

AudioGeneration Extension

User Guide & Installation Instructions

Overview

The Audio Generation Extension brings powerful AI-driven audio creation capabilities directly into your Adobe workflow. This comprehensive guide covers all features, from basic text-to-speech generation to advanced automation with MCP Server integration.

Key Features



Voiceover:
Generate text-to-speech audio files from text.



Sound Effect:
Create sound effects based on textual descriptions.



Storyboard:
Create a sequence of voiceover and sound effect entries.



Speech to Speech:
Convert existing audio into a different voice.



Voice Isolation:
Remove background noise from audio to isolate the voice.



Music:
Generate Music from a single text prompt.

System Requirements

The Audio Generation Extension allows you to create audio files using voiceovers, sound effects, voice isolation, and speech-to-speech conversion directly within Adobe Premiere Pro. You need an ElevenLabs API Key to use this Extension!

Compatible Adobe Software

- Adobe Premiere Pro (version 20.0 or later)
- Adobe After Effects (version 20.0 or later)
- Adobe Audition (version 20.0 or later) - Audio generation only (saves to output directory)

Operating System

- Windows 10 or later
- macOS 10.14 (Mojave) or later

Additional Requirements

- Active internet connection
- ElevenLabs API Key (obtain free key from <https://elevenlabs.io>)

Optional

- OpenAI API Key (obtain key from <https://openai.com/api/>)

Recommended

- 8 GB RAM or more
- Multi-core processor

Installation

Download the Extension

- Visit [Audio Generation Plugin](#)
- Download the Audio Generation Extension (the ZIP contains the extension and this guide)

Install an Extension Manager

You have two options for installing the extension:

Option 1: ZXP Installer

- Go to [ZXP Installer](#).
- Download and install the ZXP Installer appropriate for your operating system (Windows or macOS).

Option 2: Anastasiy's Extension Manager

- Go to [Anastasiy's Extension Manager](#).
- Download and install the Extension Manager for your operating system.

Install the Extension

Using ZXP Installer

- Launch the ZXP Installer.
- Drag and drop the downloaded Audio Generation Extension ZXP file into the ZXP Installer window.
- Click "Install" and follow any prompts.

Using Anastasiy's Extension Manager

- Launch Anastasiy's Extension Manager.
- Click "Install" and navigate to the downloaded Audio Generation Extension ZXP file.
- Follow the prompts to complete the installation.

Activate the Extension

- Launch Adobe Premiere Pro (or your compatible Adobe application).
- Go to Window > Extensions.
- Select "Audio Generation Extension" from the list.

Set Up the Extension

- When first launched, enter your email address to unlock all premium features. You'll receive a confirmation email; click the link and the extension activates itself. Use the same email on Premiere Pro, After Effects, and Audition.
- Enter your ElevenLabs API Key.
- Follow the on-screen instructions to complete the setup.
- Set your output directory
- Configure MCP Server directory (optional)
- Click "SAVE SETTINGS"

Optional: Dock the Extension

You can dock the Audio Generation Extension panel to your workspace sidebar for easy access.

Getting Started

Interface Overview

- Voiceover: For generating voiceover audio files
- Sound Effect: For generating sound effects
- Storyboard: Create a series of voiceover and sound effect entries
- Speech to Speech: Convert existing audio to a different voice
- Voice Isolation: Remove background noise from audio
- Music: Generate music from a single text prompt

Setting Up

- Your ElevenLabs API Key: Enter your ElevenLabs API key in the settings
- Your OpenAI API Key: Enter your OpenAI API key in the settings to use the inspiration feature
- Output Directory: Specify the path where you want to save the generated audio files. Use the "Browse" button to easily select a folder

Voiceover Generation

Generate professional text-to-speech audio with precise control over voice characteristics and delivery.

Basic Setup

- Select "Voiceover" as audio type
- Enter your text in the prompt field
- Choose your voice from the dropdown
- Select audio track for timeline placement

Voice Settings

- Voice: Select from your available ElevenLabs voices
- Voice Model: Choose AI model (Multilingual v2, Turbo v2.5, or v3)
- Stability (0.0-1.0): Lower = more variable, Higher = more stable. Recommended: 0.5-0.7. Below 0.3 may cause instability
- Similarity (0.0-1.0): Match to original voice characteristics
- Style (0.0-1.0): Amount of style exaggeration. Keep below 0.5 to avoid instability
- Speed (0.7-1.2): Speaking rate adjustment
- Speaker Boost: Enhance similarity to the voice (recommended ON)

Voice Model Selection

Choose the appropriate AI model for your needs:

- Eleven Multilingual v2: Standard model, supports most languages and delivers high-quality voiceovers
- Eleven Turbo v2.5: Faster generation time, optimized for English language content
- Eleven v3: Advanced model with audio tag support for emotional depth and expressive delivery

Character Limits

The extension enforces different character limits based on your activation status and selected voice model:

- Free (not activated): 100 characters per generation
- Activated users: 5,000 characters per generation
- Eleven v3 model: 3,000 characters per generation (due to API limitations)
- Control tags (break tags and audio tags) do not count toward these limits

Advanced Text-to-Speech Control

For precise control over speech delivery, the extension supports special control tags:

Break Tags (Pauses)

Add natural pauses to your voiceover using SSML break tags:

- `<break time="1s" />` - 1 second pause
- `<break time="500ms" />` - 500 millisecond pause
- Supports any duration (e.g., "0.5s", "2s", "750ms")

Example usage: Hello there `<break time="1s" />` How are you?

Audio/Emotion Tags (Eleven v3 only)

Add emotional depth and sound effects with audio tags (only works with Eleven v3 model):

- `[breathing heavily]` - Heavy breathing effect
- `[interrupts]` - Interruption effect
- `[laughs]`, `[sighs]`, `[whispers]` - Various emotional expressions
- Place these tags within your text where the effect should occur

Example usage: I can't believe it `[laughs]` this is amazing `[whispers]`

Character Counter

The character counter intelligently handles control tags:

- Break tags and audio tags are automatically excluded from the character count
- Only actual spoken text counts toward your character limit
- Example: "Hello `<break time="1s" />` `[laughs]` World" counts as only "Hello World" (11 characters)
- This gives you more creative freedom without hitting character limits
- The counter displays as "current/max" format (e.g., "450/5000")
- Color changes: Yellow at 80% capacity, Red at 95%

Complete Example with Control Tags

Hello there `<break time="0.5s" />` `[laughs]` I can't believe it! `<break time="1s" />` This is amazing `[whispers]` absolutely amazing.

This example demonstrates pauses, laughter, and whispers for expressive voiceover delivery.

Sound Effects

Create realistic sound effects from textual descriptions.

Basic Setup

- Select "Sound Effect" as audio type
- Enter a description of the sound you want
- Select audio track for timeline placement

Sound Effect Settings

- Duration: Set the duration of the sound effect in seconds (Automatically pick the best length / Fixed length)
- Prompt Influence: Adjust the influence of the prompt on the sound effect (More Creative / Follow Prompt)

Speech to Speech

Convert existing audio files to different voices while preserving timing and emotion.

- Select Audio File: Choose the source audio file you want to convert
- Target Voice: Select the voice you want to convert the audio to
- Adjust voice settings similar to the Voiceover section

Voice Isolation

Remove background noise and isolate the voice from audio recordings.

- Select Audio File: Choose the audio file you want to clean

Music

Generate original music tracks from text descriptions.

- Length: Auto-detect duration based on your prompt, or set a specific length up to 5 minutes (300 seconds)
- Instrumental: Toggle this to create purely instrumental music without vocals

Storyboard

Create and manage complex audio sequences with multiple voiceover and sound effect entries.

Creating a Storyboard

- Select the "Storyboard" option
- Click "Add Entry" to create a new segment
- Configure each entry with voiceover or sound effect settings
- Click "PROCESS ALL ENTRIES" to generate all audio files sequentially
- Place your timeline marker and confirm the position for each audio

CSV Import

- Click the folder icon to upload a CSV file containing multiple prompts and settings
- An example CSV can be found in the extension folder
- After uploading, click the "PROCESS CSV" button to generate audio files based on the CSV data

Storyboard Export

Save your current storyboard entries as a CSV file for later use, sharing, or batch processing.

- When you have two or more entries, an "Export Storyboard as CSV" button appears
- Click the button to create a CSV file containing all your storyboard entries and settings
- The exported CSV file is saved in the output directory specified in your settings
- You can reimport this CSV file to recreate your storyboard setup or process it as a batch job
- The exported CSV is fully compatible with the extension's CSV import feature

Marker Placement

- Set a marker on the Adobe Premiere Pro timeline where you want the audio to start
- Click the "MARKER SET" button to confirm the marker placement

Settings

The Settings area allows you to customize and optimize how the extension functions.

API Settings

- OpenAI API Key: Required for the Inspiration and Enhance features
- Select OpenAI Model: Choose the AI model for Inspiration/Enhance features
- ElevenLabs API Key: Essential for all core functionalities. You can obtain a free API key from the ElevenLabs website
- Show ElevenLabs Legacy Voices: Enable this to display older voice models in the voice selection dropdown

General Settings

- MP3 Output Quality: Select quality from lower (22.05kHz, 32kbps) to high quality (44.1kHz, 192kbps)
- Output Directory: Choose where generated audio files will be saved. Use the "Browse" button to select a directory

OpenAI Inspiration / Enhance Settings

These settings control how the extension enhances text inputs:

- Voiceover Prompt: Customize the default text for voiceover inspiration. The {PROMPT} variable will be replaced by the user's input
- Sound Effect Prompt: Default text for sound effect inspiration. Creates vivid descriptions of desired sounds
- Enhance Prompt: Default prompt for the ElevenLabs v3 Audio Tagger. The Auto-Tagger analyzes your text and adds suitable audio tags
- Enhance Music Prompt: Creates lyrics and music style. The AI creates a fitting music style and writes lyrics based on your description

Clip Label Color Settings

Customize Adobe Premiere Pro label colors for different audio types in your timeline:

- Voiceover (TTS) Label Color
- Sound Effect (SFX) Label Color
- Speech to Speech (STS) Label Color
- Audio Isolation (ISO) Label Color
- Music (MUSIC) Label Color

Activation

Unlock all premium features for free. Enter your email address, confirm via the link we send, and the extension unlocks automatically. Your activation status is shown here (and in the activate banner at the top of the panel).

Settings Management

Export and import your settings for backup or sharing with team members:

- Export Settings: Click "EXPORT SETTINGS" to save your configuration as a .json file
- Import Settings: Click "IMPORT SETTINGS" and select a .json file to restore all settings instantly

Voice Design

Create custom voices by providing detailed voice descriptions and sample text. The AI will analyze your inputs and generate unique voice previews that match your specifications.

Creating Custom Voices

- Access: Click the voice design icon in the header to open the voice creation interface
- Voice Description: Enter a detailed description (minimum 40 characters) including age, gender, accent, emotional tone, speaking style, professional background, and unique vocal qualities
- Generation Text: Provide sample text (minimum 100 characters) that the AI will use to generate voice variations
- Preview Options: Multiple voice variations will be generated based on your inputs
- Listening: Use the play button to preview each generated voice variation
- Adding to Library: Click "ADD TO LIBRARY" for your preferred voice variation
- Naming: Give your custom voice a name and description for easy identification

Best Practices for Voice Design

- Be specific in your voice descriptions
- Use detailed characteristics rather than vague terms
- Provide natural, conversational sample text
- Test different variations of descriptions for optimal results

History

Access, view, and reuse your previously generated audio content.

- Accessing History: Click the history icon in the extension interface
- Voice Filter: Use the voice dropdown to filter history entries by specific voices
- Entry Count: Adjust the number of entries to display (5 to 1000)
- Navigation: Use pagination controls to browse through your history
- Restoring Settings: Click the "Restore" button next to any history entry to apply its settings
- Inserting Clips: Use the "Insert Clip" button to add a past generation to your timeline at the current playhead position
- API Usage: This feature utilizes the ElevenLabs API. Ensure you have entered a valid ElevenLabs API key

Inspiration Feature

The inspiration feature uses AI to improve your input and enhance creative content.

How It Works

- For Voiceover: Refines dialogue, making it more engaging and natural while preserving the original meaning
- For Sound Effects: Creates vivid, detailed descriptions of sounds, including unique characteristics
- For Eleven v3: Analyzes and creates matching audio tags for your script
- For Music: Creates matching music/instrumental style or lyrics

Using the Feature

- Enter your initial prompt
- Click the lightbulb/star icon
- The AI will generate an enhanced version of your text
- You can use or further modify the enhanced text

Note: This feature uses the OpenAI API. Ensure you have entered a valid OpenAI API key in the settings.

MCP Server Integration

The extension includes a Model Context Protocol (MCP) server for advanced automation and integration with external tools.

Setup

- Set the MCP Server Directory in the extension settings
- The server runs automatically when the extension is active
- External tools can connect via the configured directory

Capabilities

- Batch Processing: Generate multiple audio files from scripts automatically
- Timeline Automation: Programmatic timeline manipulation and control
- External Script Integration: Control the extension from external applications
- Workflow Optimization: Automate repetitive tasks and streamline production

Example Use Cases

- Generate voiceovers for an entire video script automatically
- Integrate with video editing workflows for seamless production
- Create automated audio pipelines for content production
- Build custom tools and integrations for your specific workflow needs

For technical details, refer to the MCP documentation in the extension folder.

Best Practices

Voiceover Tips

- Use shorter sentences for more natural speech rhythm
- Add break tags between paragraphs for better pacing
- With Eleven v3, use audio tags sparingly for maximum impact
- Preview voices before committing to long generations
- Start with stability around 0.6 and adjust based on results
- Keep style exaggeration below 0.5 to avoid instability

Sound Effect Tips

- Be specific and descriptive in your prompts
- Include environment details (indoor/outdoor, room size, acoustics)
- Mention materials and actions for realistic sounds
- Use the Inspiration feature to improve descriptions
- Test different prompt influence values for varied results

Storyboard Workflow

- Plan your audio sequence before creating entries
- Use consistent voices for character continuity
- Export storyboards as CSV for backup and reusability
- Test timing with markers before bulk generation
- Keep entries organized with clear naming

Performance

- Use lower quality settings for quick previews
- Save high-quality rendering for final exports
- Keep your API keys secure and never share them
- Monitor your ElevenLabs API usage in their dashboard
- Export your settings regularly for backup

Frequently Asked Questions

Q: Why isn't the Inspiration button showing?

A: The Inspiration feature requires an OpenAI API key. Add your key in Settings to enable this feature.

Q: Can I use the extension offline?

A: No, an active internet connection is required for all AI generation features as they connect to ElevenLabs and OpenAI APIs.

Q: How do I get more characters per generation?

A: Activate the extension with your email to unlock 5,000 characters for standard models (3,000 for Eleven v3).

Q: What's the difference between voice models?

A: Eleven v3 supports audio tags for emotions, Turbo v2.5 is faster, Multilingual v2 is the standard all-purpose model.

Q: Can I import my own voices?

A: Yes, voices created in your ElevenLabs account automatically appear in the voice dropdown.

Q: Do control tags use up my character limit?

A: No! Break tags (<break>) and audio tags ([emotion]) don't count toward your character limit.

Q: How do I use break tags?

A: Insert <break time="1s" /> into your text where you want pauses. Use seconds (s) or milliseconds (ms).

Q: Can I use audio tags with any voice model?

A: Audio tags like [laughs] and [whispers] only work with the Eleven v3 model. Other models will ignore them.

Q: How do I activate the extension?

A: Enter your email in the panel and confirm the link we email you; all premium features unlock instantly. Use the same email across all three Adobe apps.

Tips and Tricks

- First Entry: When selecting Storyboard, the first entry is created automatically
- Removing Entries: Use the "Remove Entry" button to delete any unwanted segments
- Preview Voices: Use the play button next to the voice dropdown to preview the selected voice
- Dynamic Button Text: The "GENERATE" button text changes based on the selected audio type for clearer action indication
- Character Counter: Watch the color changes - yellow means you're at 80% capacity, red means 95%
- Control Tags: Experiment with different break durations and audio tags to find the perfect delivery

Troubleshooting

- API Errors: Ensure your ElevenLabs or OpenAI API key is correct and has sufficient credits
- File Saving Issues: Verify the save path is correct and accessible. Use the "Browse" button to ensure a valid directory is selected
- Marker Placement: Ensure you have set and confirmed the marker before generating audio
- Audio File Selection: For Speech to Speech and Voice Isolation, make sure you've selected a valid audio file before generating
- Control Tags Not Working: Verify you're using the correct syntax and, for audio tags, that you have Eleven v3 selected
- Character Counter Issues: If the counter seems incorrect, check for hidden characters or formatting issues in your text

Support

For further assistance, please refer to the support documentation or visit the [Changelog page](#) to stay updated on the latest features and improvements. If you need direct support, feel free to contact me at info@tbitdesign.com.

Copyright © Thomas Breher - TBIT DESIGN / tbitdesign.com

Note: This extension was independently developed and is not affiliated with or endorsed by Adobe Systems Incorporated or ElevenLabs®. Adobe®, Adobe Premiere Pro®, Adobe After Effects® and Adobe Audition® and are registered trademarks of Adobe Systems Incorporated. ElevenLabs is not responsible for the content or functionality of this extension. All other trademarks are the property of their respective owners.

powered by

ElevenLabs