

La transcription générée automatiquement

Lingua e traduzione – lingua francese – corso avanzato (RISE-SA), a.a. 2023-2024

Prof.ssa Alida Maria Silletti

SOURCES: <https://www.coe.int/fr/web/artificial-intelligence/history-of-ai> ;
<https://www.lebigdata.fr/intelligence-artificielle> ; <https://www.techno-science.net/glossaire-definition/Reconnaissance-vocale.html> ;
<https://www.journaldunet.fr/web-tech/guide-de-l-intelligence-artificielle/1501849-reconnaissance-vocale/> ; <https://shs.hal.science/halshs-02917916/document>

Les origines de la transcription automatique

- Le premier système relevant de la « reconnaissance vocale » date de 1952 comme un outil de répression au service de Staline
- Années 1980-1990 : depuis la démocratisation des outils d'enregistrement, la pratique de la transcription de l'intégralité de la parole enregistrée devient presque la norme, mais elle demande beaucoup de temps et elle donne lieu à des résultats souvent décevants - pour 1 heure d'enregistrement, la durée de transcription peut aller de 4 à 6 h, voire 30 h en fonction de la personne qui transcrit, des caractéristiques de l'enregistrement et de la transcription
- Depuis 2010 : des modules d'intelligence artificielle sont intégrés aux algorithmes de reconnaissance automatique de la parole – les algorithmes progressent rapidement

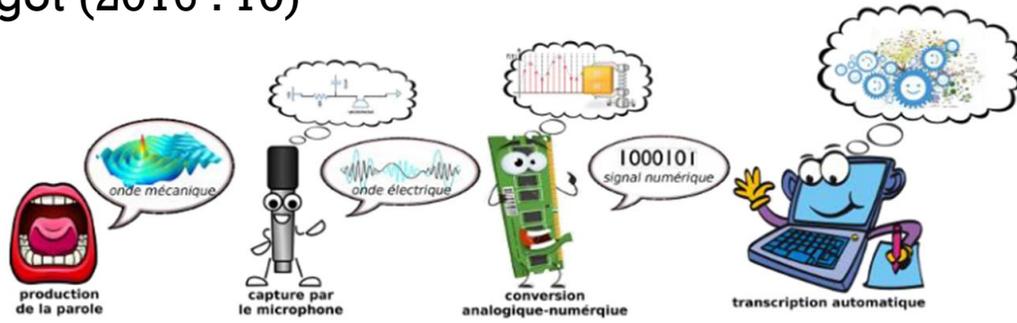


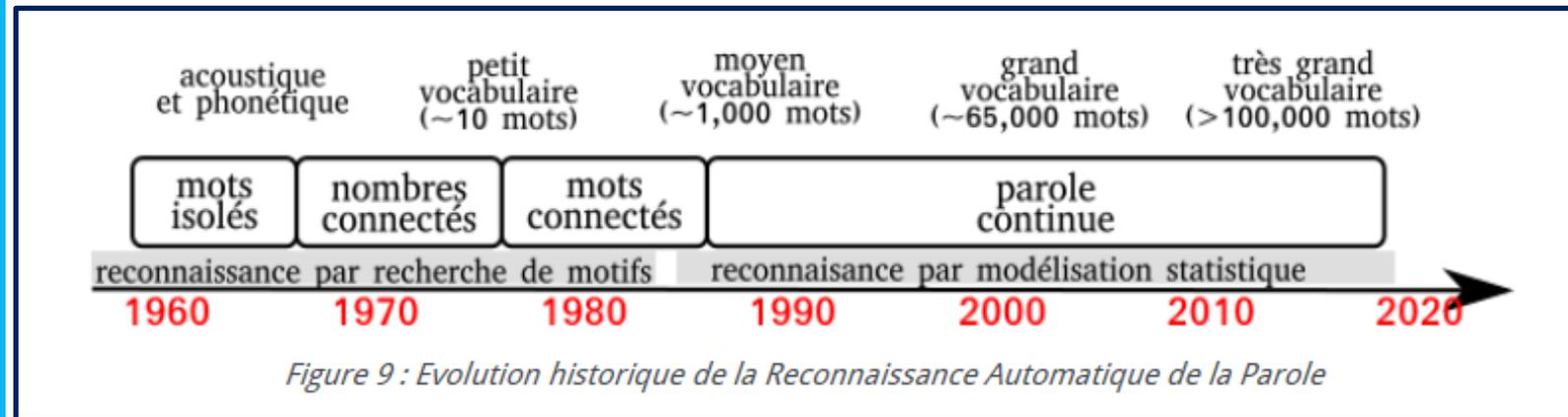
Figure 1 : Chaîne d'acquisition et de traitement du signal de parole

La reconnaissance vocale ou reconnaissance automatique de la parole (RAP)

- Une technique informatique qui permet d'analyser un mot ou une phrase captée au moyen d'un microphone pour la transcrire sous la forme d'un texte exploitable par une machine
- Elle est possible par le biais de la voix et elle passe par diverses opérations :
 - 1 – les fréquences sonores de la voix sont captées par un microphone ;
 - 2 – elles sont traduites en texte exploitable par la machine – le *speech-to-text* ;
 - 3 – les fichiers sonores sont analysés par les technologies de l'intelligence artificielle – le *deep learning* et le *natural language understanding*
 - 4 - une traduction optimale de la voix en des données utilisables par la machine

Un processus de plus en plus sophistiqué et perfectionné

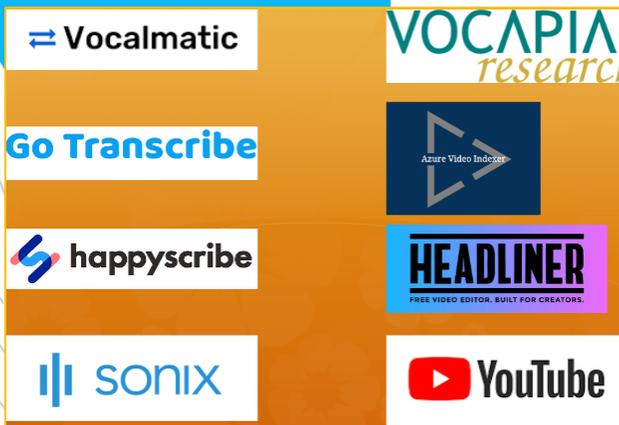
Un modèle qui s'entraîne à partir d'une grande quantité de données relevant d'exemples vocaux étiquetés



Bigot (2016 : 10)

La transcription automatique : un rêve enfin accessible ? Analyse et comparaison d'outils pour les SHS. Nouvelle méthodologie et résultats
(<https://shs.hal.science/halshs-02917916/document>)

- Un rapport de recherche publié à septembre 2020 par des spécialistes sur la transcription automatique en langue française
- Un travail de comparaison entre 8 outils de transcription automatique : Go Transcribe, Happy Scribe, Headliner, Sonix, Video Indexer, Vocalmatic, Vocapia, YouTube à partir de 4 extraits de fichiers audio en langue française :
 - un texte lu ;
 - un cours magistral enregistré en situation ;
 - un entretien avec deux personnes ;
 - une réunion associative avec de nombreuses personnes



La transcription automatique : un rêve enfin accessible ? Analyse et comparaison d'outils pour les SHS. Nouvelle méthodologie et résultats
(<https://shs.hal.science/halshs-02917916/document>)

■ Objectifs :

- évaluer les fonctionnalités des plateformes en termes de sécurité et confidentialité des données, tarification, interopérabilité, simplicité d'utilisation, outil d'édition ;
- évaluer les transcriptions obtenues pour un classement et une compréhension des erreurs générées par les logiciels et pour effectuer des estimations du potentiel de gain de temps de transcription par fichier et par logiciel

■ Résultats :

- la qualité de la transcription dépend du type de fichier soumis en entrée (des discours planifiés VS la parole spontanée) ;
- tous les logiciels échouent à retranscrire le fichier le plus complexe, la réunion associative, à cause de la présence de nombreux chevauchements, d'extraits inaudibles et de bruits de fond ;
- un temps de réécoute et de correction reste indispensable ;
- le gain de temps final observé peut aller jusqu'à 75 % par rapport à une transcription manuelle

La transcription automatique : un rêve enfin accessible ? Analyse et comparaison d'outils pour les SHS. Nouvelle méthodologie et résultats
(<https://shs.hal.science/halshs-02917916/document>)

- **Donc, un rêve enfin accessible ?**
 - ... en devenir ...
 - la transcription automatique est variable en fonction des besoins et des corpus ;
 - il faut un temps indispensable de réécoute et de correction

La transcription automatique : un rêve enfin accessible ? Analyse et comparaison d'outils pour les SHS. Nouvelle méthodologie et résultats
(<https://shs.hal.science/halshs-02917916/document>)

■ Résultats généraux pour le logiciel de YouTube :

- le seul logiciel 100 % gratuit ;
- ses données sont hébergées par Google ;
- il présente des balises temporelles ;
- la ponctuation y est absente ;
- L'alternance des personnes qui parlent n'est pas respectée ;
- 80-85 % des mots transcrits sont partagés avec le texte de référence

L'exemple du logiciel de sous-titrage automatique de Google

Utiliser les sous-titres automatiques de YouTube (Google)

Utiliser les sous-titres automatiques

Les sous-titres sont un excellent moyen de rendre vos contenus accessibles aux spectateurs. YouTube utilise sa technologie de reconnaissance vocale afin de créer automatiquement des sous-titres pour vos vidéos.

Remarque : Ces sous-titres automatiques étant générés par des algorithmes de machine learning (apprentissage automatique), leur qualité peut varier. Nous encourageons les créateurs à recourir à des services professionnels pour leurs sous-titres. YouTube améliore sans cesse sa technologie de reconnaissance vocale. Toutefois, il peut arriver que les sous-titres automatiques ne reflètent pas correctement le contenu audio en raison d'erreurs de prononciation, d'accents, de l'utilisation de dialectes ou de bruits de fond. Nous vous recommandons de vérifier les sous-titres automatiques et de modifier les phrases qui n'ont pas été correctement transcrites.

<https://support.google.com/youtube/answer/6373554?hl=fr#zippy=%2Cmots-potentiellement-inappropri%C3%A9s-dans-les-sous-titres-automatiques%2Csous-titres-automatiques-pour-les-vid%C3%A9os-diffus%C3%A9es-en-direct>