VOICE2TEXT AI

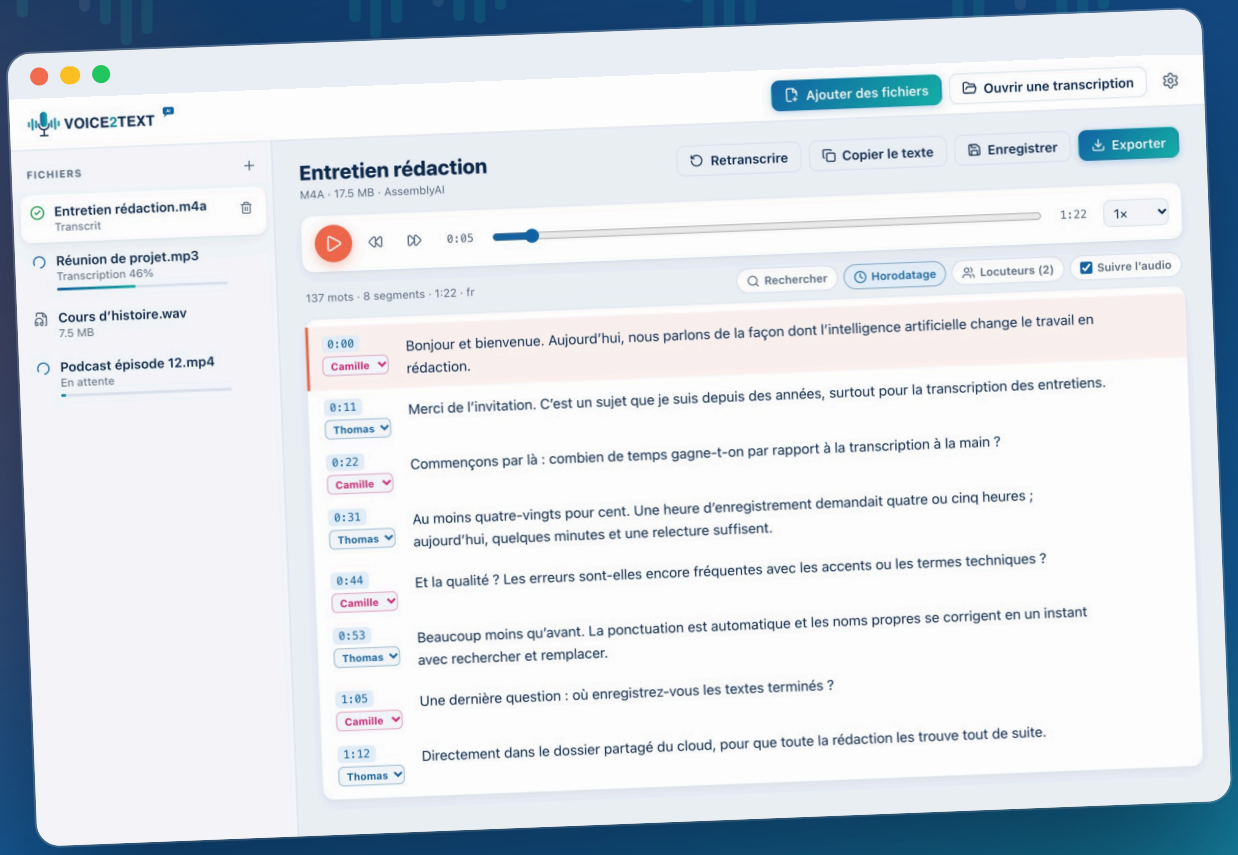
MANUEL D'UTILISATION

Voice2Text

De la voix au texte, dans toutes les langues.

Transcrivez vos fichiers audio et vidéo grâce à l'intelligence artificielle, corrigez le texte et enregistrez-le où vous voulez.

Whisper hors ligne | 12 formats | Cloud | 9 langues



Version 1.1 | macOS · Windows · Linux | Français

Sommaire

01	Bienvenue dans Voice2Text	3
02	Installation	4
03	La fenêtre principale	5
04	Choisir le moteur et la langue	6
05	Services en ligne : votre audio quitte l'ordinateur	8
06	Pendant la transcription	10
07	Corriger la transcription	11
08	Enregistrer et exporter	13
09	Enregistrer dans le cloud	15
10	Réglages et langues	17
11	Raccourcis clavier	19
12	Questions fréquentes	20

01 Bienvenue dans Voice2Text

> Voice2Text transforme les enregistrements audio et vidéo en texte écrit. Entretiens, réunions, cours, podcasts, messages vocaux : déposez le fichier dans la fenêtre, choisissez le moteur d’IA et, en quelques minutes, vous obtenez le texte, prêt à être corrigé et enregistré.

Ce que vous pouvez faire

- **Tous les formats.** MP3, M4A, WAV, FLAC, OGG, AAC, AIFF, vidéos MP4, MOV, MKV, WebM et bien d’autres ; avec ffmpeg installé, pratiquement tous.
- **Même sans Internet.** Whisper fonctionne directement sur l’ordinateur : l’audio ne quitte jamais votre bureau.
- **Qui parle.** Avec AssemblyAI et Deepgram, les voix sont séparées et vous pouvez donner un nom à chaque personne.
- **Correction facile.** Le texte est découpé en phrases liées à l’audio : un clic et vous réécoutez le passage exact.
- **Tous les formats de sortie.** Word, PDF, ODT, RTF, TXT, Markdown, HTML, sous-titres SRT et VTT, CSV, JSON et le projet réouvrable.
- **Directement dans le cloud.** Google Drive, OneDrive, Dropbox, Box, Proton Drive, iCloud et d’autres, via leurs dossiers sur l’ordinateur.
- **Neuf langues.** L’interface parle italien, anglais, français, espagnol, allemand, chinois, japonais, russe et arabe ; les traductions sont des fichiers XML que vous pouvez corriger.

Configuration requise

macOS 11 ou ultérieur (Mac Apple ou Intel), Windows 10/11 64 bits, ou Linux 64 bits (Ubuntu, Debian, Fedora et dérivées). Pour Whisper hors ligne, au moins 8 Go de mémoire sont conseillés ; les services en ligne demandent seulement une connexion.

02 Installation

> Téléchargez le fichier pour votre système et installez-le comme n'importe quel logiciel.

macOS

Ouvrez `Voice2Text_1.1.1_universal.dmg` et faites glisser l'icône de Voice2Text dans le dossier **Applications**. Au premier lancement, macOS peut indiquer que le développeur n'est pas vérifié : ouvrez **Réglages Système** → **Confidentialité et sécurité** et choisissez **Ouvrir quand même**, ou faites un clic droit sur l'app et choisissez **Ouvrir**. Une seule fois suffit.

Windows

Lancez `Voice2Text_1.1.1_x64-setup.exe`. Si la fenêtre bleue de Windows SmartScreen apparaît, cliquez sur **Informations complémentaires** puis sur **Exécuter quand même**. L'installation ne demande pas de droits d'administrateur.

Linux

Sur Ubuntu et Debian, installez le paquet `.deb` (double-clic ou `sudo apt install ./Voice2Text_1.1.1_amd64.deb`) ; sur Fedora, le paquet `.rpm`. Vous pouvez aussi rendre l'`.AppImage` exécutable (`chmod +x`) et la lancer.

Facultatif : ffmpeg. Voice2Text lit seul les formats les plus courants. Pour les formats rares (OPUS, WMA, AMR, AVI…), installez ffmpeg : `brew install ffmpeg` sur Mac, `winget install ffmpeg` sur Windows, `sudo apt install ffmpeg` sur Linux. Il est détecté automatiquement.

03 La fenêtre principale

> Tout se passe dans une seule fenêtre : la liste des fichiers à gauche, le fichier sélectionné au centre.

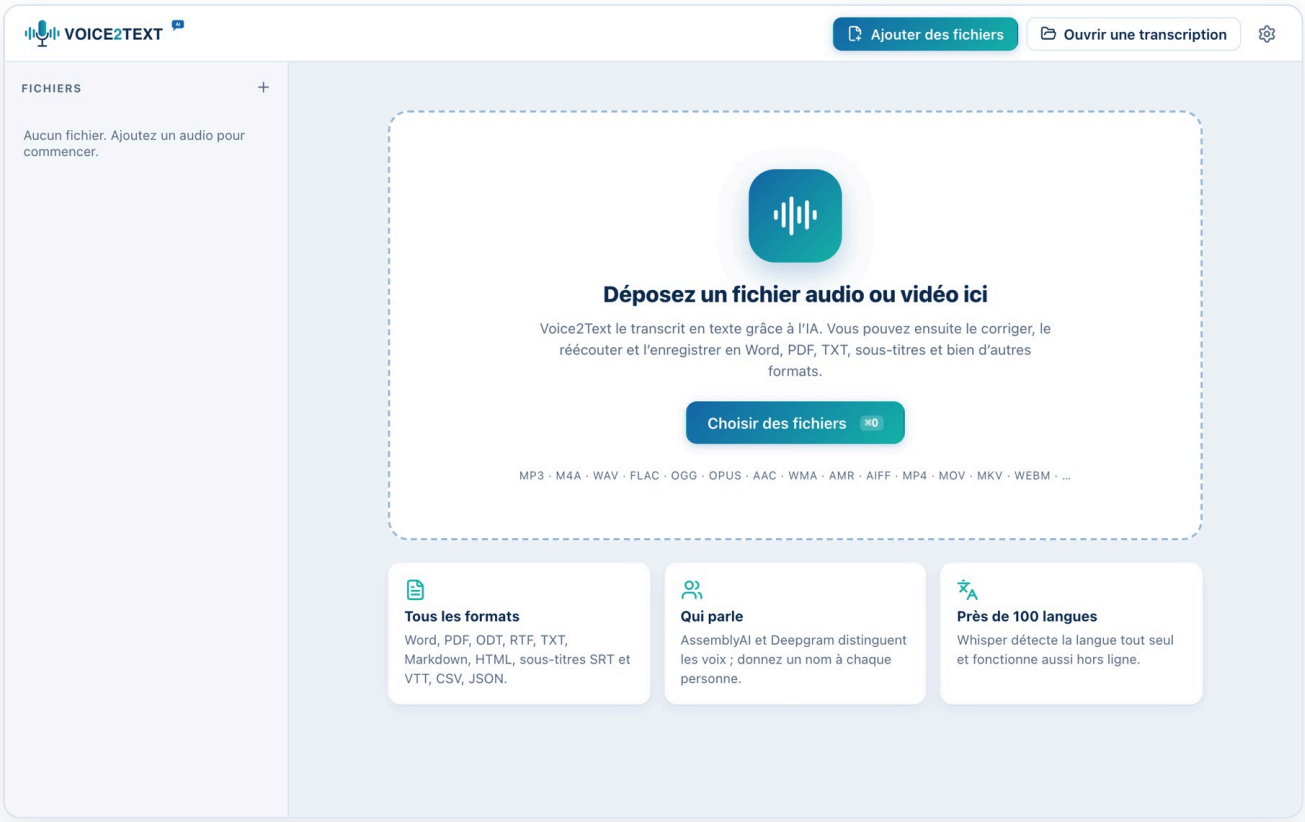


Figure 1. L'écran d'accueil : déposez un fichier dans le cadre ou utilisez « Choisir des fichiers ».

- **Barre du haut** : **Ajouter des fichiers**, **Ouvrir une transcription** (un projet `.v2t` enregistré) et l'engrenage des **Réglages**.
- **Liste des fichiers** à gauche : chaque fichier affiche son état (prêt, en attente, transcription, transcrit, erreur) et sa progression. Le point après le nom signale des modifications non enregistrées.
- **Zone centrale** : les réglages de la transcription, la progression ou le texte à corriger.

Ajouter des fichiers

Faites glisser un ou plusieurs fichiers audio ou vidéo sur la fenêtre, ou utilisez **Ajouter des fichiers** (`⌘O` / `Ctrl+O`). Ajoutez-en autant que vous voulez : ils sont transcrits l'un après l'autre. Déposer un fichier `.v2t` rouvre une transcription enregistrée.

04 Choisir le moteur et la langue

> Avant de transcrire, choisissez qui fera le travail : Whisper sur votre ordinateur ou un service en ligne.

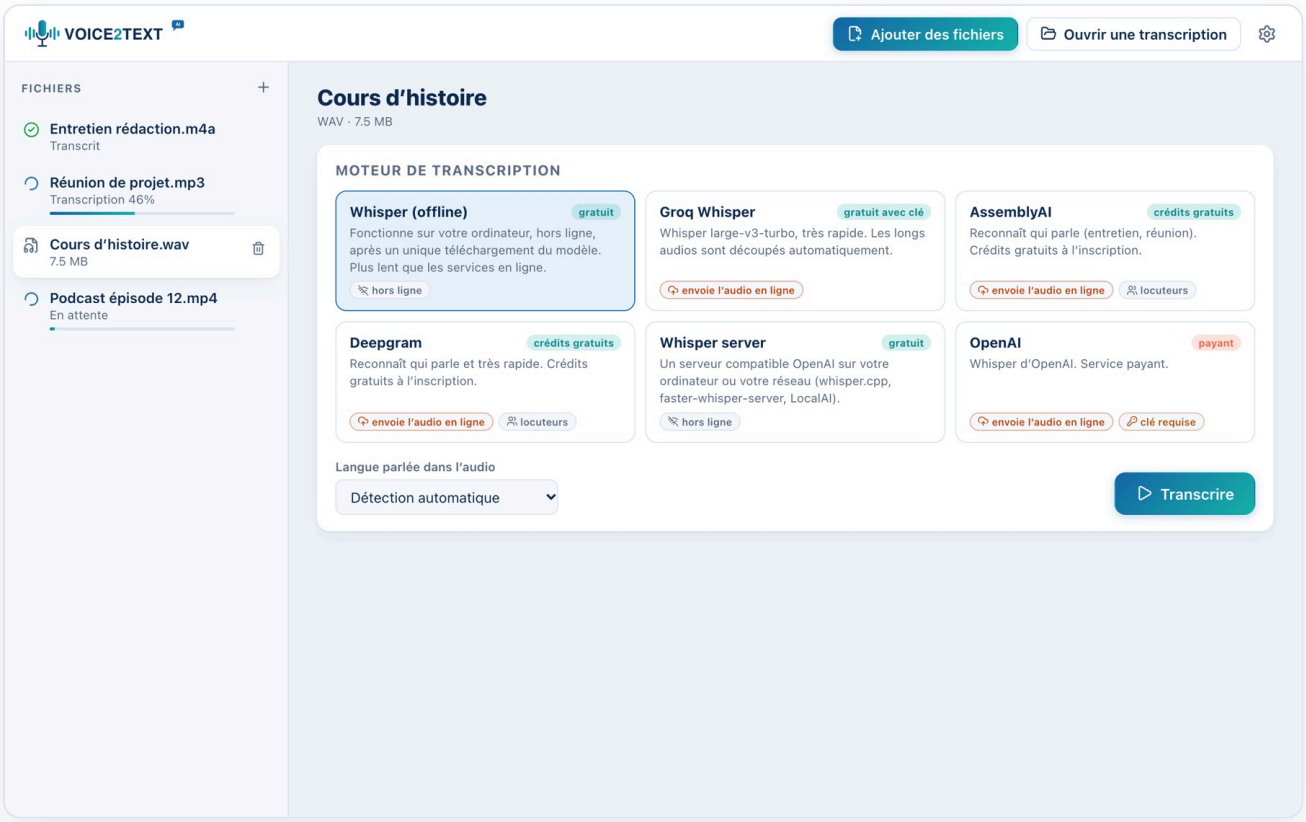


Figure 2. Choix du moteur et de la langue pour le fichier sélectionné.

Moteur	Coût	Caractéristiques
Whisper (hors ligne)	gratuit	Fonctionne sans Internet ; le modèle se télécharge une fois (45 à 600 Mo). Plus lent que les services en ligne.
Groq Whisper	gratuit avec clé	Très rapide ; les longs audios sont découpés automatiquement.
AssemblyAI	crédits gratuits	Reconnaît qui parle. Idéal pour les entretiens et les réunions.
Deepgram	crédits gratuits	Reconnaît qui parle et très rapide.
Whisper server	gratuit	Un serveur Whisper sur votre ordinateur ou votre réseau (whisper.cpp, faster-whisper-server, LocalAI).
OpenAI	payant	Whisper d'OpenAI.

La langue de l'audio

Avec **Détection automatique**, Voice2Text reconnaît seul la langue parlée. Si vous la connaissez, choisissez-la : le résultat est plus précis, surtout dans les premières secondes.

Les clés des services en ligne

Groq, AssemblyAI, Deepgram et OpenAI demandent une clé personnelle : dans **Réglages** → **Moteurs**, vous trouvez un champ pour chacun et le lien « obtenir une clé ». L'inscription à Groq est gratuite ; AssemblyAI et Deepgram offrent des crédits à l'inscription.

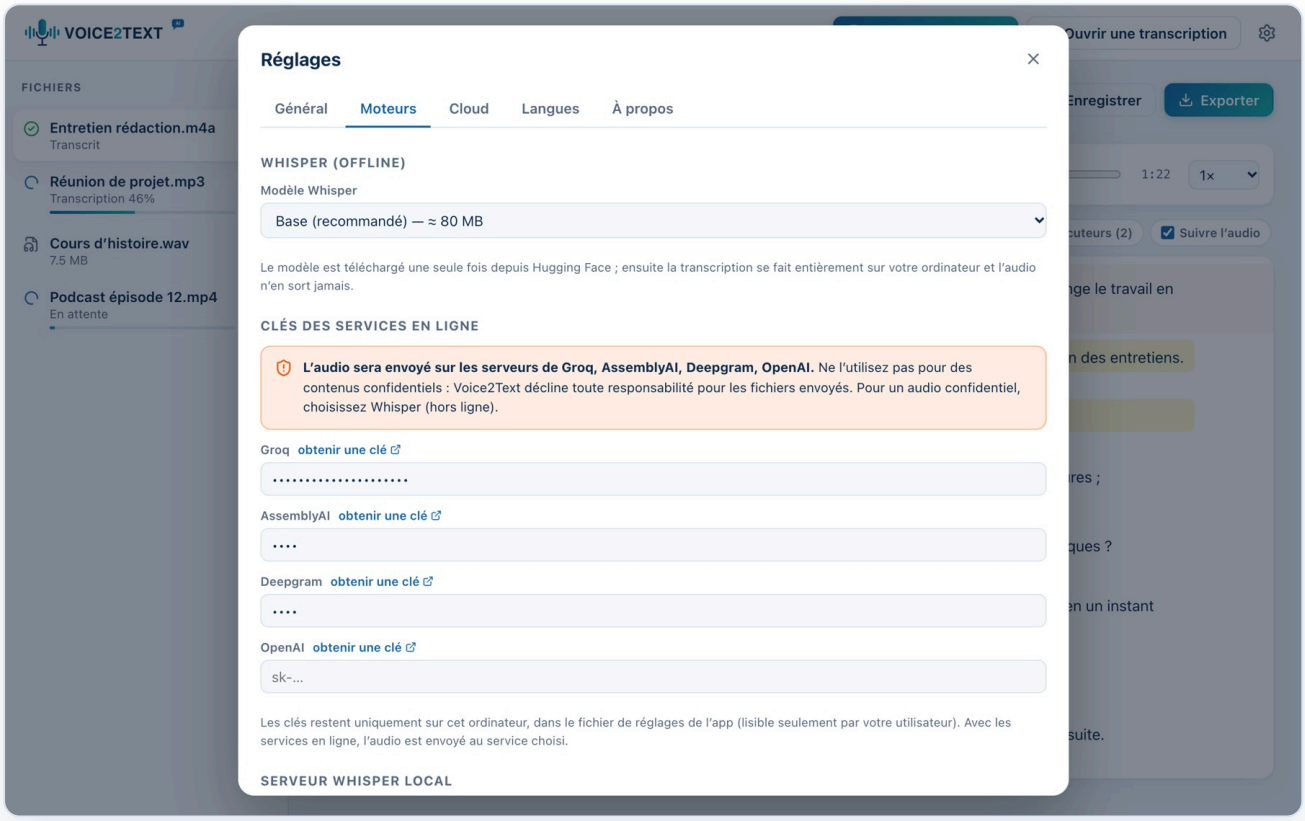


Figure 3. Réglages → Moteurs : modèle Whisper et clés des services.

Appuyez sur **Transcrire** pour commencer. Si plusieurs fichiers attendent, **Tout transcrire** apparaît aussi dans la barre latérale.

05 Services en ligne : votre audio quitte l’ordinateur

> Lorsque vous utilisez un service d’IA en ligne, le fichier audio est envoyé sur Internet. Voice2Text vous le dit clairement et demande votre confirmation.

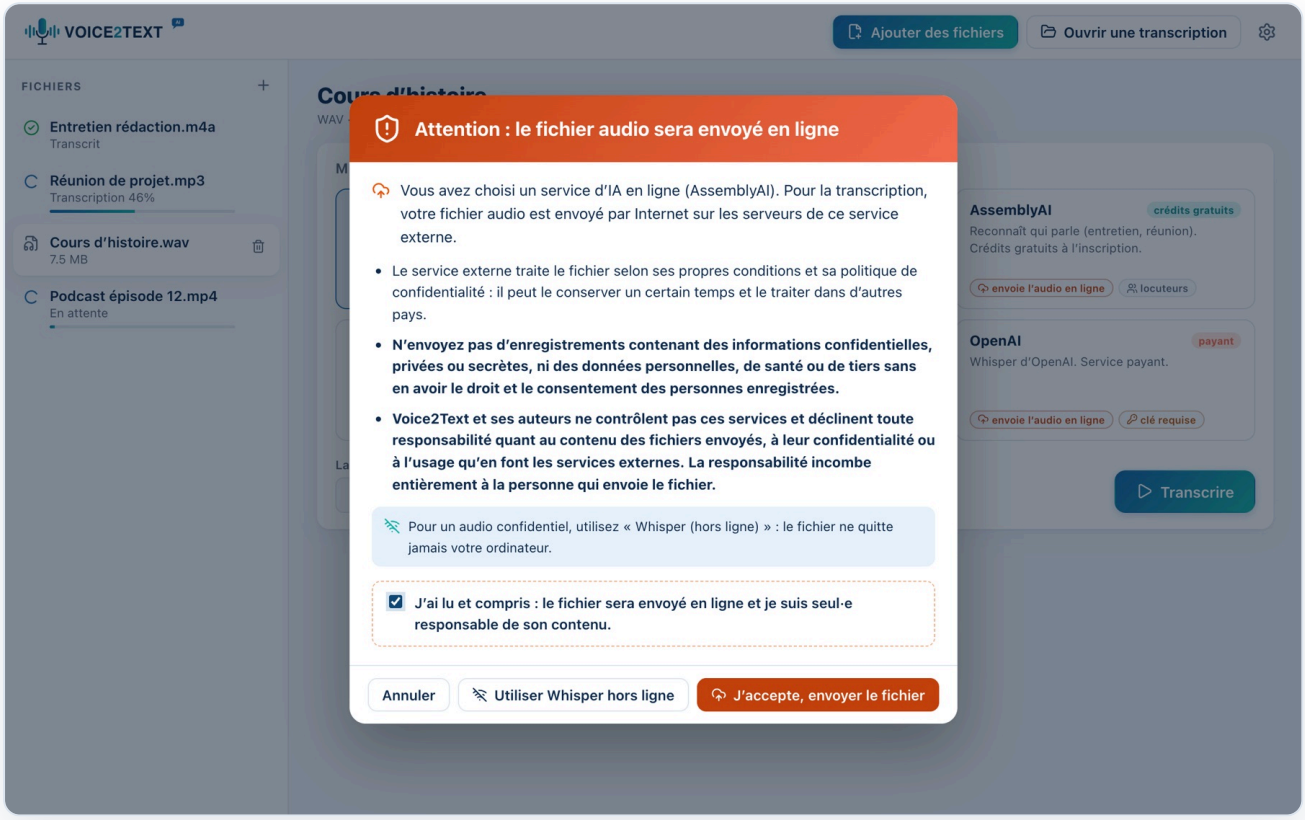


Figure 4. L’avertissement obligatoire avant le premier envoi à un service en ligne.

La première fois que vous utilisez Groq, AssemblyAI, Deepgram ou OpenAI, cet avertissement apparaît. Pour continuer, vous devez cocher **« J’ai lu et compris »** ; vous pouvez aussi passer directement à **Whisper hors ligne**.

Le fichier audio est envoyé sur les serveurs du service externe, qui le traite selon ses propres conditions et sa politique de confidentialité : il peut le conserver un certain temps et le traiter dans d’autres pays.

N’envoyez pas d’enregistrements contenant des informations confidentielles, privées ou secrètes, ni des données personnelles, de santé ou de tiers sans en avoir le droit et le consentement des personnes enregistrées.

Voice2Text et ses auteurs déclinent toute responsabilité quant au contenu des fichiers envoyés, à leur confidentialité ou à l’usage qu’en font les services externes. La responsabilité incombe entièrement à la personne qui envoie le fichier.

L’acceptation vaut pour un seul service et est enregistrée avec la date. Tant qu’un service en ligne est sélectionné, un encadré orange vous le rappelle. Dans **Réglages** → **À propos**, vous pouvez relire l’avertissement et le faire réapparaître.

> **Audio confidentiel ?** Utilisez **Whisper (hors ligne)** : la transcription se fait entièrement sur votre ordinateur.

06 Pendant la transcription

> Voice2Text montre à tout moment où en est le travail.

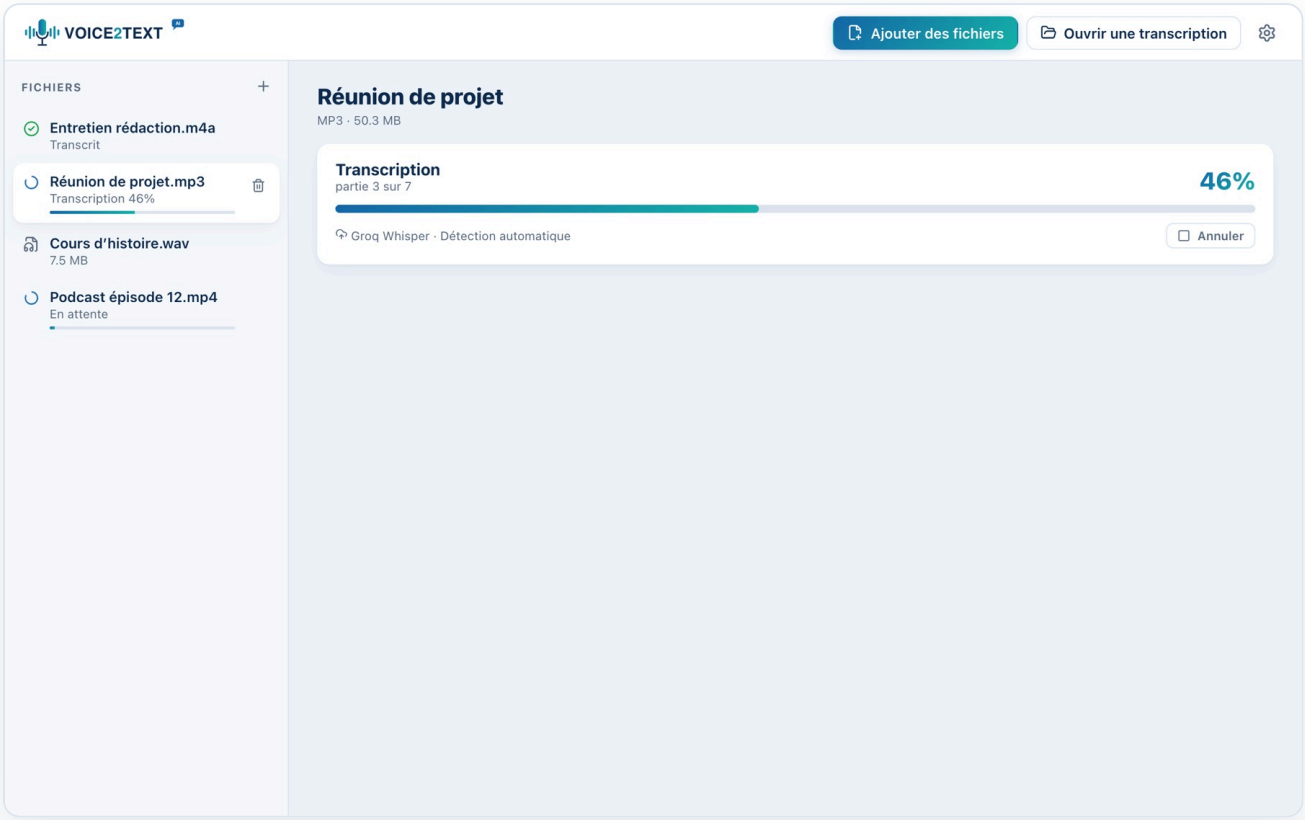


Figure 5. Une transcription en cours avec Groq : progression et partie en traitement.

- **Lecture de l'audio** : le fichier est décodé sur l'ordinateur.
- **Préparation du modèle** (Whisper hors ligne uniquement, la première fois) : le modèle choisi est téléchargé et conservé.
- **Envoi de l'audio** et **Transcription** : pour les longs fichiers, vous voyez la partie en cours (« partie 3 sur 7 »).
- **En attente du service** : avec les forfaits gratuits, si la limite de requêtes est dépassée, Voice2Text attend et réessaie tout seul.

Vous pouvez annuler à tout moment avec **Annuler**, continuer à travailler sur d'autres fichiers et laisser la transcription tourner : une notification vous prévient quand le texte est prêt.

07 Corriger la transcription

> Le texte est découpé en phrases, chacune avec son horodatage et, si elle est reconnue, la personne qui parle.

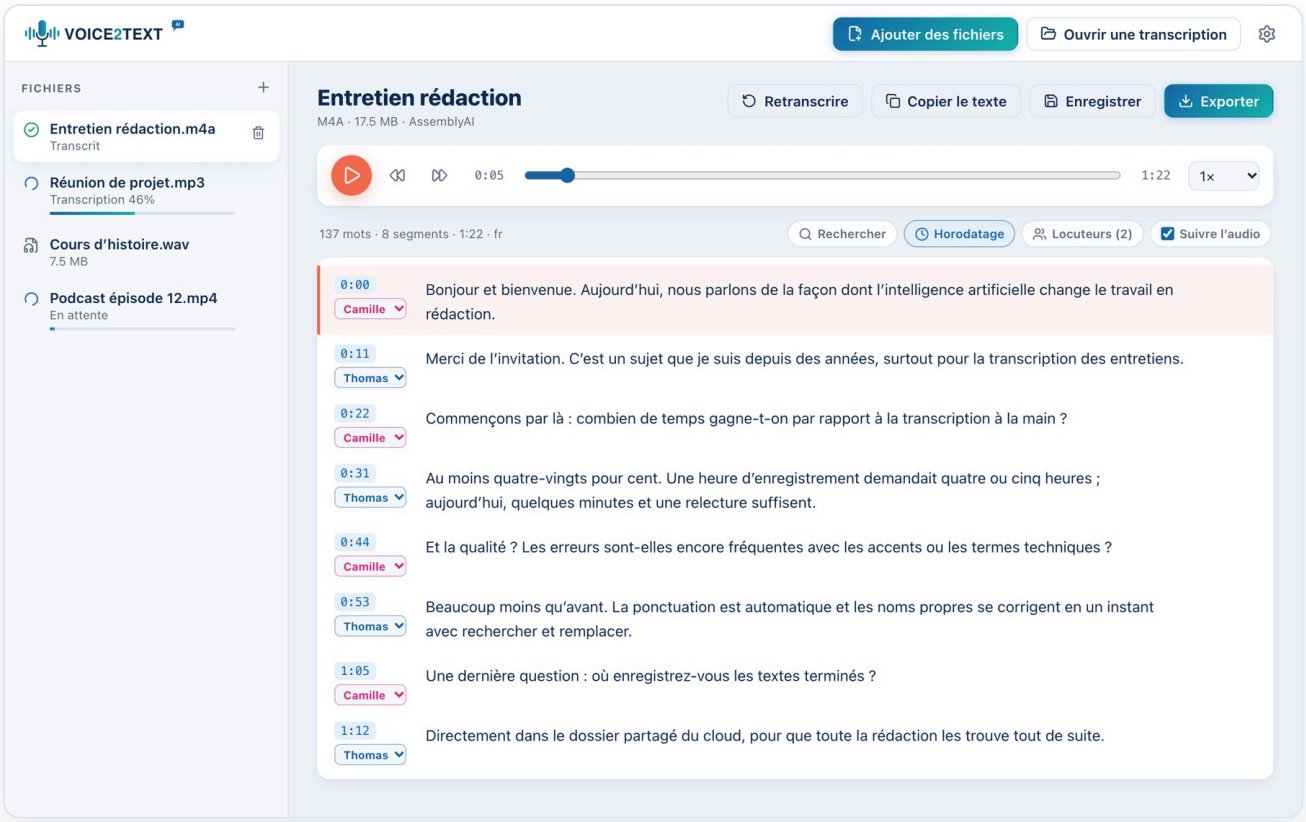


Figure 6. L'éditeur : lecteur audio en haut et phrases modifiables ; la phrase en cours de lecture est mise en évidence.

Écouter

Le lecteur en haut lit l'audio original : **Espace** lance et met en pause (quand vous n'écrivez pas), les boutons sautent de 5 secondes et vous pouvez choisir la vitesse de 0,75× à 2×. Cliquez sur l'heure d'une phrase pour l'écouter ; avec **Suivre l'audio**, la phrase en lecture reste toujours visible.

Modifier

- Cliquez sur une phrase et écrivez : les modifications sont immédiates.
- **Entrée** coupe la phrase à l'endroit du curseur ; **Maj+Entrée** ajoute un retour à la ligne dans la même phrase.
- **Retour arrière** au début d'une phrase la fusionne avec la précédente ; le bouton **Fusionner** la joint à la suivante.
- La corbeille supprime une phrase (par exemple un bruit reconnu comme un mot).

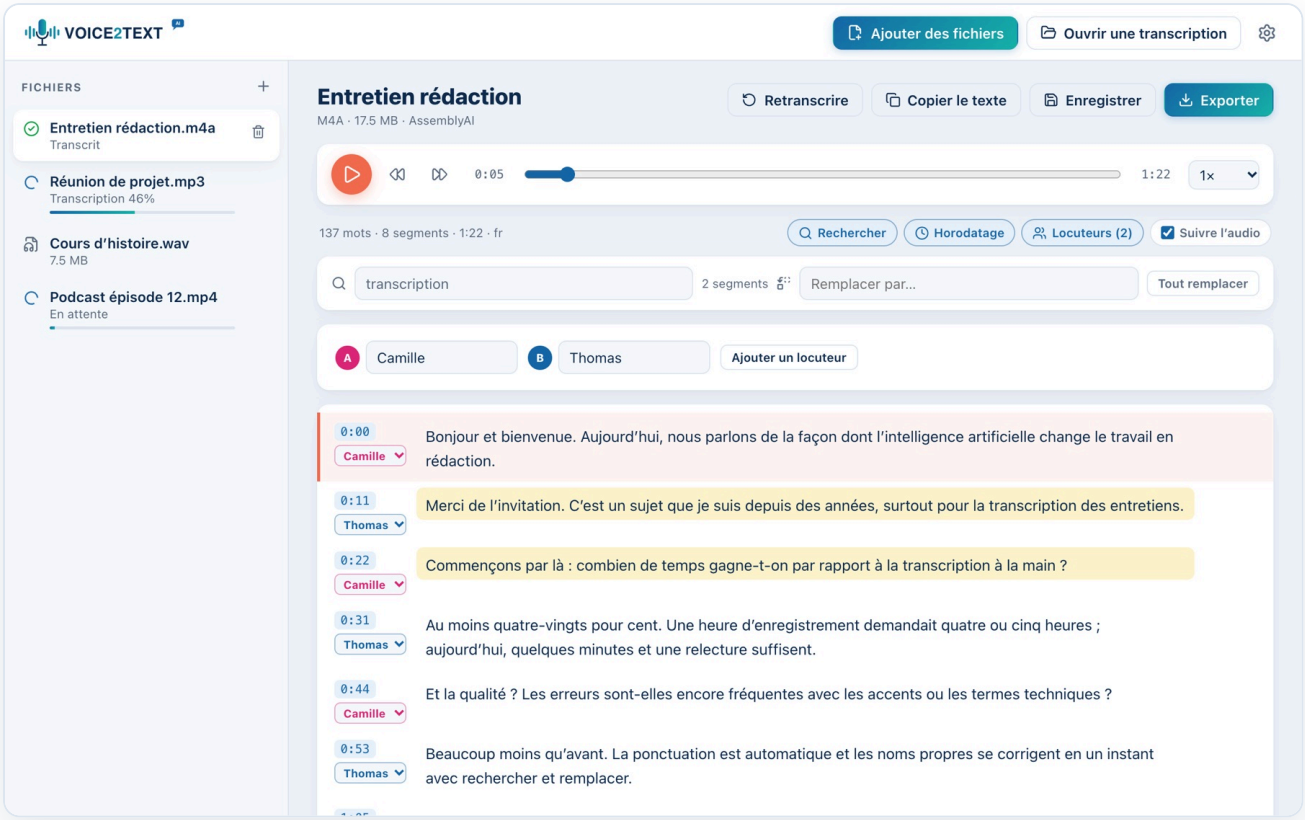


Figure 7. Rechercher et remplacer, et les noms des locuteurs.

Rechercher et remplacer

Rechercher (⌘F / Ctrl+F) met en évidence les phrases qui contiennent le mot ; **Tout remplacer** corrige d'un coup un nom mal écrit dans toute la transcription.

Locuteurs

Avec AssemblyAI et Deepgram, chaque phrase porte une étiquette (Locuteur A, Locuteur B…). Dans le panneau **Locuteurs**, saisissez le vrai nom : il apparaît partout, y compris dans les fichiers exportés. Vous pouvez changer le locuteur d'une phrase depuis le menu à côté de l'heure, en ajouter d'autres ou les attribuer à la main même quand le moteur ne les reconnaît pas.

08 Enregistrer et exporter

> Enregistrez le projet pour reprendre les corrections plus tard et exportez le texte dans le format dont vous avez besoin.

Le projet .v2t

Enregistrer (⌘S / Ctrl+S) crée un fichier `.v2t` avec la transcription, les corrections et les noms des locuteurs. Rouvrez-le avec **Ouvrir une transcription** (⇧⌘O / Ctrl+Maj+O) ou en le déposant sur la fenêtre. ⇧⌘S enregistre sous un autre nom.

Exporter

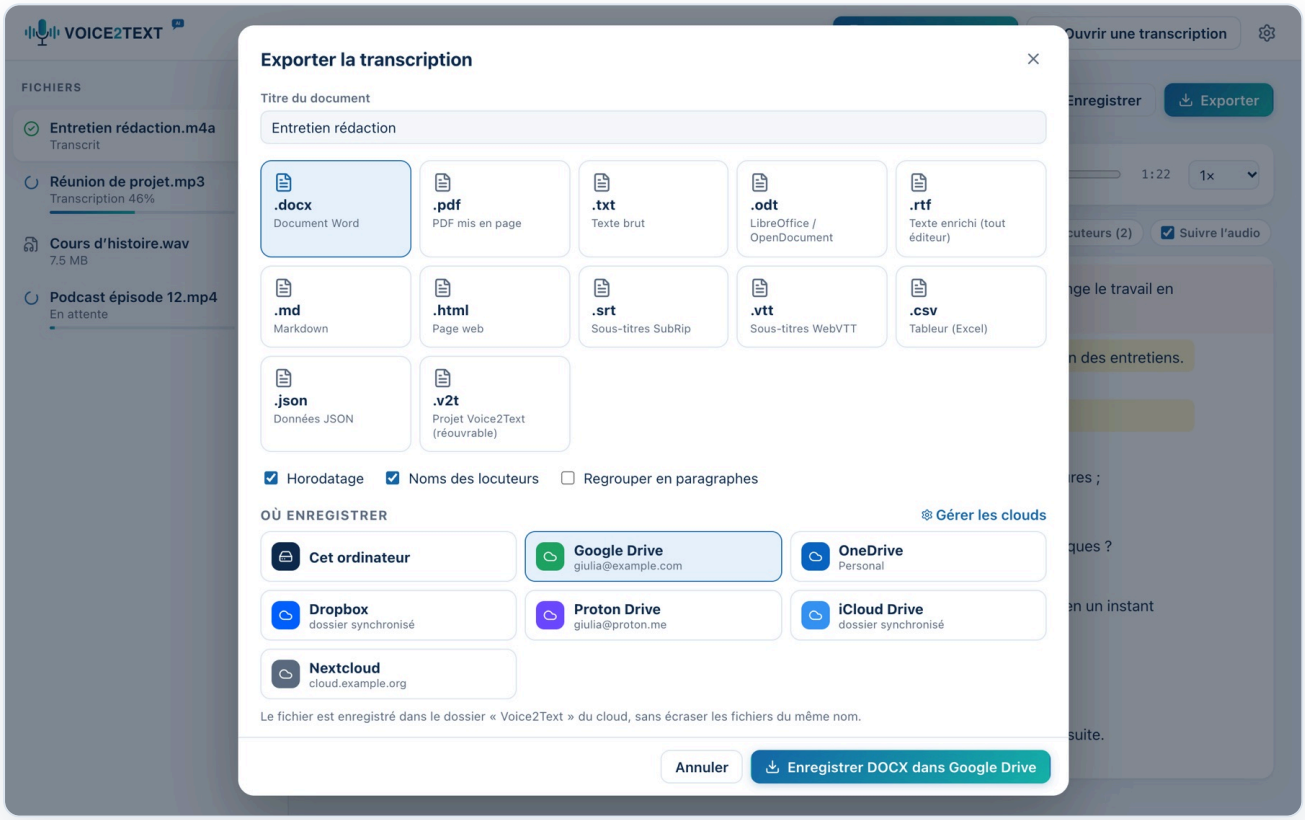


Figure 8. La fenêtre Exporter : format, options et destination.

Appuyez sur **Exporter** (⌘E / Ctrl+E), saisissez le titre du document et choisissez le format :

Format	Idéal pour
.docx Word · **.odt** LibreOffice · **.rtf**	documents à modifier, mis en page avec le titre et les données de l’enregistrement
.pdf	documents à imprimer ou envoyer ; créé via la fenêtre d’impression du système (« Enregistrer au format PDF »)
.txt · **.md** · **.html**	texte brut, Markdown, page web
.srt · **.vtt**	sous-titres vidéo (YouTube, Premiere, DaVinci, VLC…)
.csv · **.json**	tableurs (Excel, Numbers) et données pour d’autres logiciels
.v2t	le projet réouvrable, à enregistrer aussi dans le cloud

Les options permettent d’inclure l’**horodatage**, les **noms des locuteurs** et de **regrouper les phrases en paragraphes**. L’arabe et les autres langues écrites de droite à gauche sont mis en page correctement.

09 Enregistrer dans le cloud

> Voice2Text enregistre les fichiers directement dans votre cloud grâce aux dossiers que les applications des services créent sur l'ordinateur. Aucun compte à connecter, aucun mot de passe à saisir.

Fonctionnement

Quand vous installez l'application de Google Drive, OneDrive, Dropbox, Box, **Proton Drive**, iCloud Drive, pCloud, MEGA, Nextcloud ou Sync.com, un dossier synchronisé apparaît sur l'ordinateur. Voice2Text le trouve tout seul : dans la fenêtre **Exporter**, sous **Où enregistrer**, choisissez le cloud et appuyez sur le bouton. Le fichier est écrit dans le sous-dossier `Voice2Text` et l'application du service l'envoie aussitôt en ligne.

- Les fichiers ne sont jamais écrasés : si un fichier du même nom existe, le nouveau devient « nom (2) ».
- Vous pouvez enregistrer n'importe quel format dans le cloud, y compris le projet `.v2t`.
- Le PDF passe par la fenêtre d'impression : choisissez-y « Enregistrer au format PDF » et le dossier du cloud.

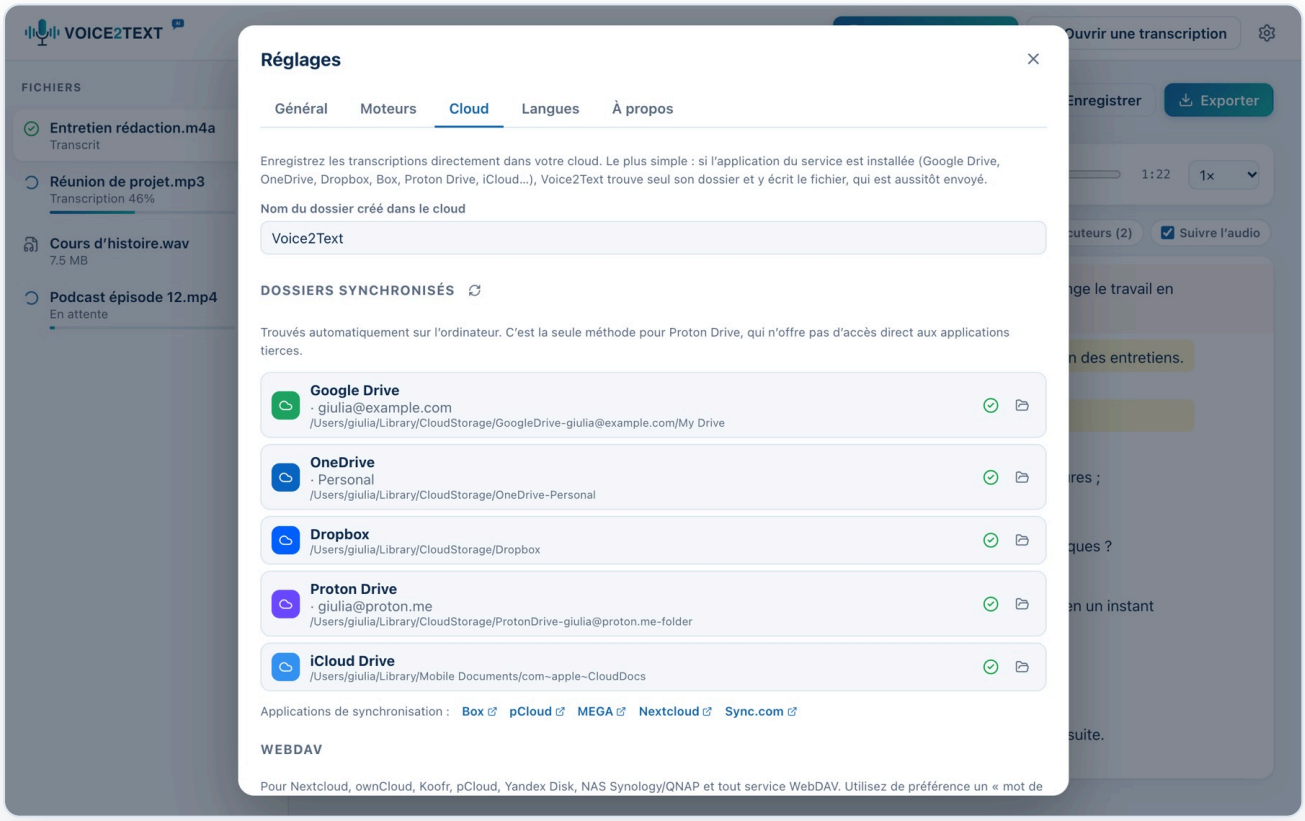


Figure 9. Réglages → Cloud : les dossiers trouvés sur l'ordinateur et la connexion WebDAV.

Réglages → Cloud

- **Nom du dossier** : le sous-dossier créé dans le cloud (par défaut `Voice2Text`).
- **Dossiers synchronisés** : la liste des clouds trouvés ; l'icône de dossier l'ouvre. S'il manque un service, les liens en bas mènent aux applications à installer.

- **Ajouter un autre dossier** : pour d’autres clouds ou un disque réseau (NAS).
- **WebDAV** : pour Nextcloud, ownCloud, Koofr, pCloud, Yandex Disk et NAS Synology/QNAP sans application installée. Choisissez le service, saisissez l’adresse, l’utilisateur et un mot de passe d’application, puis appuyez sur **Tester la connexion**.

> **Proton Drive** ne permet pas aux autres applications d’accéder directement au compte : c’est pourquoi Voice2Text utilise son dossier synchronisé, qui conserve le même chiffrement de bout en bout.

10 Réglages et langues

L'engrenage en haut à droite (`⌘,` / `Ctrl+,`) ouvre les réglages.

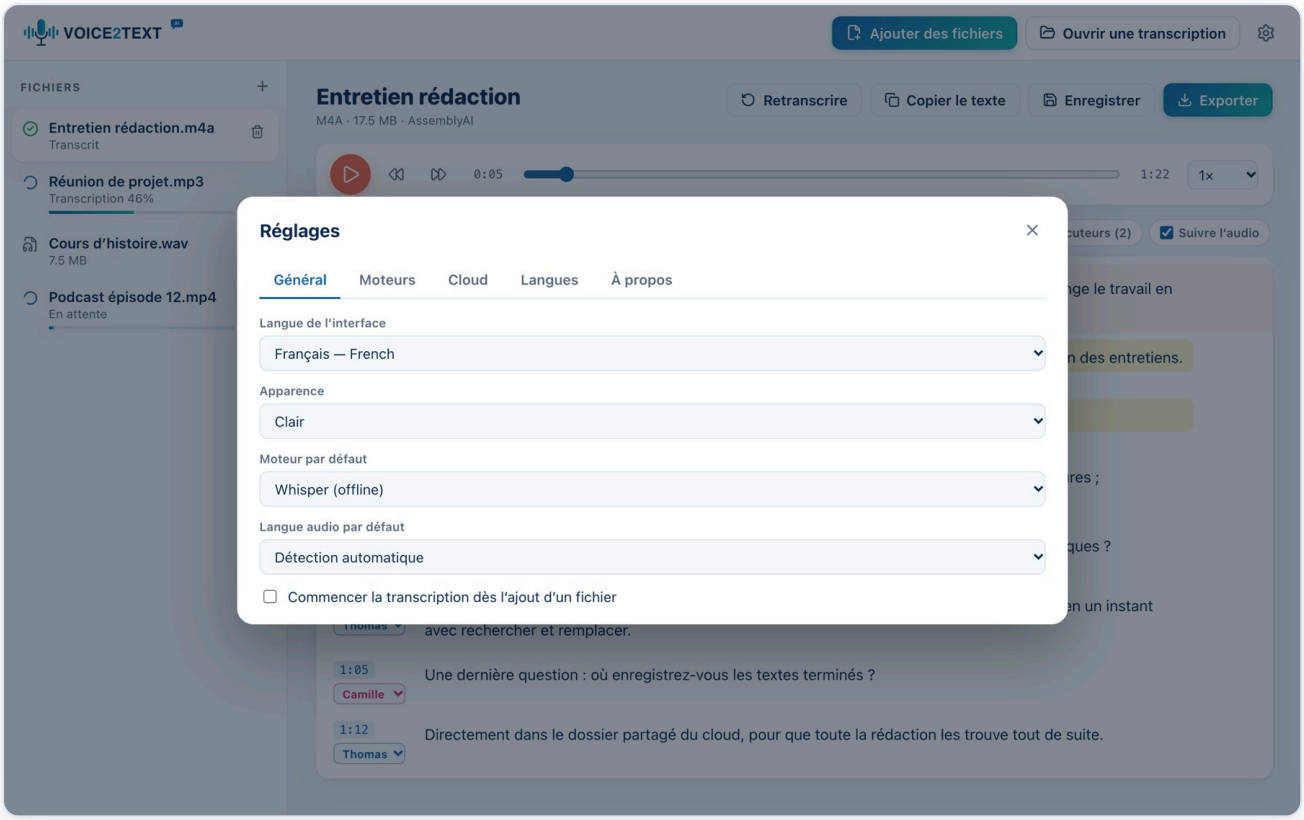


Figure 10. Réglages → Général.

- **Général** : langue de l'interface, apparence claire, sombre ou automatique, moteur et langue audio par défaut, démarrage automatique de la transcription.
- **Moteurs** : modèle de Whisper hors ligne (minimal, base, petit, grand turbo), espace occupé par les modèles, clés des services, serveur Whisper local, présence de ffmpeg.
- **Cloud** : dossiers synchronisés et WebDAV (chapitre précédent).
- **Langues** : les fichiers de traduction de l'interface.
- **À propos** : version, licence, avertissement sur les services en ligne.

Les langues sont des fichiers XML

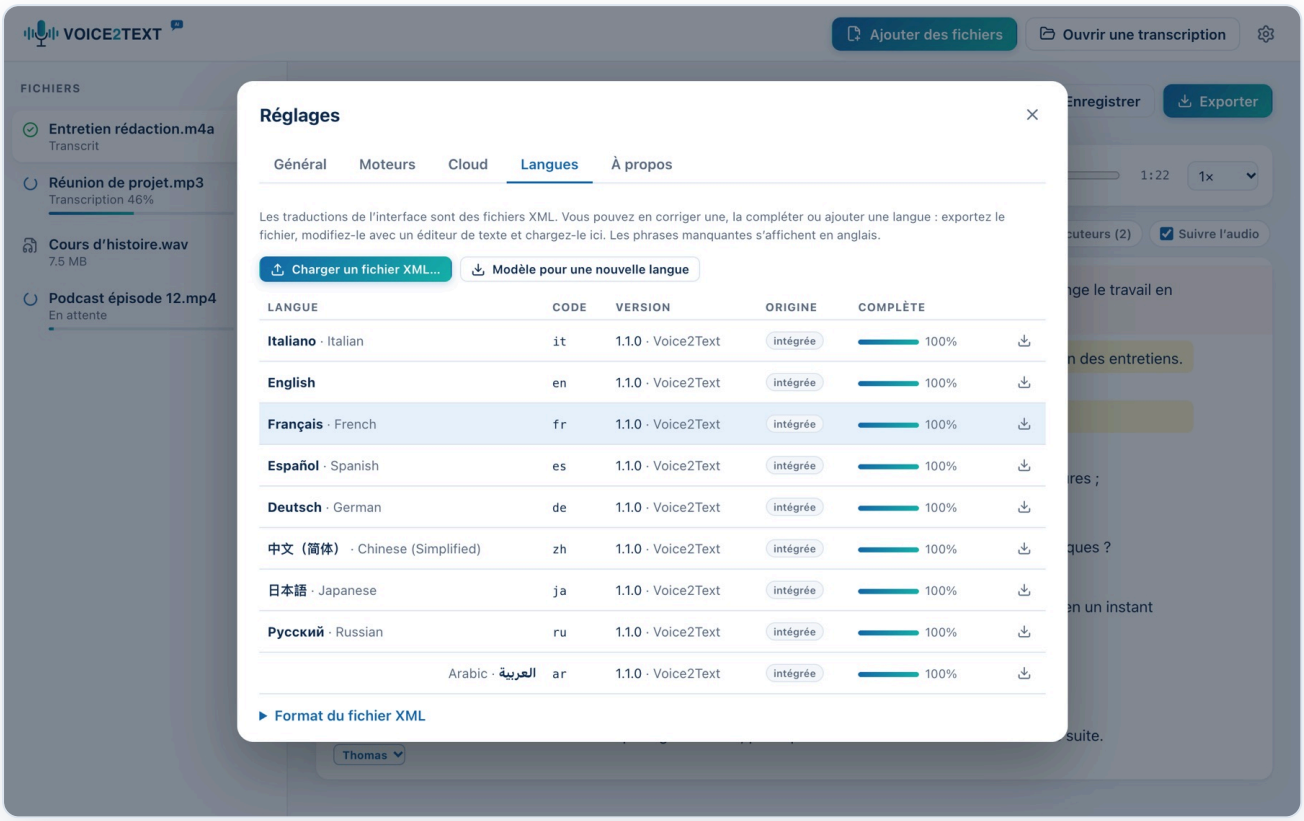


Figure 11. Réglages → Langues : langues disponibles, taux de traduction et outils pour les fichiers XML.

Chaque langue de l'interface est un fichier XML. Pour corriger une traduction ou en ajouter une :

1. Exportez le fichier de la langue (flèche sur sa ligne) ou téléchargez le **Modèle pour une nouvelle langue**.
2. Ouvrez-le avec un éditeur de texte et traduisez uniquement le texte entre les balises, sans modifier les attributs `key` ni les espaces réservés entre accolades comme `{n}`.
3. Appuyez sur **Charger un fichier XML…** : la langue apparaît aussitôt dans la liste et dans le menu des langues.

```
<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>
```

Un fichier ayant le même code qu'une langue intégrée la remplace (pour les corrections) ; vous pouvez toujours revenir à l'original. Les phrases manquantes s'affichent en anglais ; la colonne **Complète** indique combien sont traduites. Pour les langues de droite à gauche, utilisez `dir="rtl"`.

11 Raccourcis clavier

> Sous Windows et Linux, utilisez **Ctrl** au lieu de ⌘.

Touches	Action
⌘O	Ajouter des fichiers audio ou vidéo
⇧⌘O	Ouvrir une transcription enregistrée (.v2t)
⌘S · ⇧⌘S	Enregistrer · Enregistrer sous
⌘E	Exporter
⌘F	Rechercher et remplacer
⌘,	Réglages
Espace	Lecture / pause (hors du texte)
Entrée · Maj+Entrée	Couper la phrase · Retour à la ligne dans la phrase
Retour arrière en début de phrase	Fusionner avec la phrase précédente
Échap	Quitter le texte · fermer la fenêtre ouverte

12 Questions fréquentes

> Les réponses aux questions les plus courantes.

Le format de mon fichier n'est pas lu.

Installez ffmpeg (chapitre Installation) ou utilisez AssemblyAI ou Deepgram, qui acceptent tous les formats car ils reçoivent le fichier original.

Whisper hors ligne est lent.

Choisissez le modèle **Base** ou **Minimal** dans Réglages → Moteurs. « Grand turbo » est le plus précis mais prend du temps sans carte graphique. Pour les longs fichiers, Groq est bien plus rapide.

Le texte est dans une autre langue ou traduit.

Indiquez la langue de l'audio au lieu de la détection automatique et utilisez **Retranscrire**.

« Clé de service non valide ».

Vérifiez la clé dans Réglages → Moteurs : copiez-la de nouveau depuis le site du service, sans espaces.

Mon cloud n'apparaît pas.

Vérifiez que l'application du service est installée et synchronisée, puis appuyez sur **Rechercher à nouveau** dans Réglages → Cloud. Sinon, ajoutez le dossier à la main.

Le lecteur indique « audio indisponible ».

Le fichier audio original a été déplacé ou renommé après l'enregistrement du projet. Le texte reste disponible ; pour réécouter, remettez le fichier à sa place.

Où sont mes clés et mes réglages ?

Uniquement sur votre ordinateur, dans le fichier de réglages de l'application, lisible seulement par votre utilisateur. Ils ne sont jamais envoyés à Voice2Text.



VOICE2TEXT

De la voix au texte, dans toutes les langues.

Voice2Text est un logiciel d'overminder. Licence PolyForm Strict 1.0.0 : usage non commercial autorisé ; redistribution, modifications et œuvres dérivées non autorisées.

© 2026 overminder · Version 1.1