

Generate sound file from text

Generate sound files from text (aka *Text to speech for IVR Audio*) is a subscription based add-on module.

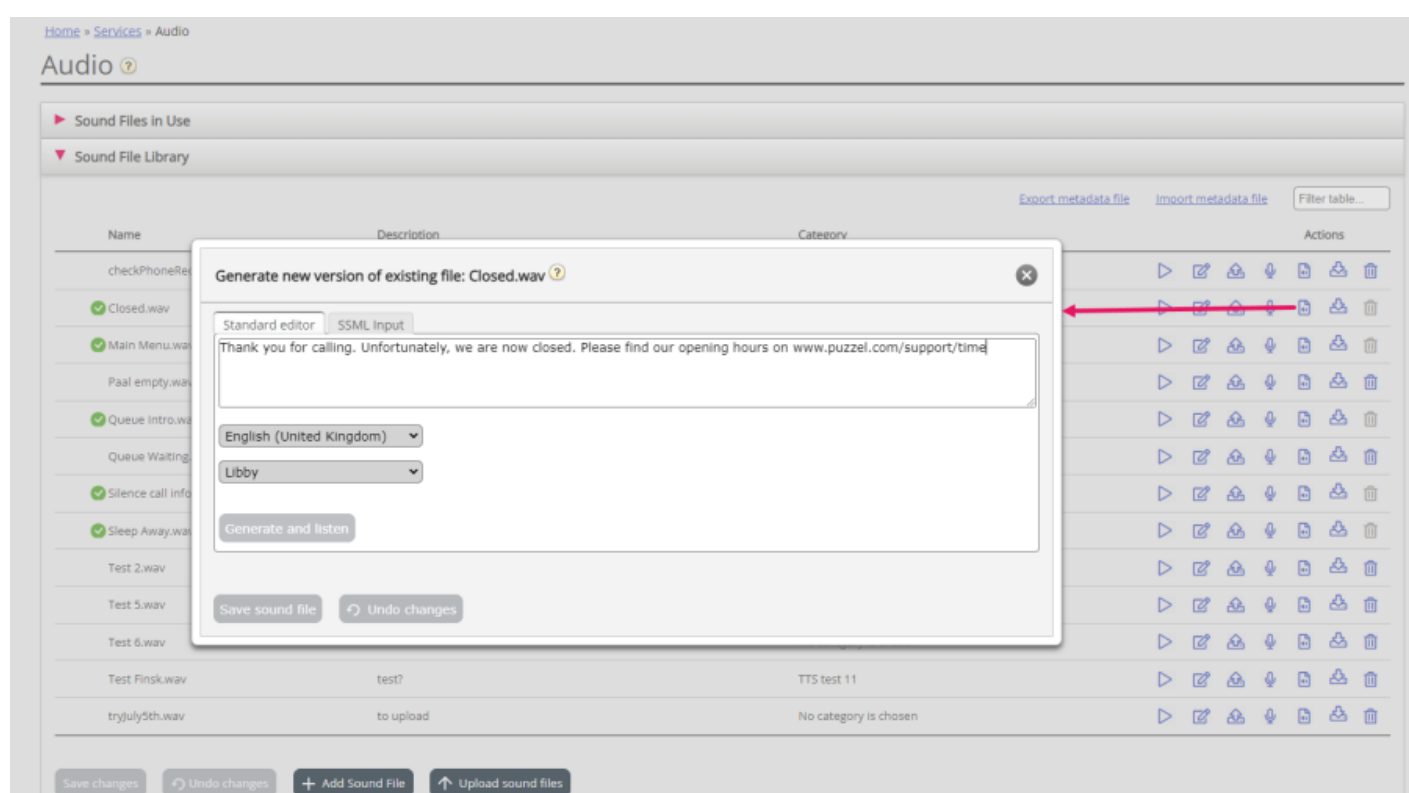
With Generate sound files from text you can generate consistent, high quality recordings in a few minutes without any external costs related to doing recordings.

We recommend that you create a new sound file row and **test** the Generate sound file functionality to find your preferred voice before you replace an existing uploaded or recorded sound file with a sound file generated from text.

You could also consider putting sound files generated from text in a separate **Category**, or possibly add a special prefix or suffix to the name for the sound files you generate from text.

Click the 'Generate' button for the relevant sound file row in the sound file library to open the window where you can generate a sound file based on text.

In the Generate window's **Standard editor** tab, add the **Script** (or update it), then select a **language** (most European languages are supported) and a **voice**, and then click 'Generate and listen'.



- A player will be shown (between Voice and the Generate button) and the generated sound file will automatically be played for you. Click the play button in the player once more if you want to listen to it again. Strange or special words might not sound so good.
- If you adjust the script or select another voice, you need to click 'Generate and listen' again (before you are allowed to save).
- If you are happy with the generated sound file, click 'Save sound file' and then confirm.

If you generated a sound file from text for a sound file row not currently in use, you can now start using this sound file in a module.

Sound files generated from text might be most useful for sound files you need to update often and/or sound files that need to be changed with a very short notice, e.g. emergency messages.

Please make sure you enter a script (the text) in the same language as the selected voice.

What about numbers and special characters?

- Check out how a (phone)**number** sounds depending on how you write it ("21898989" vs "21 89 89 89" vs "2 1 8. 9 8. 9 8 9")
- In general, brackets and quotation marks are ignored, so no point using such characters in the script.
- Special characters like <, >, #, &, % are not allowed.
- If you want to include a **web address** in your script (e.g. www.puzzel.com/support), make sure to listen to how it sounds in your language and adjust it if needed so that it sounds ok before you save it.
 - Is "/" announced as "forward slash", and is "." announced as "dot" in your language?

Please note that if you open the *Upload* or *Record* popup window for a sound file row that now uses a sound file generated from text, the script is "correct", so don't change the Script (from *Upload/Record*) unless you also upload a file or record a new file to replace this sound file.

SSML input

In some cases, a generated sound file using the Standard editor might not be good enough. In these cases, you can adjust the wording to make it sound ok, or record the sound file manually (and upload it). If you want to enhance the quality of a generated sound file, you can provide the SSML (Speech Synthesis Markup Language) input to adjust speed and pitch for words/phrases, add pauses etc. If you use SSML input instead of the Standard editor, you also need to include what voice to use.

The screenshot shows the 'Audio' section of the Puzzel interface. A modal window titled 'Generate new version of existing file: Test 2.wav' is open, displaying the 'SSML Input' tab. The SSML script is as follows:

```
<speak xmlns="http://www.w3.org/2001/10/synthesis" xmlns:mstts="http://www.w3.org/2001/mstts"
xmlns:emo="http://www.w3.org/2009/10/emotionml" version="1.0" xml:lang="en-GB"><voice name="Microsoft Server Speech
Text to Speech Voice (en-GB, RyanNeural)">That's remarkable! You're a genius! Mom said to her youngest son.
Customize output by <prosody rate="-40.00%"> slowing-down the speed rate.</prosody>
Add a break <break time="600ms" /> between words.</voice></speak>
```

Below the script, there are buttons for 'Generate and listen', 'Show script', 'Save sound file', and 'Undo changes'. The background shows a table of sound files with columns for Name, Description, Category, and Actions.

Puzzel's Text-To-Speech (TTS) is based on Microsoft Cognitive Services, so the Microsoft-specific elements are supported. Information about the supported syntax and elements can be found here: <https://docs.microsoft.com/en-us/azure/cognitive-services/speech-service/speech-synthesis-markup?tabs=csharp>

You can use an SSML editor, e.g. Microsoft's Audio content creator (<https://speech.microsoft.com/audiocontentcreation>) to do adjustments in speed, speaking style, add breaks etc, and then copy the SSML-formatted script and paste it into Puzzel's

SSML input tab.

If the content in the SSML input tab begins with '<' and ends with '>', you can click 'Generate and listen', but if the SSML format is not correct, an error message will be given after you click 'Generate and listen'.

If you're happy with the generated sound file after having listened to it, you can click 'Save sound file'.

If you have a sound file in your library generated based on SSML input, you can click "Show script" to see the plain text script. This works only if the syntax is correct, since we simply remove the parts within the 'brackets'.