

# Table des matières

1	Introduction et problématique	2
2	Présentation générale	3
3	Prétraitement : préparation des données	3
4	Utilisation des réseaux de neurones	3
5	Conclusion	3

# Références

[1] Jean-François Jodouin. Présentation des modèles connexionnistes. 1990.

## 1 Introduction et problématique

Aujourd'hui, on dispose de moyens assez efficaces pour transformer automatiquement un texte prononcé dans un fichier audio en sa transcription écrite (par exemple, l'outil de sous-titrage automatique de Youtube). Le problème avec ces outils est qu'ils nécessitent toujours de savoir en quelle langue le texte est prononcé pour pouvoir le transcrire efficacement.

En effet, pour faire de la reconnaissance de la parole (fig. 1), il est essentiel de connaître le langage à traduire au préalable car il faut pouvoir utiliser les éléments propres à celui-ci.

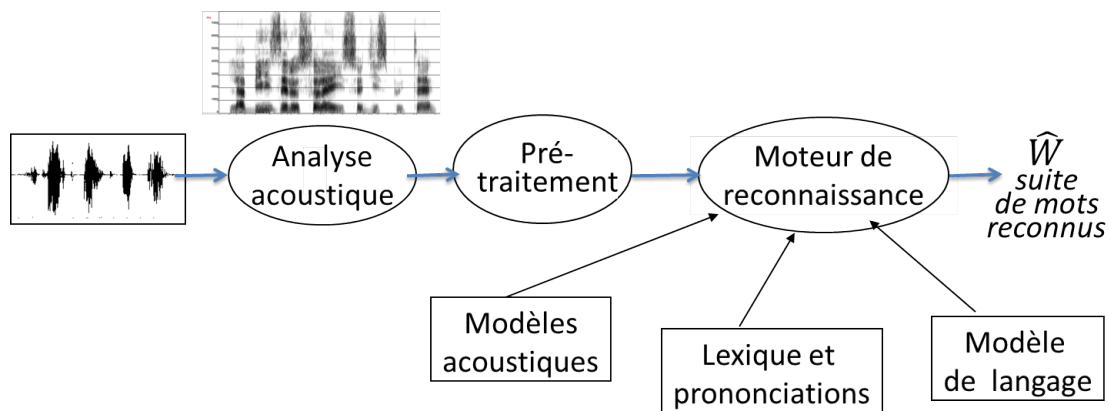


FIGURE 1 – Schéma d'un système classique de reconnaissance de la parole

Ici, les éléments qui vont changer d'un langage à l'autre sont les modèles acoustiques, le lexique et les prononciations et le modèle de langage.[1]

L'objectif de ce projet d'initiation à la recherche est donc de parvenir à détecter automatiquement la langue parlée dans un fichier audio. Cette détection automatique pourrait permettre de directement transcrire la parole prononcée dans le fichier, en utilisant la "bibliothèque" vocale correspondant à la langue parlée.

Pour ce faire, on utilisera des méthodes d'apprentissage automatique (machine learning) basées sur des réseaux de neurones profonds. Dans un premier temps, le but sera de construire un réseau de neurones capable de déterminer, pour un fichier de parole homogène (même langue dans le fichier) donné en entrée, s'il s'agit de parole allemande, anglaise, arabe ou française.

Ensuite, selon les performances du réseau précédent, nous tenterons de créer un système capable de détecter la langue d'un contenu audio de manière dynamique. Ainsi, nous pourrions repérer automatiquement, par exemple, une citation prononcée en anglais dans un flux de parole globalement allemande, pour pouvoir ainsi faciliter la transcription automatique d'un tel texte.

- 2 Présentation générale
- 3 Prétraitement : préparation des données
- 4 Utilisation des réseaux de neurones
- 5 Conclusion