VOICE2TEXT AI

ユーザーガイド

Voice2Text

声を文字に、どんな言語でも。

音声・動画ファイルを AI で文字起こしし、テキストを修正して好きな場所に保存できます。

オフライン Whisper　12 形式　クラウド　9 言語



バージョン 1.1　　macOS · Windows · Linux　　日本語

目次

01	Voice2Text へようこそ	3
02	インストール	4
03	メインウィンドウ	5
04	エンジンと言語の選択	6
05	オンラインサービス：音声はパソコンの外へ送られます	8
06	文字起こし中	9
07	文字起こしの修正	10
08	保存とエクスポート	12
09	クラウドへの保存	14
10	設定と言語	16
11	キーボードショートカット	18
12	よくある質問	19

01 Voice2Text へようこそ

> Voice2Text は音声や動画の録音を文字に変換します。インタビュー、会議、講義、ポッドキャスト、ボイスメッセージ。ファイルをウィンドウにドロップして AI エンジンを選べば、数分で編集・保存できるテキストが得られます。

できること

- **あらゆる形式に対応。** MP3、M4A、WAV、FLAC、OGG、AAC、AIFF、MP4・MOV・MKV・WebM の動画など。ffmpeg を入れればほぼすべての形式に対応します。
- **インターネットなしでも。** Whisper はパソコン上で直接動作し、音声が外部に出ることはありません。
- **話者の識別。** AssemblyAI と Deepgram では声が区別され、各話者に名前を付けられます。
- **かんたん修正。** テキストは音声と結び付いた文に分かれています。クリックひとつで該当箇所を聞き直せます。
- **多彩な出力形式。** Word、PDF、ODT、RTF、TXT、Markdown、HTML、SRT・VTT 字幕、CSV、JSON、そして再編集できるプロジェクト。
- **クラウドに直接保存。** Google Drive、OneDrive、Dropbox、Box、Proton Drive、iCloud などに、パソコン上のフォルダ経由で保存します。
- **9 つの言語。** イタリア語、英語、フランス語、スペイン語、ドイツ語、中国語、日本語、ロシア語、アラビア語。翻訳は修正可能な XML ファイルです。

動作環境

macOS 11 以降（Apple シリコンまたは Intel の Mac）、64 ビット版 Windows 10/11、または 64 ビット版 Linux（Ubuntu、Debian、Fedora とその派生版）。オフライン Whisper には 8 GB 以上のメモリを推奨します。オンラインサービスはインターネット接続だけで使えます。

02 インストール

> お使いのシステム用のファイルをダウンロードし、ほかのアプリと同じようにインストールします。

macOS

`Voice2Text_1.1.1_universal.dmg` を開き、Voice2Text のアイコンを**アプリケーション**フォルダにドラッグします。初回起動時に「開発元を確認できません」と表示された場合は、**システム設定** → **プライバシーとセキュリティ**で**このまま開く**を選ぶか、アプリを右クリックして**開く**を選んでください。操作は一度だけです。

Windows

`Voice2Text_1.1.1_x64-setup.exe` を実行します。青い Windows SmartScreen の画面が表示されたら、**詳細情報**をクリックし、**実行**を選びます。管理者権限は不要です。

Linux

Ubuntu と Debian では `.deb` パッケージ（ダブルクリック、または `sudo apt install ./Voice2Text_1.1.1_amd64.deb`）、Fedora では `.rpm` パッケージをインストールします。`.AppImage` を実行可能にして（`chmod +x`）起動することもできます。

> **任意：ffmpeg。**一般的な形式は Voice2Text だけで読み込めます。まれな形式（OPUS、WMA、AMR、AVI…）には ffmpeg をインストールしてください。Mac は `brew install ffmpeg`、Windows は `winget install ffmpeg`、Linux は `sudo apt install ffmpeg`。自動的に検出されます。

03 メインウィンドウ

> 作業はすべてひとつのウィンドウで行います。左にファイル一覧、中央に選択中のファイルが表示されます。



図 1. スタート画面：枠内にファイルをドロップするか、「ファイルを選択」を使います。

- **上部バー**：**ファイルを追加**、**文字起こしを開く**（保存済みの `.v2t` プロジェクト）、**設定**の歯車。
- **左のファイル一覧**：各ファイルの状態（準備完了、待機中、文字起こし中、文字起こし済み、エラー）と進捗が表示されます。名前の後ろの点は未保存の変更があることを示します。
- **中央エリア**：文字起こしの設定、進捗、または修正するテキスト。

ファイルの追加

音声・動画ファイルをウィンドウにドラッグするか、**ファイルを追加**（`⌘O` / `Ctrl+O`）を使います。いくつでも追加でき、順番に文字起こしされます。`.v2t` ファイルをドロップすると、保存済みの文字起こしが開きます。

04 エンジンと言語の選択

> 文字起こしの前に、作業を担当するエンジンを選びます。パソコン上の Whisper か、オンラインサービスです。



図 2. 選択中のファイルのエンジンと言語を選ぶ画面。

エンジン	料金	特長
Whisper（オフライン）	無料	インターネット不要。モデルは一度だけダウンロード（45〜600 MB）。オンラインサービスより低速。
Groq Whisper	キーで無料	非常に高速。長い音声は自動で分割されます。
AssemblyAI	無料クレジット	話者を識別。インタビューや会議に最適。
Deepgram	無料クレジット	話者を識別し、非常に高速。
Whisper server	無料	パソコンまたはネットワーク上の Whisper サーバー（whisper.cpp、faster-whisper-server、LocalAI）。
OpenAI	有料	OpenAI の Whisper。

音声の言語

自動判別を選ぶと、Voice2Text が話されている言語を自動で判断します。言語がわかっている場合は選択してください。特に冒頭部分の精度が上がります。

オンラインサービスのキー

Groq、AssemblyAI、Deepgram、OpenAI には個人のキーが必要です。**設定** → **エンジン**に各サービスの入力欄と「キーを取得」リンクがあります。Groq の登録は無料で、AssemblyAI と Deepgram は登録時に無料クレジットがもらえます。



図 3. 設定 → エンジン：Whisper のモデルと各サービスのキー。

文字起こしを押すと開始します。複数のファイルが待機中のときは、サイドバーに**すべて文字起こし**も表示されます。

05 オンラインサービス：音声はパソコンの外へ送られます

> オンラインの AI サービスを使うと、音声ファイルはインターネット経由でアップロードされます。Voice2Text はそのことをはっきり伝え、確認を求めます。



図 4. オンラインサービスへの初回送信前に表示される必須の警告。

Groq、AssemblyAI、Deepgram、OpenAI を初めて使うときにこの警告が表示されます。続行するには**「読んで理解しました」**にチェックを入れる必要があります。すぐに**オフライン Whisper** に切り替えることもできます。

音声ファイルは外部サービスのサーバーにアップロードされます。外部サービスは独自の利用規約とプライバシーポリシーに従ってファイルを扱い、一定期間保存したり、他の国で処理したりする場合があります。

機密情報を含む録音はアップロードしないでください。非公開・秘密の情報、個人データ・健康データ・第三者のデータも、権利と録音された人の同意がない限りアップロードしないでください。

Voice2Text とその作者は一切の責任を負いません。アップロードされたファイルの内容、機密性、外部サービスによる利用についても同様です。責任はすべてファイルをアップロードする人にあります。

同意はサービスごとに有効で、日付とともに記録されます。オンラインサービスが選択されている間は、オレンジ色の枠で注意が表示されます。**設定** → **情報**で警告を読み直し、再表示させることができます。

機密性の高い音声ですか？Whisper（オフライン）を使えば、文字起こしはすべてパソコン上で行われます。

06 文字起こし中

> Voice2Text は作業の進み具合を常に表示します。

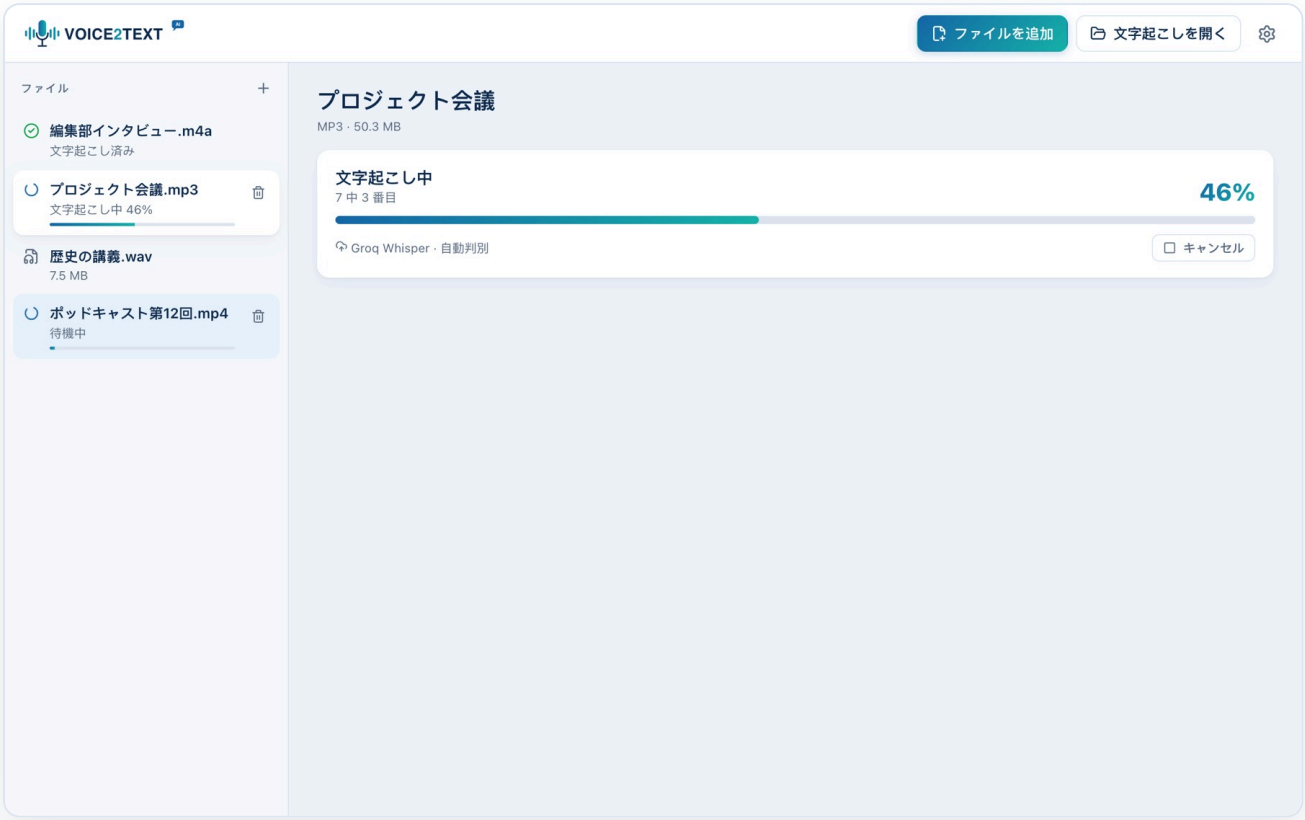


図 5. Groq による文字起こし中：進捗と処理中の区間。

- **音声を読み込み中**：パソコン上でファイルをデコードします。
- **モデルを準備中**（オフライン Whisper の初回のみ）：選んだモデルをダウンロードして保存します。
- **音声をアップロード中・文字起こし中**：長いファイルでは処理中の区間（「7 中 3 番目」）が表示されます。
- **サービス待ち**：無料プランでリクエスト上限を超えた場合、Voice2Text が待機して自動で再試行します。

キャンセルでいつでも中止できます。ほかのファイルの作業を続けてもかまいません。テキストができあがると通知が届きます。

07 文字起こしの修正

テキストは文ごとに分かれ、それぞれに時間と、識別できた場合は話者が付いています。



図 6. エディタ：上部に音声プレーヤー、その下に編集可能な文。再生中の文が強調表示されます。

聞く

上部のプレーヤーで元の音声を再生します。`スペース`で再生・一時停止（入力中でないとき）、ボタンで 5 秒移動、速度は 0.75× から 2× まで選べます。文の時間をクリックするとその箇所を再生します。**音声に追従**をオンにすると、再生中の文が常に表示されます。

編集する

- 文をクリックして入力すると、すぐに反映されます。
- `Enter` でカーソル位置で文を分割、`Shift+Enter` で同じ文の中で改行します。
- 文の先頭で `Backspace` を押すと前の文と結合し、**結合**ボタンで次の文と結合します。
- ごみ箱で文を削除します（言葉として認識された雑音など）。



図 7. 検索と置換、話者の名前。

検索と置換

検索（⌘F / Ctrl+F）は語を含む文を強調表示し、**すべて置換**で誤記された名前を文字起こし全体で一度に修正できます。

話者

AssemblyAI と Deepgram では、各文にラベル（話者 A、話者 B...）が付きます。**話者**パネルで実際の名前を入力すると、書き出したファイルを含むすべての場所に反映されます。時間の横のメニューで文の話者を変更したり、話者を追加したり、エンジンが識別しない場合でも手動で割り当てたりできます。

08 保存とエクスポート

> プロジェクトを保存すれば後から修正を続けられます。テキストは必要な形式でエクスポートできます。

.v2t プロジェクト

保存（`⌘S` / `Ctrl+S`）で、文字起こし・修正内容・話者名を含む `.v2t` ファイルを作成します。**文字起こしを開く**（`⇧⌘O` / `Ctrl+Shift+O`）またはウィンドウへのドロップで再び開けます。`⇧⌘S` で別名保存します。

エクスポート

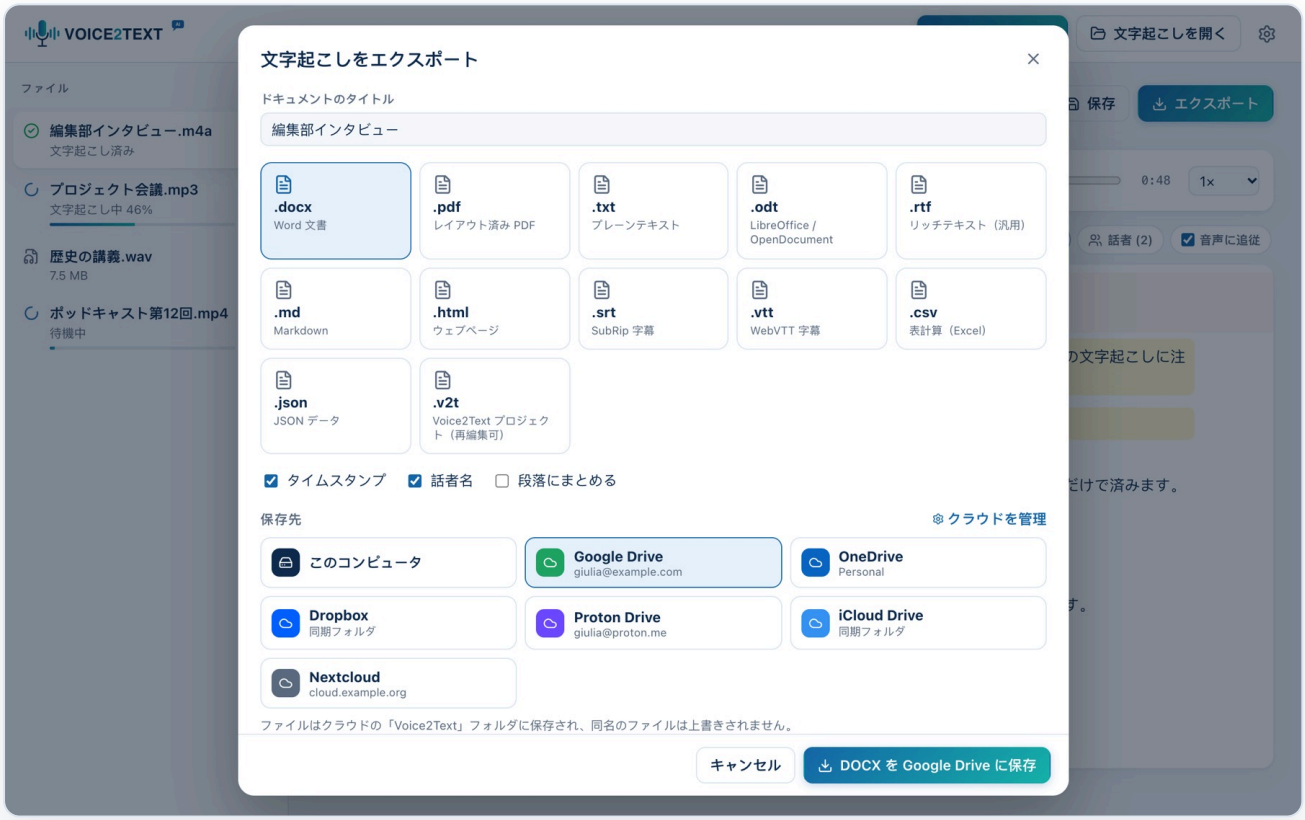


図 8. エクスポート画面：形式、オプション、保存先。

エクスポート（`⌘E` / `Ctrl+E`）を押し、ドキュメントのタイトルを入力して形式を選びます。

形式	用途
.docx Word · **.odt** LibreOffice · **.rtf**	編集用の文書。タイトルと録音情報付きでレイアウト
.pdf	印刷・送付用。システムの印刷ダイアログ（「PDF として保存」）で作成
.txt · **.md** · **.html**	プレーンテキスト、Markdown、ウェブページ
.srt · **.vtt**	動画の字幕（YouTube、Premiere、DaVinci、VLC…）
.csv · **.json**	表計算（Excel、Numbers）やほかのソフト用のデータ
.v2t	再編集できるプロジェクト。クラウドにも保存可能

オプションで**タイムスタンプ**や**話者名**を含めるか、**段落にまとめる**かを選べます。アラビア語など右から左に書く言語も正しくレイアウトされます。

09 クラウドへの保存

> Voice2Text は、各サービスのアプリがパソコン上に作るフォルダを使って、クラウドへ直接保存します。アカウントの接続やパスワードの入力は不要です。

しくみ

Google Drive、OneDrive、Dropbox、Box、**Proton Drive**、iCloud Drive、pCloud、MEGA、Nextcloud、Sync.com のアプリをインストールすると、パソコンに同期フォルダができます。Voice2Text はこれを自動で見つけます。**エクスポート**画面の**保存先**でクラウドを選んでボタンを押すと、ファイルは `Voice2Text` サブフォルダに書き込まれ、サービスのアプリがすぐにアップロードします。

- ファイルが上書きされることはありません。同じ名前のファイルがある場合、新しいファイルは「名前 (2)」になります。
- `.v2t` プロジェクトを含め、どの形式でもクラウドに保存できます。
- PDF は印刷ダイアログ経由です。そこで「PDF として保存」とクラウドのフォルダを選びます。



図 9. 設定 → クラウド：パソコン上で見つかったフォルダと WebDAV 接続。

設定 → クラウド

- **フォルダ名**：クラウドに作成するサブフォルダ（既定は `Voice2Text`）。
- **同期フォルダ**：見つかったクラウドの一覧。フォルダのアイコンで開きます。見つからないサービスは、下のリンクからアプリを入手できます。
- **別のフォルダを追加**：ほかのクラウドやネットワークドライブ（NAS）用。

- **WebDAV**：アプリを入れていない Nextcloud、ownCloud、Koofr、pCloud、Yandex Disk、Synology/QNAP の NAS 用。サービスを選び、アドレス、ユーザー、アプリ パスワードを入力して**接続テスト**を押します。

> **Proton Drive** はほかのアプリからアカウントへの直接アクセスを許可していないため、Voice2Text は同じエンドツーエンド暗号化が保たれる同期フォルダを使います。

10 設定と言語

右上の歯車（⌘, / Ctrl+,）で設定を開きます。



図 10. 設定 → 一般。

- **一般**：表示言語、ライト・ダーク・自動の外観、既定のエンジンと音声の言語、文字起こしの自動開始。
- **エンジン**：オフライン Whisper のモデル（最小、ベース、小、大 turbo）、モデルの使用容量、サービスのキー、ローカル Whisper サーバー、ffmpeg の検出。
- **クラウド**：同期フォルダと WebDAV（前の章）。
- **言語**：インターフェースの翻訳ファイル。
- **情報**：バージョン、ライセンス、オンラインサービスに関する警告。

言語は XML ファイル



図 11. 設定 → 言語：利用できる言語、完成度、XML ファイル用ツール。

インターフェースの各言語は XML ファイルです。翻訳を修正したり新しい言語を追加したりするには：

1. 言語ファイルをエクスポートする（行の矢印）か、**新しい言語のテンプレート**をダウンロードします。
2. テキストエディタで開き、タグの間のテキストだけを翻訳します。`key` 属性や `{n}` のような波括弧のプレースホルダーは変更しないでください。
3. **XML ファイルを読み込む…**を押すと、言語がすぐに一覧と言語メニューに表示されます。

```
<language code="pt" name="Português" english="Portuguese" dir="ltr" version="1.0">
  <string key="app.addFiles">Adicionar ficheiros</string>
  <string key="progress.piece">parte {i} de {n}</string>
</language>
```

内蔵言語と同じコードのファイルはそれを置き換えます（修正用）。いつでも元に戻せます。未翻訳の文は英語で表示され、**完成度**の列で翻訳済みの割合がわかります。右から左に書く言語には `dir="rtl"` を使います。

11 キーボードショートカット

> Windows と Linux では ⌘ の代わりに Ctrl を使います。

キー	操作
⌘O	音声・動画ファイルを追加
⇧⌘O	保存した文字起こし（.v2t）を開く
⌘S · ⇧⌘S	保存 · 別名で保存
⌘E	エクスポート
⌘F	検索と置換
⌘,	設定
スペース	再生 / 一時停止（テキスト外で）
Enter · Shift+Enter	文を分割 · 文の中で改行
文の先頭で Backspace	前の文と結合
Esc	テキストから出る · 開いている画面を閉じる

12 よくある質問

> よくある疑問への回答です。

ファイルの形式が読み込めません。

ffmpeg をインストールする（「インストール」の章）か、元のファイルをそのまま受け取るためどの形式にも対応する AssemblyAI または Deepgram を使ってください。

オフライン Whisper が遅いです。

設定 → エンジンで**ベース**または**最小**のモデルを選んでください。「大 turbo」は最も高精度ですが、グラフィックカードがないと時間がかかります。長いファイルには Groq がずっと速いです。

テキストが別の言語になっている、または翻訳されている。

自動判別ではなく音声の言語を指定し、**再度文字起こし**を使ってください。

「サービスのキーが無効です」と表示される。

設定 → エンジンでキーを確認し、サービスのサイトから空白を含めずにコピーし直してください。

クラウドが表示されません。

サービスのアプリがインストールされ同期中であることを確認し、設定 → クラウドで**再検索**を押してください。フォルダを手動で追加することもできます。

プレーヤーに「音声を利用できません」と表示される。

プロジェクトの保存後に元の音声ファイルが移動または名前変更されています。テキストはそのまま使えます。聞き直すにはファイルを元の場所に戻してください。

キーや設定はどこに保存されますか。

お使いのパソコンのアプリ設定ファイルにのみ保存され、自分のユーザーだけが読み取れます。Voice2Text に送信されることはありません。



VOICE2TEXT

声を文字に、どんな言語でも。

Voice2Text は overminder のソフトウェアです。PolyForm Strict 1.0.0 ライセンス：非商用利用は可能。再配布・改変・二次的著作物は不可。

© 2026 overminder · バージョン 1.1