



Module

Applications of Artificial Intelligence

AAI



Reconnaissance automatique de la parole

(Automatic Speech Recognition: ASR)



13:14



Modalités biométriques

**Caractéristiques
biologiques**

**Caractéristiques
comportementales**

**Caractéristiques
morphologiques**

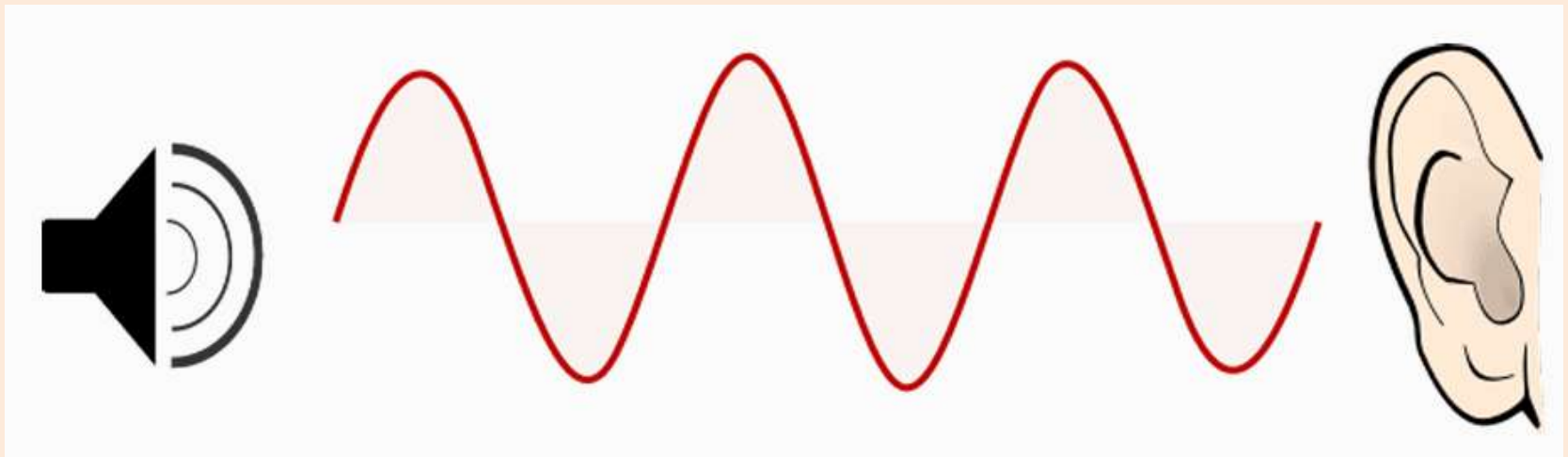


**Reconnaissance automatique
de la parole**



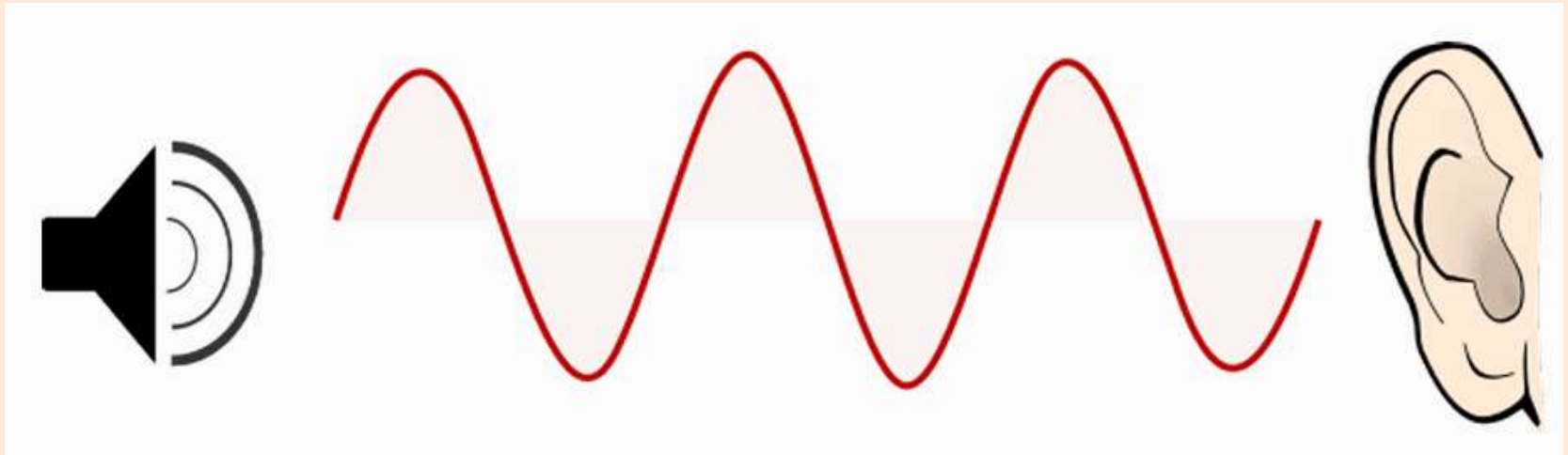
Reconnaissance automatique de la parole

- **Le son** est une **onde mécanique** qui se propage dans tous les milieux physiques (gaz, liquide, solide).



Reconnaissance automatique de la parole

- Le système auditif humain est en mesure de **percevoir** les sons entre 20Hz et 20kHz (dans l'air).



Reconnaissance automatique de la parole

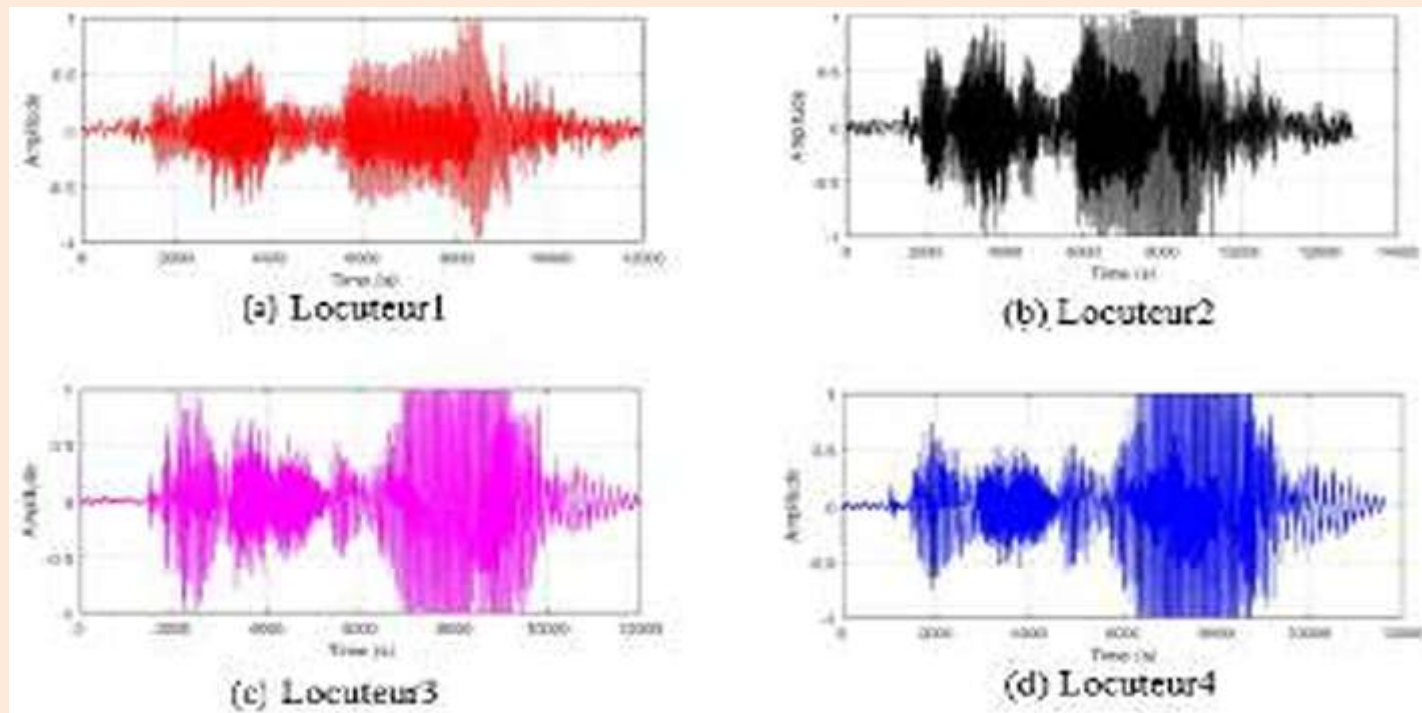
- La parole est: le principal moyen de communication dans toute société humaine.
- La parole permet de se dégager de toute obligation de **contact physique** avec la machine.

Reconnaissance automatique de la parole

La parole est:

- Un processus naturel, variable dans le temps.
- Ce dernier est un vecteur acoustique porteur d'informations d'une grande complexité, variabilité et redondance.

Reconnaissance automatique de la parole



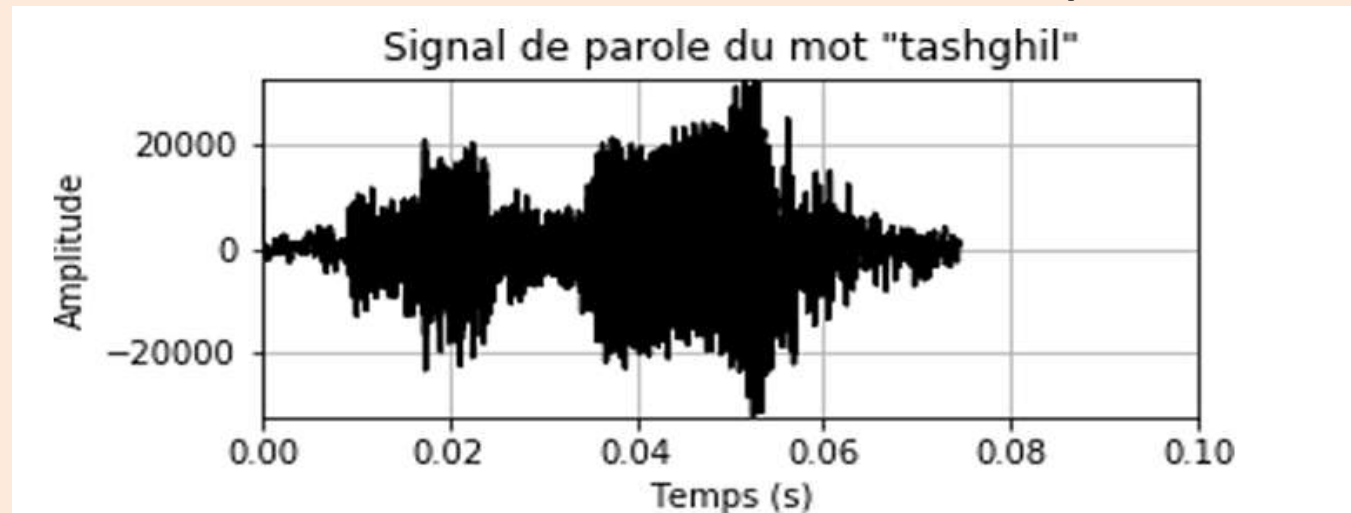
Différents signaux du même mot « allumer » prononcé par différents locuteurs

Reconnaissance automatique de la parole

Elle donne à la machine la capacité de:

- Comprendre.
- Traiter

Des informations fournies oralement par un être humain.



Reconnaissance automatique de la parole

La « **reconnaissance de la parole** » se réfère à plusieurs types de systèmes dont la mission est: de **décoder l'information** portée par le signal vocal.

Reconnaissance automatique de la parole

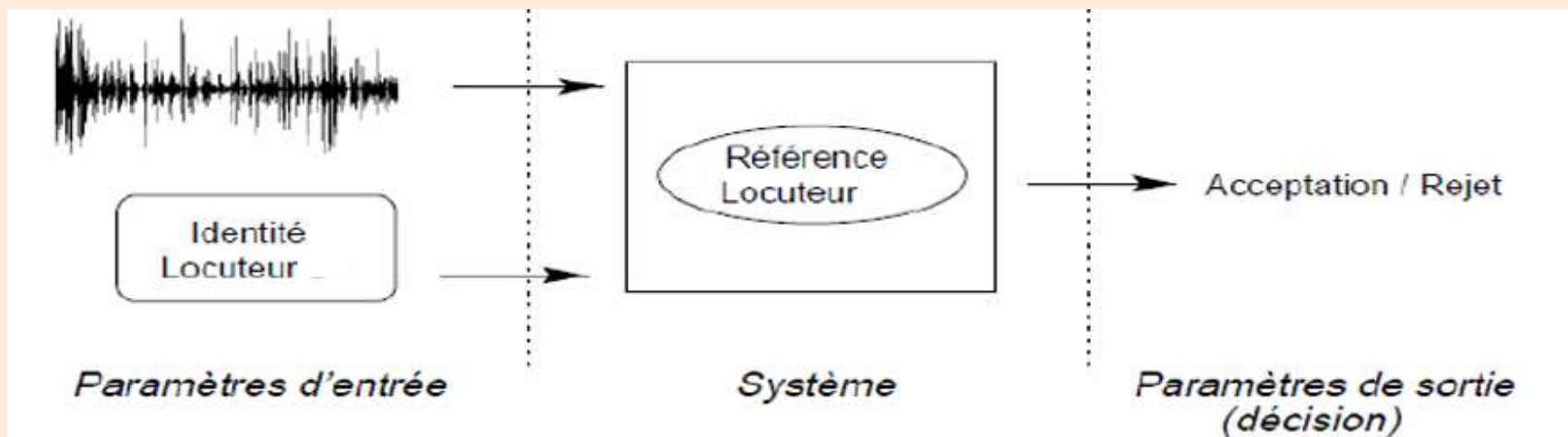
On distingue **plusieurs** types systèmes de reconnaissance:



Reconnaissance automatique de la parole

On distingue plusieurs types de reconnaissance :

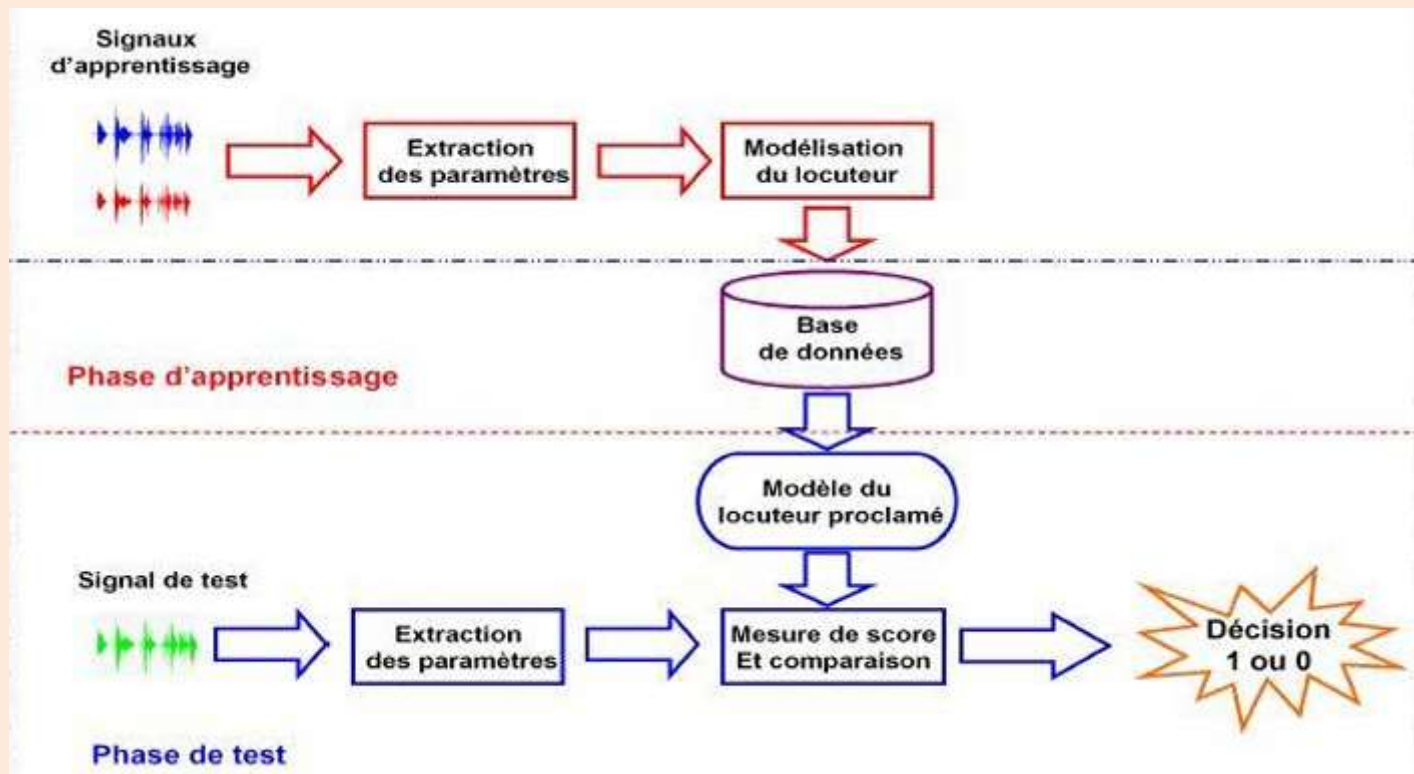
- 1. La reconnaissance du locuteur:** dont le but est de reconnaître la personne qui parle parmi une population de locuteurs (identificateur) ou de vérifier son identité (vérificateur).



Reconnaissance automatique de la parole

On distingue plusieurs types de reconnaissance :

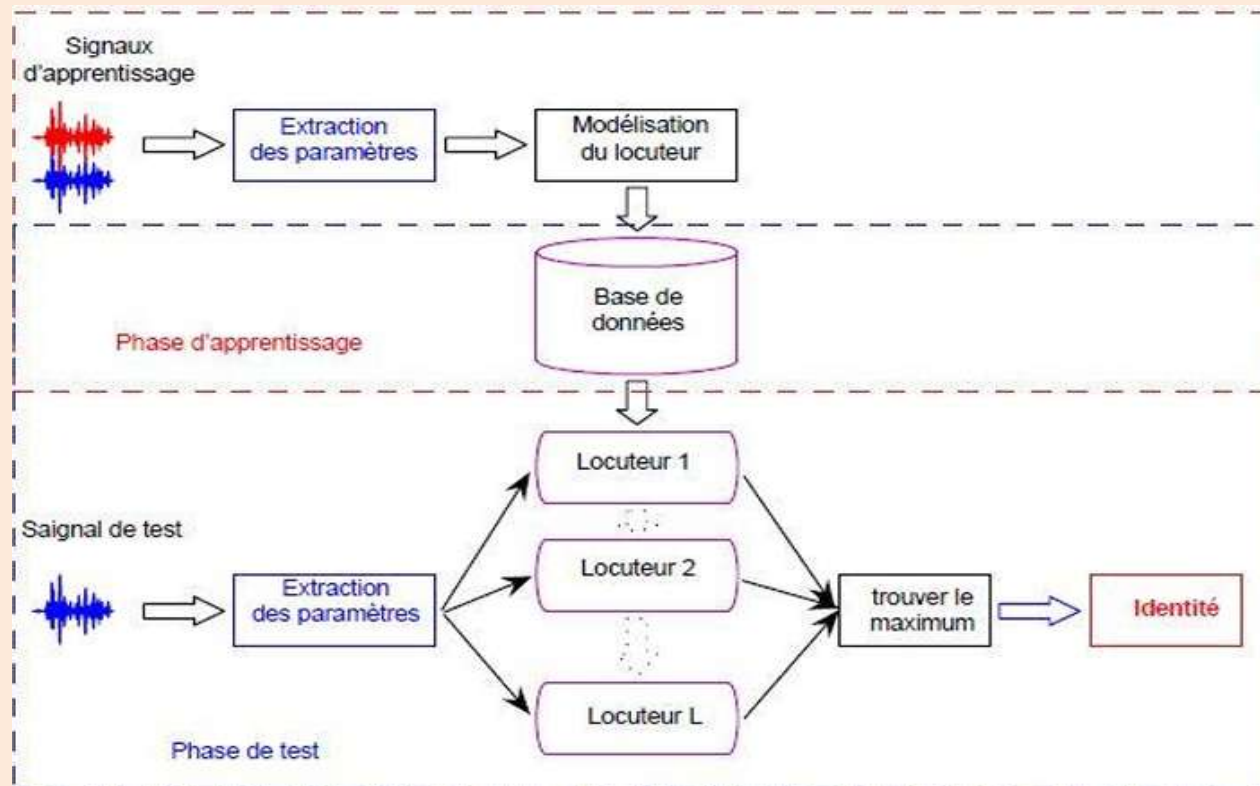
1. La reconnaissance du locuteur:



Reconnaissance automatique de la parole

On distingue plusieurs types de reconnaissance :

1. La reconnaissance du locuteur:



Reconnaissance automatique de la parole

On distingue plusieurs types de reconnaissance :

2. La reconnaissance de parole: dont le but est de transcrire l'information symbolique exprimée par le locuteur (Monolocuteur, Multilocuteur, ou Indépendant du Locuteur).



Reconnaissance automatique de la parole

On distingue plusieurs types de reconnaissance :

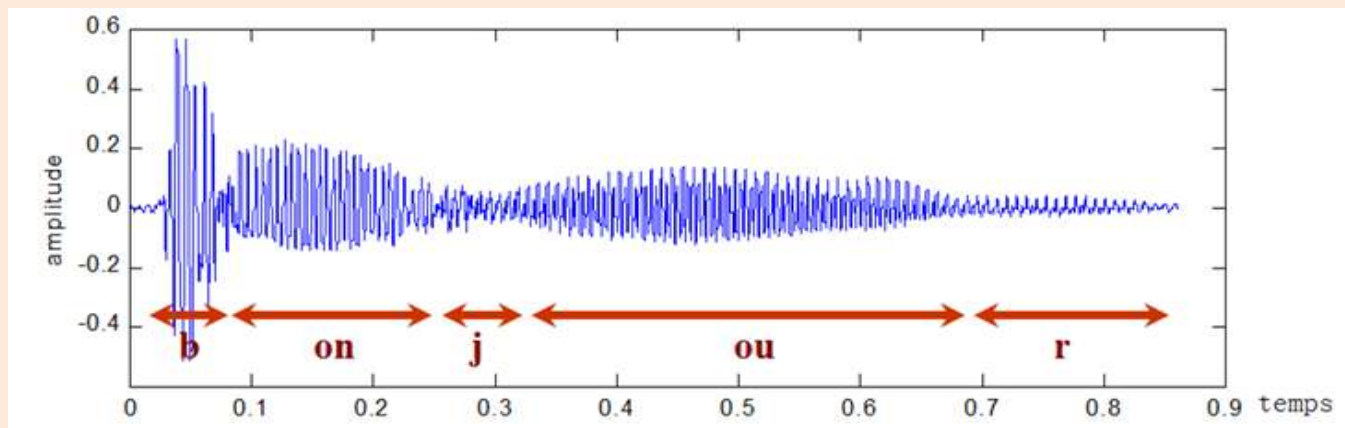
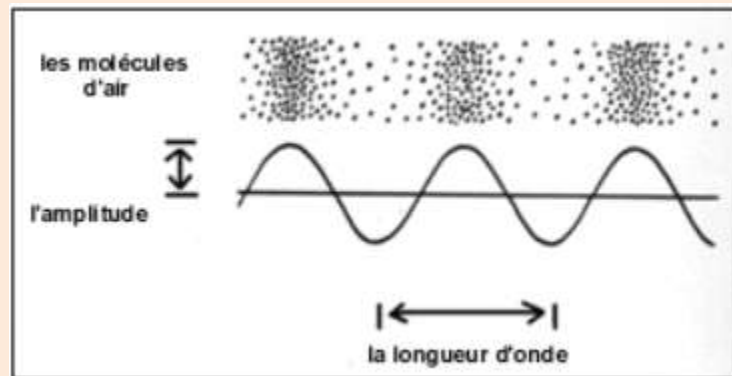
3. La reconnaissance de la Langue: Déterminer dans quelle langue la parole est prononcée.

4. La reconnaissance des émotions: Déterminer l'état psychologique du locuteur (triste, nerveux, content, peur ,etc.)

Caractéristiques du son

Caractéristiques du son

- L'**amplitude** d'un son correspond à la variation de pression maximale de l'air engendrée par les oscillations (volume sonore).



Caractéristiques du son

Une intensité:

Elle dépend de l'amplitude de la vibration.

Plus elle est importante, plus le son est fort.

Plus elle est faible, plus le son est faible.

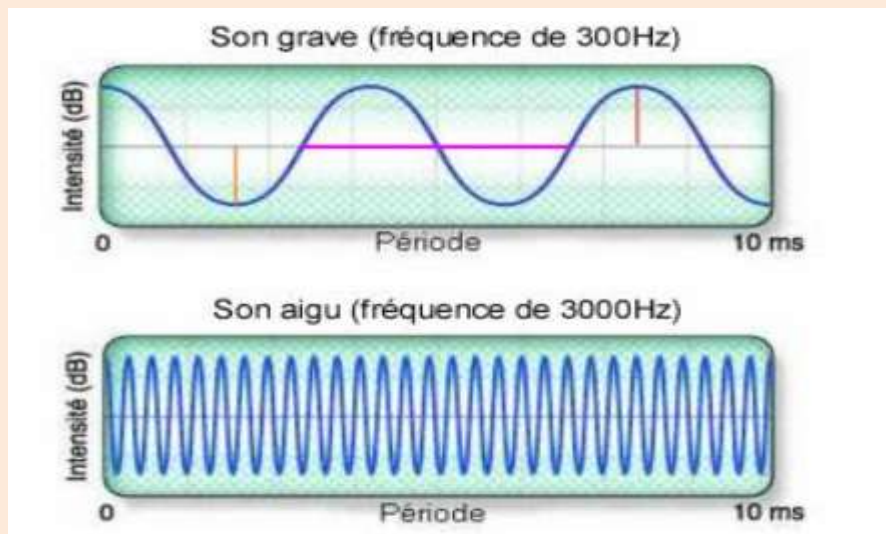
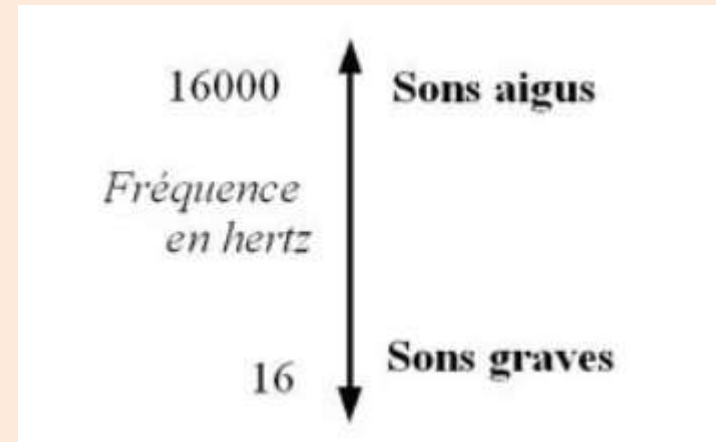
On l'exprime couramment en décibel (dB).



Caractéristiques du son

Une fréquence:

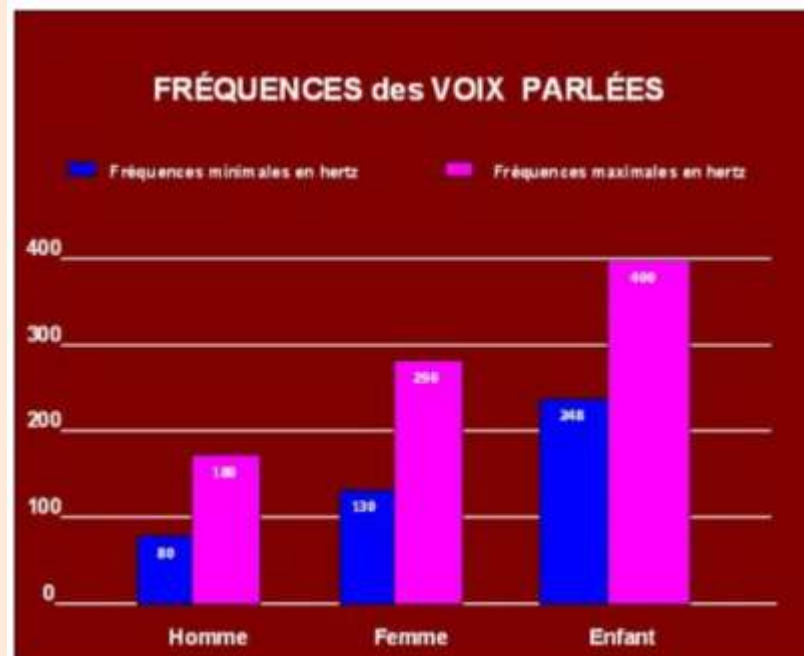
- Le nombre de vibrations par seconde.
- S'il y en a **peu** on entend un son **grave**.
- Sinon, on entend un son **aigu**.
- On exprime la fréquence en **Hertz Hz**.



Caractéristiques du son

Une fréquence:

- De 80Hz à 200Hz pour une voix d'homme.
- De 150Hz à 450Hz pour une voix de femme.
- De 200Hz à 600Hz pour une voix d'enfant.



Reconnaissance automatique de la parole

Speech recognition

Reconnaissance automatique de la parole

Définition

- Une fonctionnalité.
- Elle permet à un programme de:
- Traiter la parole humaine dans un format écrit.



Reconnaissance vocale

Voice recognition

Reconnaissance vocale

Définition

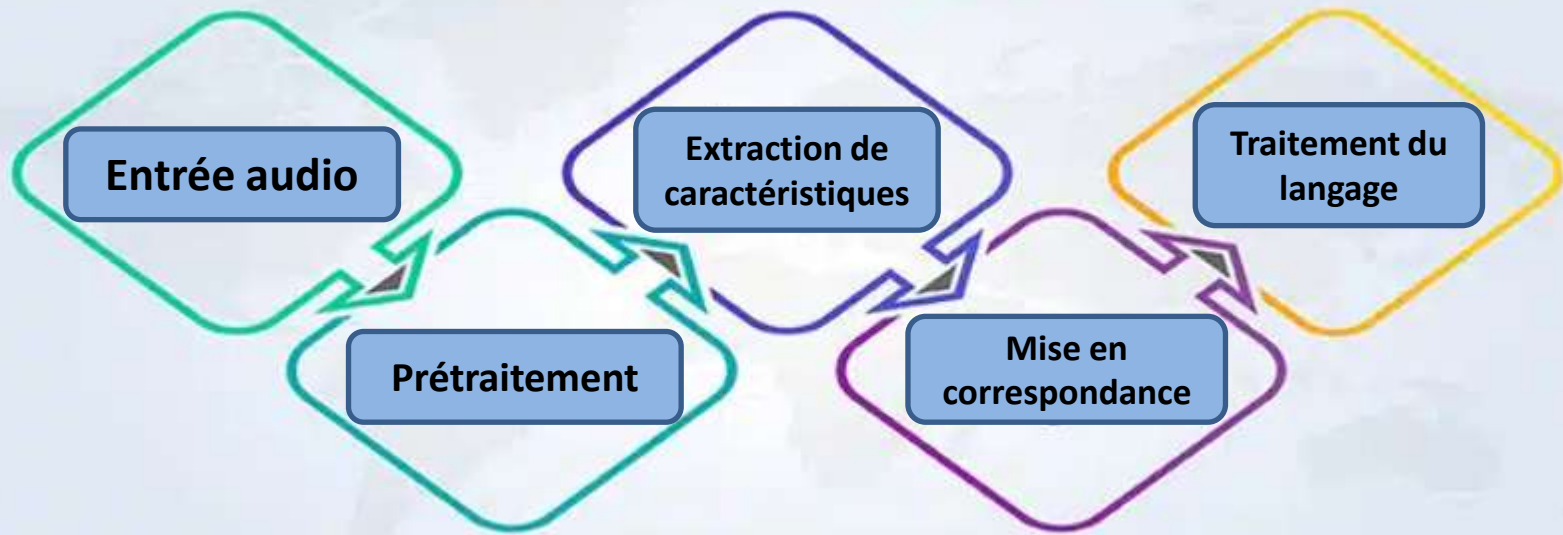
- La **reconnaissance vocale**, également connue sous le nom de reconnaissance du locuteur, permet:

D'identifier, décoder, distinguer et authentifier la voix d'une personne en fonction de son empreinte vocale distincte.



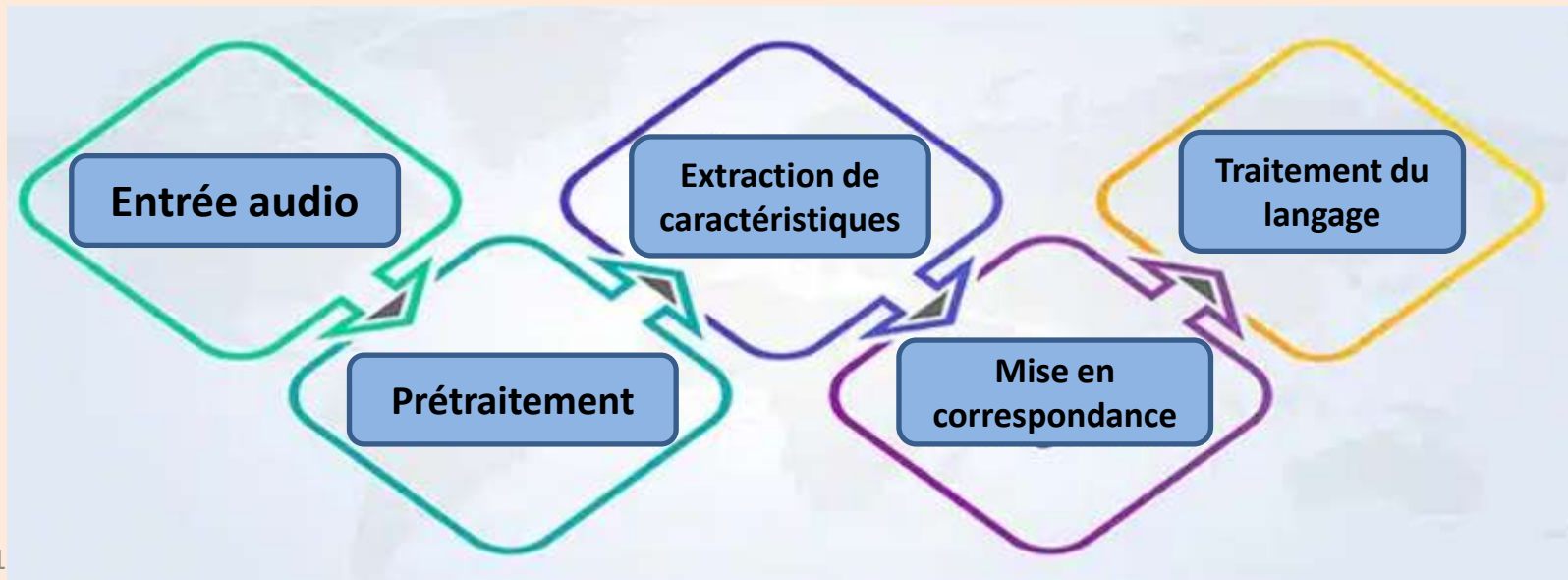
Comment fonctionne la reconnaissance de la parole ?

Comment fonctionne la reconnaissance de la parole ?



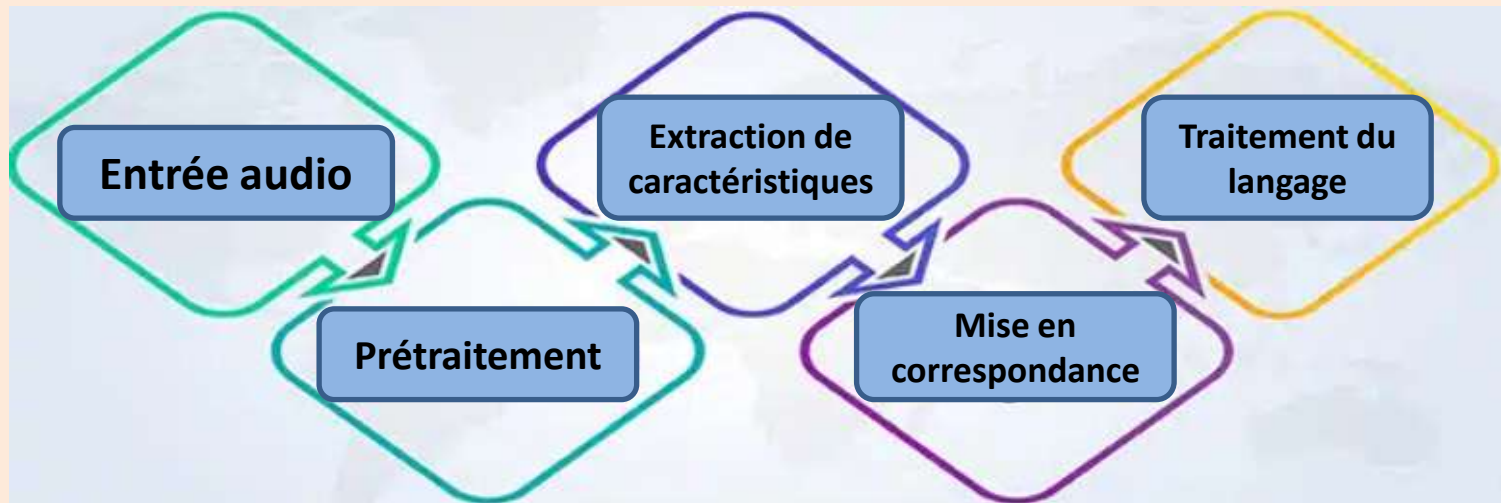
Comment fonctionne la reconnaissance de la parole ?

- **Entrée audio:** Le processus commence par la capture de l'entrée audio à l'aide d'un microphone.
- **Prétraitement:** Le signal audio est nettoyé en diminuant le bruit et en normalisant le volume. Il permet aussi d'éliminer le silence.



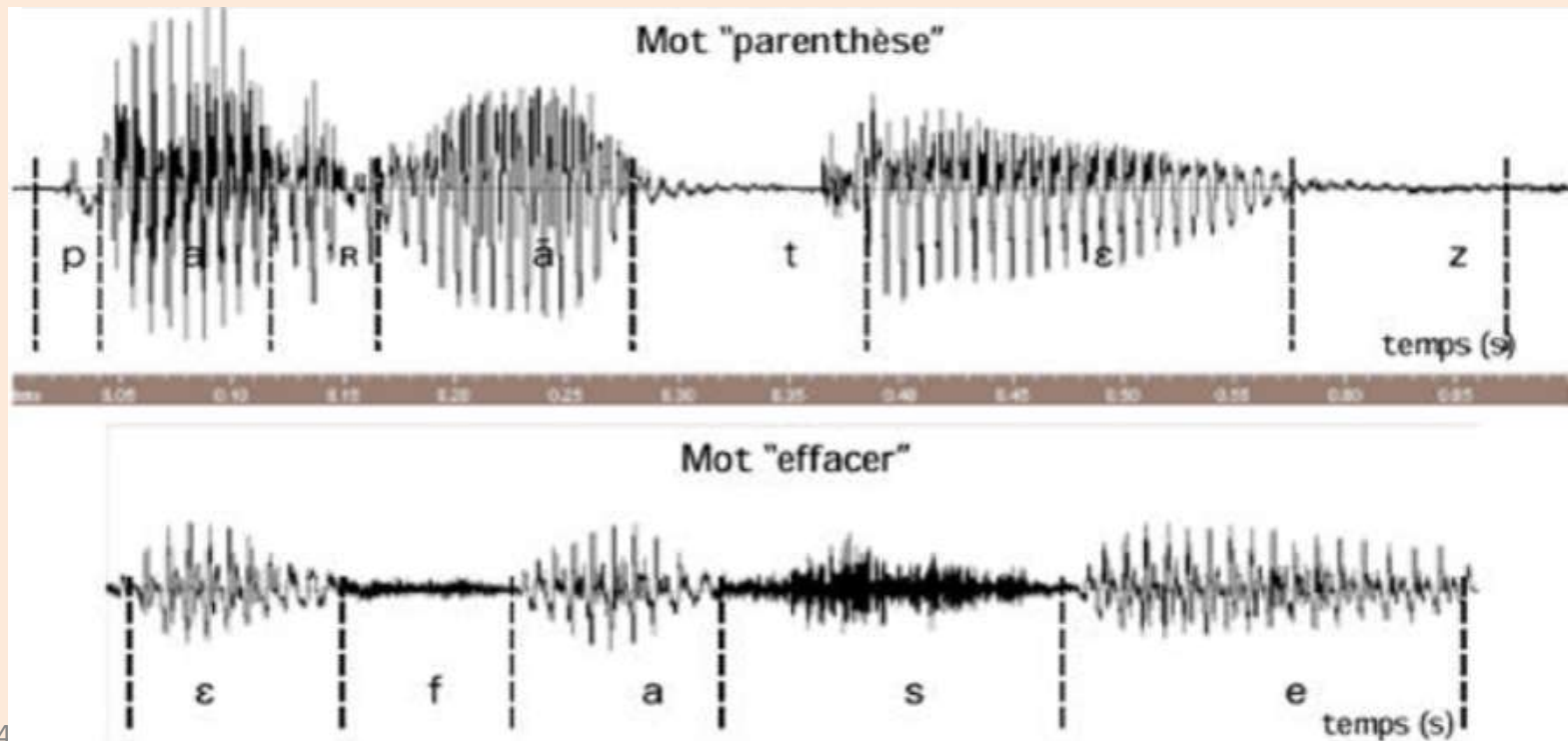
Comment fonctionne la reconnaissance de la parole ?

- **Extraction de caractéristiques:** Le système analyse l'audio pour en extraire les caractéristiques clés telles que la hauteur, la tonalité et la fréquence.



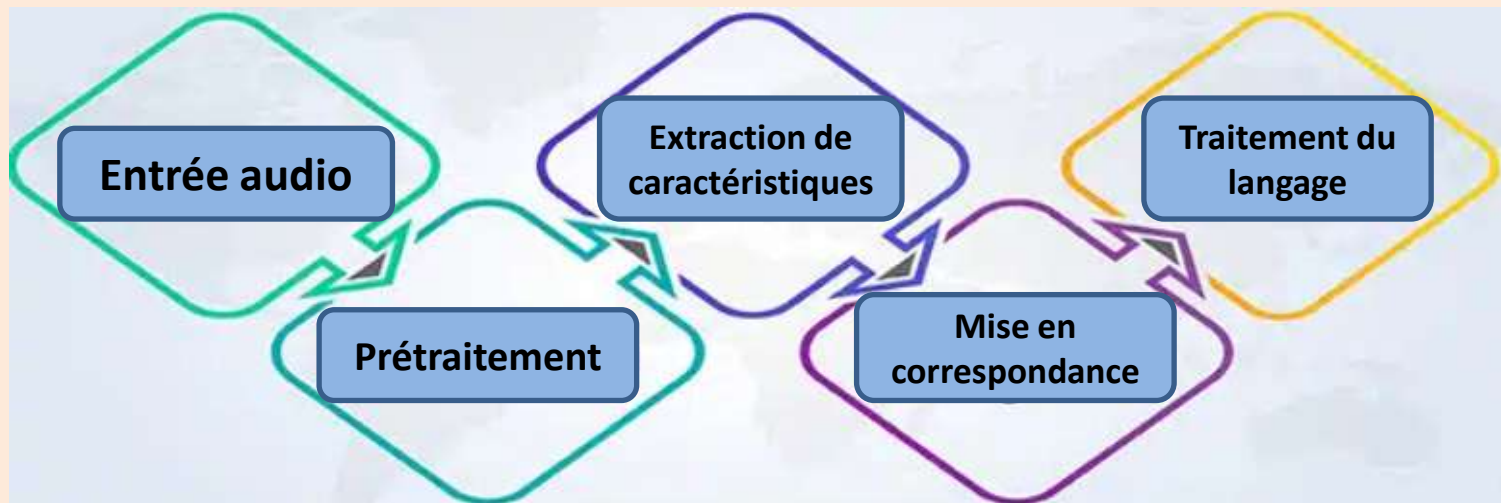
Comment fonctionne la reconnaissance de la parole ?

- **Mise en correspondance:** Les caractéristiques extraites sont comparées à des modèles de parole connus stockés dans une base de données.



Comment fonctionne la reconnaissance de la parole ?

- **Traitement du langage:** Les modèles reconnus sont convertis en texte et des algorithmes de traitement du langage naturel (NLP) interprètent la signification.



Exemples de système de reconnaissance automatique de la parole

Système de reconnaissance automatique de la parole

Reconnaissance des mots isolés :

- Les systèmes de reconnaissance de mots séparés n'acceptent qu'un seul mot à la fois.
- Ces systèmes ont généralement un état "n'écouter/ne pas écouter" qui oblige le locuteur à faire une pause entre les énoncés.

Système de reconnaissance automatique de la parole

Reconnaissance des mots isolés :

La reconnaissance de mots convient lorsque le locuteur doit donner une seule réponse au système ASR ou lorsque les mots représentent des commandes.

Système de reconnaissance automatique de la parole

Reconnaissance des mots connectés :

- Le système de reconnaissance de mots chaînés permet le traitement de mots séparés par des pauses.
- Ils sont similaires aux mots simples, mais ils permettent à des énoncés séparés de s'exécuter ensemble avec des pauses minimales.

Système de reconnaissance automatique de la parole

Reconnaissance de la parole spontanée :

La parole spontanée peut être considérée comme une parole naturelle dont le contenu est inconnu à l'avance.

Un système de reconnaissance qui gère la parole spontanée réelle doit être capable de gérer diverses caractéristiques de la parole naturelle telles que les mots prononcés simultanément (bégaiement léger et non mots, par exemple : "um", "ah", etc.).

Reconnaissance automatique de la parole

Reconnaissance vocale

- La **reconnaissance automatique** de la parole se concentre sur la conversion de la parole d'un format verbal à un format textuel.
- La **reconnaissance vocale** cherche simplement à identifier la voix d'un utilisateur.

	Reconnaissance vocale	Reconnaissance automatique de la parole
Objectif	Identifie et authentifie l'orateur	Reconnaît et transcrit les mots prononcés
Comment ça marche	Analyse les caractéristiques vocales uniques telles que la hauteur, la fréquence et l'accent pour faire correspondre la voix avec une empreinte vocale connue	Utilise des algorithmes pour convertir la langue parlée en texte écrit, en se concentrant sur la compréhension du contenu du discours
Cas d'usage	Systèmes de sécurité, expériences utilisateur personnalisées, authentification biométrique	Assistants virtuels, logiciels de dictée, services de transcription, systèmes de commande et de contrôle
Focus	Qui parle	Ce qui est dit

Reconnaissance vocale

Avantages et inconvénients

Avantages de la reconnaissance vocale	Inconvénients de la reconnaissance vocale
La reconnaissance vocale permet le multitâche et le confort mains libres.	Bien que la technologie de reconnaissance vocale s'améliore à pas de géant, elle n'est pas totalement exempte d'erreurs.
Parler et donner des commandes vocales est beaucoup plus rapide que de taper.	Bruit de fond peuvent interférer avec le fonctionnement et affecter la fiabilité du système.
Les cas d'utilisation de la reconnaissance vocale se développent avec l'apprentissage automatique et les neurones profonds. réseaux.	La confidentialité des données enregistrées est un sujet de préoccupation.

Synthèse vocale

La **synthèse vocale** est:

- Une technique informatique de synthèse sonore.
- Elle permet de générer automatiquement de la parole artificielle à partir de n'importe quel texte.

Synthèse vocale

La synthèse vocale s'appuie à la fois sur des techniques de:

- Traitement linguistique, notamment pour transformer le texte orthographique en une version phonétique prononçable sans ambiguïté.
- Des techniques de traitement du signal pour transformer cette version phonétique en son numérisé écoutable sur un haut parleur.

Synthèse vocale

Parmi les applications, on peut citer:

- La **vocalisation d'écrans** informatiques pour les personnes aveugles.
- Les **annuaires vocaux de grande taille**, où la synthèse vocale est la seule technique viable pour permettre la restitution sonore des noms et des adresses des abonnés.

Défis

Défis

1. **Traitement du Signal:** la séparation de la parole du bruit de fond.
2. **Phonétique:** la variabilité de la voix des locuteurs.
3. **Phonologie:** reconnaître et distinguer les sons.

Défis

- **Les facteurs intra-locuteurs:** variation dans la prononciation, etc.

Cette variation peut résulter:

- De l'état physique ou moral du locuteur.
- Une maladie des voies respiratoires peut ainsi dégrader la qualité du signal de parole.
- L'humeur ou l'émotion du locuteur peut influencer son rythme d'élocution.

Défis

- **Les facteurs inter-locuteurs:**
 - Physiologie.
 - Age.
 - Sexe.
 - Psychologie.
 - Familiarité avec l'application, etc.

Défis

- **L'environnement:** la variabilité due à l'environnement peut provoquer une dégradation du signal de parole sans que le locuteur ait modifié son mode d'élocution.

Défis

- **Bruit:**

les moyens de transport peuvent entraîner des déformations du signal.

Le bruit ambiant peut ainsi provoquer une déformation du signal de parole en obligeant le locuteur à accentuer son effort vocal.

- Micro.
- Canal de transmission.
- Présence d'autres locuteurs, etc.

Merci pour votre attention

