

ReadSpeaker® speechMaker Studio

ReadSpeaker speechMaker Studio (speechMaker Studio) is a browser-based text-to-speech (TTS) application for converting written text into downloadable audio files.

Product Description

ReadSpeaker speechMaker Studio (speechMaker Studio) is a browser-based Software as a Service (SaaS) application that enables multiple licensed users within a customer account to create, manage, and convert user-entered text into downloadable audio files using ReadSpeaker text-to-speech (TTS) technology.

Access to speechMaker Studio requires authentication using a username and password. A customer account includes a defined number of user licenses as specified in the applicable Order Confirmation. Users with the appropriate role (Owner or Administrator) may invite additional users within the limits of the licensed account.

To generate an audio file, the user enters or pastes text into the application, selects the language and voice, chooses the audio output format, and adjusts available voice parameters such as speaking rate, pitch, and volume. The user may preview the generated audio within the application. Previewing audio does not consume credits. Credits are deducted only when an audio file is exported.

Audio generation is performed using ReadSpeaker neural TTS technology. Exported audio files are downloaded to the user's local device.

Features

Dashboard

After authentication, the user is redirected to the dashboard. The dashboard provides an overview of:

- Recent activity within the customer account
- System notifications
- Account-related information

The information displayed may vary depending on the user's assigned role.

Content Library

Text content can be stored in speechMaker Studio's content library.

A file containing text content is referred to as a script. A script can contain one or more text blocks.

Scripts can be organized in folders. Folders and scripts can be organized into projects.

All users within a customer account can access all content stored in the content library. Users assigned the Viewer role have read-only access and cannot create, edit, or delete content.

Text to Speech

speechMaker Studio converts user-provided text into speech using ReadSpeaker's Neural Premium TTS technology.

The languages and voices available within the application depend on the voice package enabled for the customer account. The list of enabled languages and voices is displayed in the [voice gallery](#) within the application.

Content Studio

All content is stored within the customer account and is accessible according to assigned user roles.

Selecting a script from the content library or dashboard opens it in the content studio.

Each script can contain one or more text blocks.

Text blocks can have individual settings for:

- [Reading voice](#)
- [Speaking rate](#) (speed)
- [Pitch](#)
- [Volume](#)

The user can adjust the pause duration between text blocks and the duration of optional initial silence at the beginning of the audio, to align voice prompts with predefined time windows, for example.

The following tools are available for detailed control of speech output:

- [Pronunciation editor](#)
- [Intonation editor](#)
- [Pause tool](#)

Adjustable Reading Voice

A reading voice is selected from the list of voices enabled for the customer account. Different voices may be assigned to individual text blocks within the same script. A single script may therefore contain multiple speakers.

Adjustable Speaking Rate

Speaking rate can be adjusted via a sliding scale across the selected voice's supported range, from slower to faster relative to the default rate.

Adjustable Pitch

Pitch can be adjusted via a sliding scale across the selected voice's supported range, from lower to higher relative to the default pitch. Pitch can be adjusted independently for each text block.

Adjustable Volume

Volume can be adjusted via a sliding scale across the selected voice's supported range, from lower to higher relative to the default volume. Volume can be adjusted independently for each text block.

Preview: Listen to Audio Before Exporting

Generated audio can be previewed within the application before export. Preview is available for:

- The entire script
- Individual text blocks
- Pronunciation edits
- Intonation edits

Previewing audio does not consume credits. Credits are deducted only when audio is exported.

Pronunciation Editor

The pronunciation editor enables modification of word's pronunciation within a script. Pronunciation may be modified using one of the following methods:

- Alternate spelling - A word's pronunciation can be overridden by using a different spelling when the default TTS pronunciation does not match the intended pronunciation.

- **Phonetic transcription** - A word's pronunciation may be defined using phonetic symbols selected from a digital IPA keyboard.

Such pronunciation edits apply only to the specific occurrence of the word within the script.

Pronunciation edits can be previewed before application.

Intonation Editor

Intonation can be modified at the sentence level by adjusting the relative pitch and/or duration of individual words within the sentence.

The intonation editor displays the selected sentence as a curve, where each word is represented by a point that can be repositioned to modify the pitch. A duration value can be assigned to each word to modify its length.

Pitch and duration adjustments are relative to the default rendering of the selected voice. Identical adjustments may produce different results depending on the characteristics of the selected voice.

Intonation edits can be previewed before application.

Pause Tool

Pauses can be inserted anywhere in the text. Pause duration is selected from predefined values within the application. Inserted pauses are visually indicated in the editor.

Export Audio

Audio files are downloaded to the user's device upon export.

Audio can be exported at different levels:

- **Script** - Exports a single script from within the content studio or the content library.
- **Project** - Exports a project from the content library.

Exporting a project generates a ZIP archive that mirrors the project's folder structure in the content library. Only folders containing scripts are included in the exported archive.

Credits are deducted upon export.

Voice Gallery

The voice gallery displays the languages and voices enabled for the customer account. Voices can be filtered using the available search and filter functions. The voice list is sorted by language and voice name.

Each voice entry includes an avatar, a flag indicating the language, and sample text that can be played to preview the voice.

Voices can be marked as favorites. Voices marked as favorites are displayed at the top of the voice selection list in the content studio and in the voice gallery.

Additional languages and voices can be added to the customer account's voice gallery.

Credits for Production

speechMaker Studio uses a credit-based model for audio export.

Credits are deducted only when audio is exported. Previewing audio within the application does not consume credits.

One credit is charged per character when an audio file is exported. Characters are counted excluding whitespace. Characters introduced or modified through pronunciation edits are included in the character count.

Customer accounts typically include an initial credit allocation for testing purposes, as defined in the applicable Order Confirmation.

Additional credit packages may be purchased by the account Administrators. Credit quantities and validity periods are defined in the applicable Order

Confirmation. Credit purchases are subject to confirmation by ReadSpeaker before credits are applied to the customer account.

View Remaining Balance

The number of remaining credits, representing the amount of text (measured in characters excluding whitespaces) that may still be converted to speech, is displayed in the top bar of the speechMaker Studio interface.

The remaining balance is visible to users of all roles.

Additional credits may be purchased at any time by the account Administrators.

When the credit balance reaches zero, audio export is not available until additional credits are applied to the customer accounts.

Users and Roles

A customer account includes five user licenses. Additional user licenses may be added.

Users within a speechMaker Studio account can have the following roles:

- Owner - Assigned to the first user created for the customer account. The customer account must have an Owner. The Owner has the same permissions as an Administrator.
- Administrator - Can create, edit, and delete all content in the library, invite users, assign user roles, and purchase speechMaker Studio credits.
- Editor - Can create, edit, and delete all content in the library.
- Viewer - Has read-only access to content. Viewers cannot create, edit, or delete content. Viewers can preview audio in the content studio.

Technical Specifications

Device Independent

speechMaker Studio is a browser-based application that is supported on current versions of major web browsers.

Fully Server-Based, No Software Downloads or Installation

speechMaker Studio is fully server-based. No software installation or download is required for use. The application is hosted by ReadSpeaker and accessed through a supported web browser.

Product Hosting

speechMaker Studio is hosted by ReadSpeaker in a high-security environment with a high level of redundancy and failover, and with the following properties:

- Double-redundant enterprise load balancing and failover
- 100% Gigabit network (internally and to the Internet)
- IBM server hardware with virtualization engine
- Redundant power supply
- Highly scalable and high availability design
- Environmentally friendly
- Storage Area Network (SAN) storage over Gigabit iSCSI
- UPS and diesel engine power backup
- 24/7 video and audio surveillance
- Level 3 security
- Argonite-based automatic fire-fighting system
- Automatic server monitoring every 3 seconds with alarm system
- Automatic daily backup

Supported Input Formats

speechMaker Studio accepts plain text input. Text must be entered directly into the Editor, either by typing or by copying and pasting from an external source.

Text Size Limit

speechMaker Studio has a default limit of 20,000 characters, excluding spaces, for text-to-speech conversion per exported audio file produced.

Supported Audio Formats

Exported audio files are available in the following formats:

- MP3 (24, 48, and 128 Kbits/s, 22.1 KHz and 44.1 KHz; 192 Kbit/s, 44.1 KHz)
- Ogg Vorbis (variable bitrate Q=0.2)
- PCM (8 bit/16 KHz, 16 bit/8 KHz, 16 bit/16 KHz, 16 bit/44.1 KHz)
- IMA-ADPCM (4 bit/8 KHz)
- WAV (U-law 16 bit/8 KHz, A-law 16 bit/8 KHz)
- WAV (8 bit/16 KHz, 16 bit/8 KHz, 16 bit/16 KHz, 16 bit/44.1 KHz, 16 bit/48 KHz)

The formats available for a customer account are defined in the applicable Order Confirmation.

Voice Types

speechMaker Studio currently supports ReadSpeaker's Neural Premium voice packages.

Product Components

SaaS Application

The core of speechMaker Studio is the ReadSpeaker speechMaker Studio SaaS (Software as a Service) application, hosted by ReadSpeaker. The application is accessible to licensed users through authenticated login.

Login Credentials

Access to speechMaker Studio is provided through user-specific login credentials associated with a licensed customer account. No additional software components are required for end users.

User Support and Documentation

Technical support is provided by ReadSpeaker. Contact details and support hours are available at: <https://www.readspeaker.com/support/>.

Options

Optional features can be added to speechMaker Studio. Pricing information is available from your Account Manager.

Larger Maximum Size for Text Input

The default limit of 20,000 characters per audio file produced can be increased on a per-customer basis.

Requirements and Limitations

Internet Connection

An active Internet connection is required to use speechMaker Studio.

Default Size for Text Input

The default maximum size for text-to-speech conversion per audio file produced is 20,000 characters excluding spaces. This can be increased on a per-customer basis, see under Options above.

Export Audio

The customer may use the exported audio files only within the context set out in the applicable order confirmation.

Certain voices may be subject to usage restrictions. Applicable restrictions are displayed in the export interface when affected voices are included in the export. The customer is responsible for ensuring that exported audio is used in accordance with the applicable usage restrictions.

Voice Types

speechMaker Studio supports ReadSpeaker's off-the-shelf Neural Premium voices. ReadSpeaker's Neural Standard and Custom Voices will be supported in a future release.

The voice types enabled for a customer account depend on the configured voice package.

Voices of different voice types cannot be used within the same script. All text blocks within a single script must use voices of the same voice type.

Support for HTML5 and CSS3

A web browser with support for HTML5 and CSS3 is required to access the full functionality of speechMaker Studio.

Price Model

speechMaker Studio is subject to a set-up fee and a fixed monthly or annual service and maintenance fee.

Audio production is charged based on the number of characters converted, in accordance with the pricing specified in the applicable Order Confirmation.

Intended Use

speechMaker Studio may be used only to convert text segments into audio files in accordance with this product specification. Any use outside the scope defined in this product specification is not considered intended use.

Disclaimer

Features described in this product specification are guaranteed only if the requirements set out in this product specification are met. Specifications and features described in this product specification are subject to change by ReadSpeaker without prior notice.