусственного интеллекта для генерации голосового сообщения. В данном подходе текст генерируется более точно, с грамотными ударениями и уверенным прочтением сложных слов и оборотов речи.

В режиме “Regular Voices” существует выбор из 2-х вариантов Российских голосов: “Russian/Tatyana” и “Russian Maxim”. Вы можете выбрать любой голос и ввести текстовое сообщение для озвучки.

Italian / Carla

Italian / Bianca

Japanese / Takumi

Japanese / Mizuki

Korean / Seoyeon

Mexican Spanish / Mia

Norwegian / Liv

Polish / Jan

Polish / Maja

Polish / Ewa

Polish / Jacek

Portuguese / Cristiano

Portuguese / Inês

Romanian / Carmen

Russian / Tatyana

Russian / Maxim

Swedish / Astrid

Turkish / Filiz

US English / Kimberly

US English / Ivy

US English / Kimberly

Read

Download as MP3

Далее необходимо нажать кнопку “Read” для прочтения голосового сообщения или кнопку “Download as MP3” для загрузки сообщения в формате MP3.

Regular Voices

AI Voices

Пример классической озвучки текста

US English / Kimberly

Read

Download as MP3

Следующий режим “AI Voices” представляет аналогичное поле для ввода текста, необходимого для озвучки, возможность прочитать и скачать сообщение, а также выбор голоса. Однако, помимо этого добавлены настройки скорости чтения сообщения (“Speed”).

Regular Voices

AI Voices

your text here

Speed:  1.00

Add Instructions: ☐

Alloy (female)

Read

Download as MP3

Today's use: 22 / 1,000 characters

Также, сервис предоставляет возможность добавить определенное количество описательных инструкций для вашего текста с помощью кнопки “Add Instructions”, например: “Прочитай текст с драматическим оттенком” и т.д.

Add Instructions:



Enter your voice instructions here (e.g., 'Speak in a dramatic tone', 'Use a thick australian accent', etc.) - Recommended voices for this feature: Ash, Coral, Sage - results may vary!

Clear

Random ✨

Help

Поэкспериментируйте в работе с данным сервисом! Не забывайте, что бесплатное использование предоставляет только 1000 символов озвучки в день!

## Практические задания

### Задание №1. Работа с онлайн-сервисами синтеза речи

1. Попробуйте три разных сервиса из списка: **TTSMP3.com**, **Natural Readers**, **Readloud.net**.
2. Введите один и тот же текст (около 3–4 предложений).
3. Сравните полученные результаты по параметрам:
  - естественность звучания,
  - постановка ударений,
  - скорость,
  - возможность управлять голосом.
4. Составьте сравнительную таблицу (сервис → плюсы → минусы → оценка от 1 до 5).

### Задание №2. Управление качеством речи через SSML и инструкции

1. Возьмите сервис **TTSMP3.com** (режим Regular и AI Voices).
2. Сгенерируйте 3 версии одного текста:
  - обычная (без тегов),
  - с паузами и изменением скорости,
  - с акцентами и изменением интонации.
3. Опишите, как именно теги и инструкции изменили восприятие текста.

### Задание №3. Озвучивание учебного материала

1. Возьмите небольшой учебный текст по информатике (например, объяснение рекурсии или алгоритма сортировки).
2. Озвучьте его через один из TTS-сервисов (по выбору).
3. Подготовьте **мини-ролик (1–2 минуты)**, где слайды сопровождаются синтезированной озвучкой.



## **Задание № 4 (проект). Создание образовательного квеста с использованием синтеза речи**

**Задача:** разработать мини-квест для урока информатики с применением технологий TTS.

Этапы работы:

1. **Сюжет:** придумайте короткий сценарий (например, «Путешествие в мир алгоритмов», «Кибер-детективное агентство», «Поиск утёкших данных»).
2. **Персонажи:** создайте 2–3 персонажей (учёный, робот-ассистент, хакер и т. д.).
3. **Роли TTS:** озвучьте каждого персонажа разными голосами через онлайн-сервисы (например, один — через TTSMP3, другой — через Natural Readers).
4. **Интерактив:** подготовьте задания для учащихся (разгадать зашифрованное слово, исправить алгоритм, найти ошибку в коде).
5. **Презентация проекта:** соберите слайды с заданиями + вставьте озвучку персонажей.
6. **Реализация на уроке:** презентуйте мини-квест однокурсникам как «пробный урок».