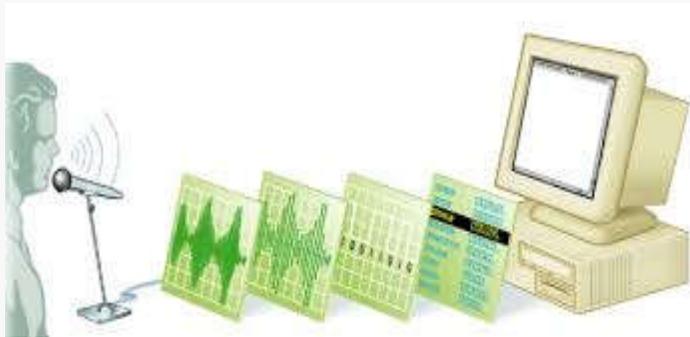
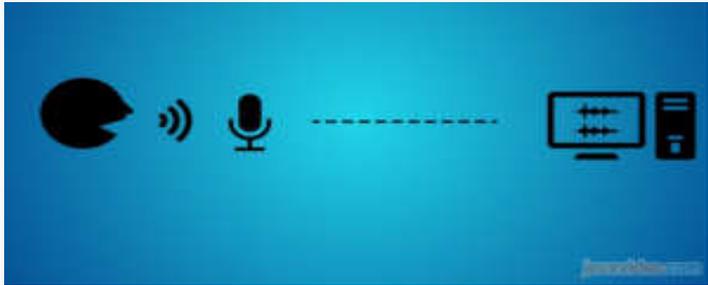
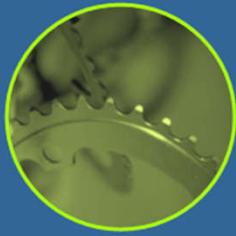


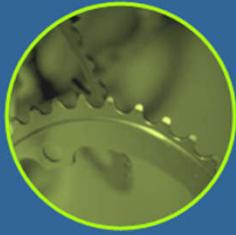
TAP

Traitement Automatique de la Parole Fondements théoriques Logiciels, exemples et applications

Encadré par :
Mme Catherine RECANATI

Présenté par :
Louiza AOUCHETA
Dofinisséri Désiré KANI





Plan

1

Définitions

2

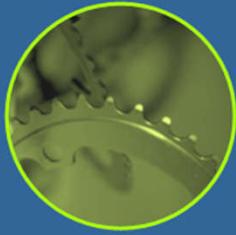
Historique

3

Fondements théoriques

4

Logiciels, exemples et applications

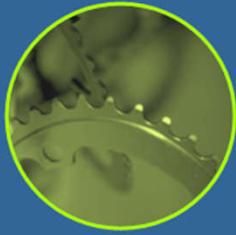


TAP

1. Définitions

Qu'est-ce que le **traitement automatique de la parole**?

Discipline technologique dont l'objectif est la captation, la transmission, l'identification et la synthèse de la parole.



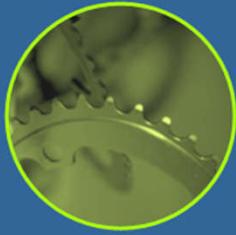
TAP

1. Définitions

Qu'est-ce que le **traitement automatique de la parole**?

Deux principaux domaines :

- Reconnaissance vocale**
- Synthèse vocale**

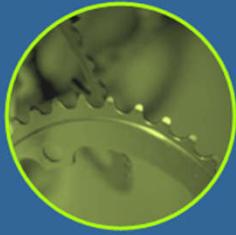


TAP

1. Définitions

Qu'est-ce que la **reconnaissance vocale**?

Permettre à la **machine** de **comprendre** et de **traiter** des **informations** fournies **oralement** par un **utilisateur humain**.

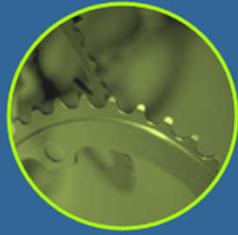


TAP

1. Définitions

Qu'est-ce que la **reconnaissance vocale**?

Utilisation des **techniques d'appariement**
pour de **comparer** une **onde sonore** à un
ensemble d'échantillons, composés de
mots ou de **phonèmes**

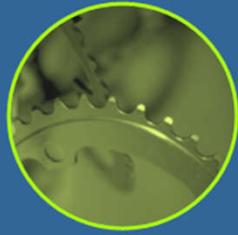


TAP

1. Définitions

Quelques connaissances utilisées en TAP

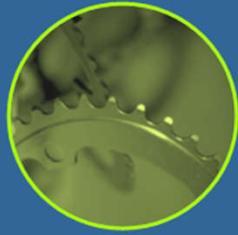
anatomie (les fonctions de l'appareil phonatoire et de l'oreille), les **signaux émis par la parole**, la **phonétique**, le **traitement du signal**, la **linguistique**, **l'informatique**, **l'intelligence artificielle** et les **statistiques**



TAP *2. Historique*

Cas de la **reconnaissance vocale** :

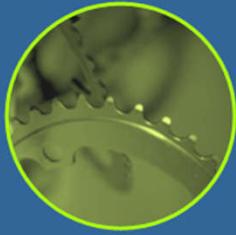
- ❑ Début XXe siècle : début des travaux
- ❑ **1952** : **premier système électronique** de reconnaissance vocale développé par **Davis, Biddulph et Balashek** aux **laboratoires Bell Labs**



TAP *2. Historique*

Cas de la **reconnaissance vocale** :

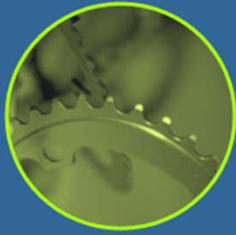
- ❑ **1970** : amélioration de la recherche avec les travaux de **Jelinek** chez IBM (1972-1993)
- ❑ **1972** : première commercialisation d'un **système de reconnaissance vocale** par la société **Threshold Technologies**



TAP *2. Historique*

Cas de la **synthèse vocale** :

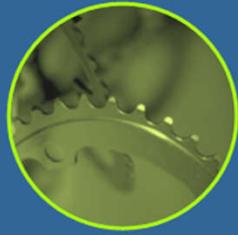
- ❑ De 1965 à 1985 : **synthèse vocale par règles** ou **synthèse vocale par formants** - basée sur la modélisation paramétrique du spectre sonore de la parole et **entièrement synthétique**



TAP *2. Historique*

Cas de la **synthèse vocale** :

- ❑ **A partir de 1985 : synthèse vocale semi-synthétique** - consiste à assembler des petits segments élémentaires de parole naturelle pour constituer n'importe quel énoncé synthétique souhaité

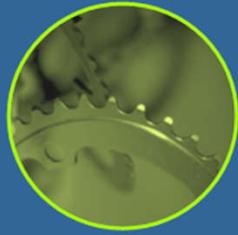


TAP

3. Fondements théoriques

Cas de la reconnaissance vocale :

- Traitement acoustique**
- Apprentissage automatique**
- Décodage**
- Modèles**



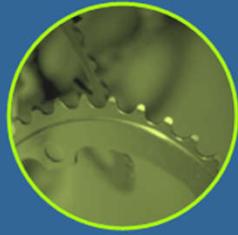
TAP

3. Fondements théoriques

Cas de la reconnaissance vocale :

□ Traitement acoustique

Permet principalement d'extraire du signal vocal une image acoustique compacte sous forme de vecteurs acoustiques correspondant à des tranches de 20 à 30ms de signal avec un pas de 10ms



TAP

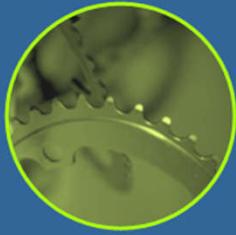
3. Fondements théoriques

Cas de la reconnaissance vocale :

□ Apprentissage automatique

Réalise une association entre les segments élémentaires de la parole et les éléments lexicaux

Utilise les modèles de Markov cachés et/ou réseaux de neurones artificiels



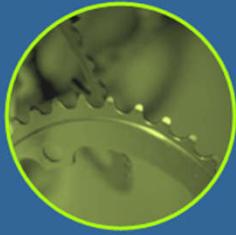
TAP

3. Fondements théoriques

Cas de la reconnaissance vocale :

□ Décodage

Concaténation des modèles élémentaires précédemment appris pour reconstituer le discours le plus probable.



TAP

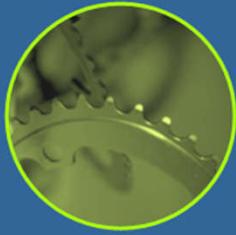
3. Fondements théoriques

Cas de la reconnaissance vocale :

□ Modèles

Le système s'appuie sur trois modèles principaux :

- **Modèle de langage**
- **Modèle de prononciation**
- **Modèle acoustico-phonétique**



TAP

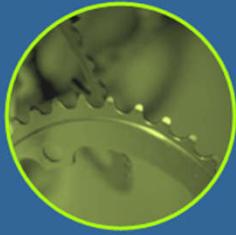
3. Fondements théoriques

Cas de la reconnaissance vocale :

□ Modèles

- **Modèle de langage** : ce modèle donne la probabilité de chaque suite de mots dans le langage cible.

$P(W)$



TAP

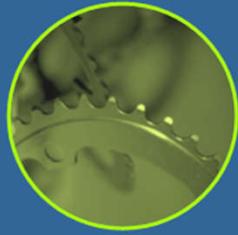
3. Fondements théoriques

Cas de la reconnaissance vocale :

□ Modèles

- **Modèle de prononciation** : ce modèle donne pour chaque suite de mots, la ou les prononciations possibles avec leurs probabilités .

$$P(H|W)$$



TAP

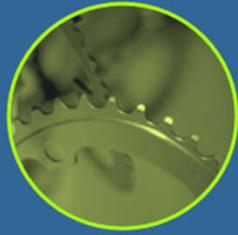
3. Fondements théoriques

Cas de la reconnaissance vocale :

□ Modèles

- **Modèle acoustico-phonétique** : ce modèle estime la probabilité de la séquence observée de vecteurs acoustiques étant donnée une prononciation possible d'une séquence de mots donnée.

$$P(X|H)$$



TAP

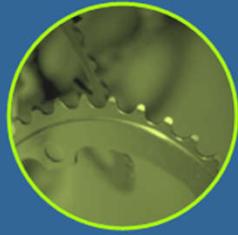
3. Fondements théoriques

Cas de la reconnaissance vocale :

□ Modèles

La combinaison des trois modèles permet de calculer la probabilité de toute suite de mots étant donné un signal vocal observé.

La reconnaissance de la parole consiste à trouver la suite de mots qui a la probabilité la plus élevée

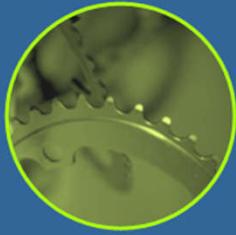


TAP

3. Fondements théoriques

Cas de la **synthèse vocale** :

- Transcription phonétique**
- Formants**
- Intonation et prosodie**



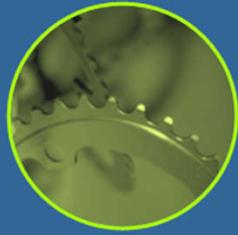
TAP

3. Fondements théoriques

Cas de la **synthèse vocale** :

□ **Transcription phonétique**

La technique classique pour réaliser la transcription phonétique est d'appliquer à la suite de lettres composant le texte une série de règles de réécritures dépendant du contexte



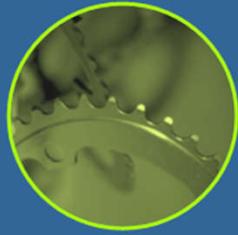
TAP

3. Fondements théoriques

Cas de la **synthèse vocale** :

□ **Transcription phonétique**

Il est aussi nécessaire de connaître la catégorie grammaticale de certains mots: par exemple dans la phrase "les présidents président" la catégorie grammaticale du mot "président" détermine sa voyelle finale ("e" muet ou le digramme "en").



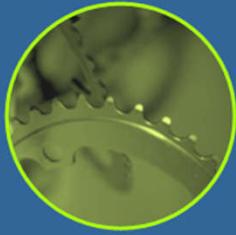
TAP

3. Fondements théoriques

Cas de la **synthèse vocale** :

Transcription phonétique

Pour les mots dont la prononciation n'est pas régulière, et notamment pour les noms propres, on stocke tout simplement leur transcription phonétique dans un lexique d'exception.



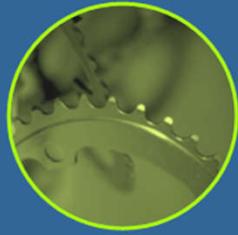
TAP

3. Fondements théoriques

Cas de la **synthèse vocale** :

Formants

La synthèse par formant repose typiquement sur la description des trois premiers formants du spectre de la parole.



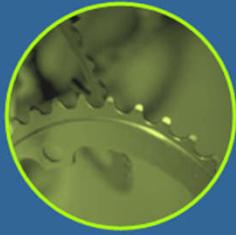
TAP

3. Fondements théoriques

Cas de la **synthèse vocale** :

□ **Formants**

Chaque **formant** (maximum du spectre de parole) est classiquement décrit par trois paramètres, sa **fréquence** (en Hz), son **amplitude** (en dB) et sa **bande passante** (en Hz).



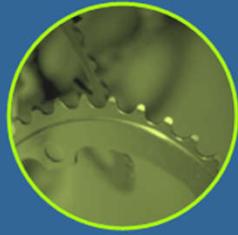
TAP

3. Fondements théoriques

Cas de la **synthèse vocale** :

Formants

L'**amplitude** représente l'**intensité du signal** à la fréquence du formant tandis que la **bande passante** représente la **largeur du spectre** autour du maximum formantique.



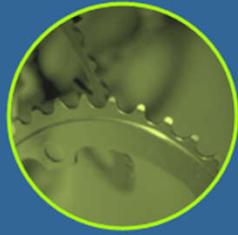
TAP

3. Fondements théoriques

Cas de la **synthèse vocale** :

Intonation et prosodie

Constitue l'un des trois éléments de la prosodie, les deux autres paramètres prosodiques étant le rythme et l'intensité.



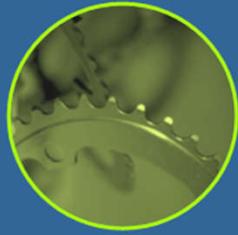
TAP

3. Fondements théoriques

Cas de la **synthèse vocale** :

Intonation et prosodie

Pour être perçue comme naturelle, la synthèse vocale nécessite d'imiter une prosodie (durée, mélodie et rythme des sons) naturelle dans son ensemble.



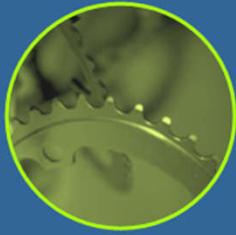
TAP

3. Fondements théoriques

Cas de la **synthèse vocale** :

Intonation et prosodie

C'est une caractéristique importante de la voix naturelle, qui dépend également très fortement de la langue concernée et du contenu linguistique de la phrase.



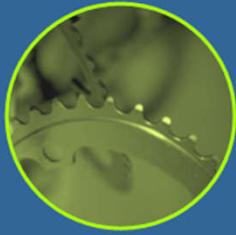
TAP

3. Fondements théoriques

Cas de la **synthèse vocale** :

Intonation et prosodie

Un énoncé synthétisé sans intonation est perçu comme anormal, et à la limite une voix synthétisée avec une fréquence fondamentale absolument constante, est perçue comme totalement artificielle ("voix de robot").



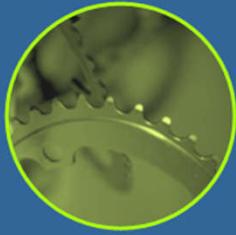
TAP

4. Logiciels, exemples et applications

Cas de la reconnaissance vocale :

□ Logiciel : fonctionnement

Il reconnaît la parole de l'utilisateur, au moyen d'un microphone, pour la retranscrire instantanément en format texte.



TAP

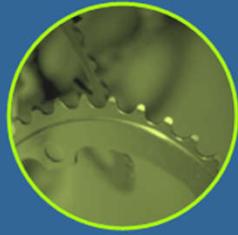
4. Logiciels, exemples et applications

Cas de la reconnaissance vocale :

□ **Logiciel : fonctionnement**

Il peut également servir comme outil de commande de l'ordinateur en se substituant par exemple à l'usage d'une souris.





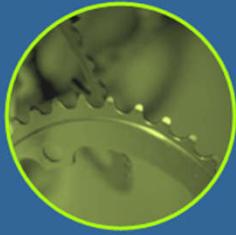
TAP

4. Logiciels, exemples et applications

Cas de la reconnaissance vocale :

□ Logiciel : fonctionnement

Un son correspondant à un mot est prononcé oralement devant le microphone, qui convertit le signal acoustique en un signal électrique. Celui-ci est soumis à un prétraitement (numérisation, séparation du signal pertinent par rapport au bruit ambiant), puis il subit le paramétrage (extraction de paramètres).



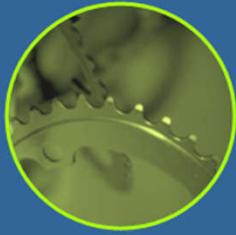
TAP

4. Logiciels, exemples et applications

Cas de la reconnaissance vocale :

Logiciel : fonctionnement

Cliquez [ici](#) pour voir une vidéo (durée : 3 minutes 25) du fonctionnement!



TAP

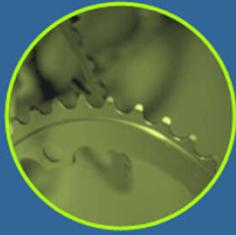
4. Logiciels, exemples et applications

Cas de la reconnaissance vocale :

□ **Logiciels : démo en vidéo de deux exemples**

Démo de **Dys vocal** : cliquez [ici](#)

Pour **Dragon naturally speaking** : cliquez [ici](#)



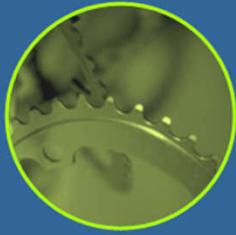
TAP

4. Logiciels, exemples et applications

Cas de la reconnaissance vocale :

□ Logiciels : applications

- La compensation personnel pour rédiger, grâce à la dictée vocale, des documents, courriels, dossiers, rapports, etc...
- La compensation des difficultés liées à la prise de notes écrites : troubles moteurs, troubles du langage et de la parole, troubles auditifs...



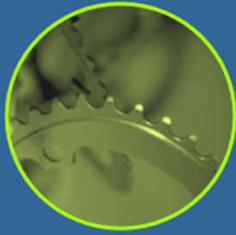
TAP

4. Logiciels, exemples et applications

Cas de la reconnaissance vocale :

Logiciels : applications

- Pour un public sourd, pour obtenir “un sous-titrage” du discours sur ordinateur.
- Pour une situation de dysorthographe
- Dans le cas d’un trouble de l’élocution, l’utilisation reste possible avec paramétrage adapté.



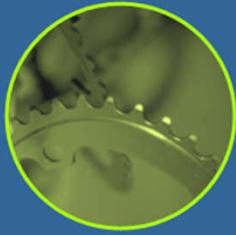
TAP

4. Logiciels, exemples et applications

Cas de la **synthèse vocale** :

Logiciels : fonctionnement

Ils utilisent à la fois des techniques de traitement linguistique et des techniques de traitement du signal.



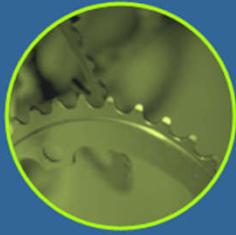
TAP

4. Logiciels, exemples et applications

Cas de la **synthèse vocale** :

Logiciels : fonctionnement

Ils opèrent d'abord une **transcription phonétique** : ils transcrivent la suite de lettres composant le texte en une série de phonèmes qui représentent exactement les sons devant être prononcés.



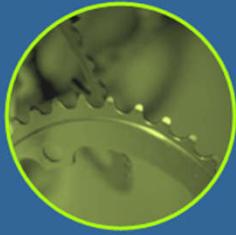
TAP

4. Logiciels, exemples et applications

Cas de la **synthèse vocale** :

Logiciels : fonctionnement

Ils restituent ensuite le contenu du texte grâce à des **sons synthétisés**. Il s'agit en fait de morceaux d'enregistrements de parole qui sont artificiellement "attachés" les uns à la suite des autres.



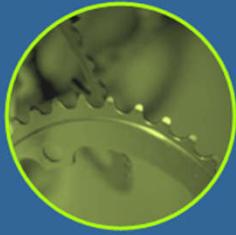
TAP

4. Logiciels, exemples et applications

Cas de la **synthèse vocale** :

Logiciels : fonctionnement

Le principal défi de la parole synthétisée est de **produire une impression de voix naturelle** et non une voix monocorde de robot.



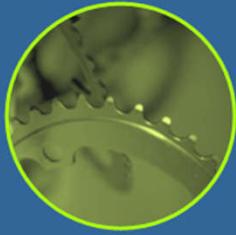
TAP

4. Logiciels, exemples et applications

Cas de la **synthèse vocale** :

Logiciels : fonctionnement

Pour obtenir un résultat proche de la parole humaine, les logiciels doivent tenir compte du **rythme, de l'intonation et de l'accent tonique.**



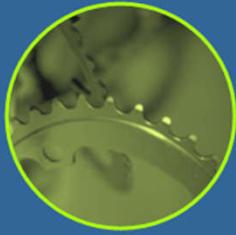
TAP

4. Logiciels, exemples et applications

Cas de la **synthèse vocale** :

Logiciels : exemples

- **Natural reader**
- **Balabolka**
- **Loquendo (en ligne)**
- **Artonik (en ligne)**



TAP

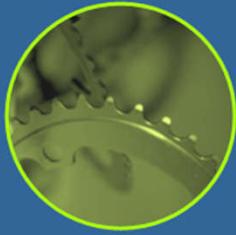
4. Logiciels, exemples et applications

Cas de la **synthèse vocale** :

□ **Logiciels : exemples**

- **Balabolka**

Il permet la lecture des fichiers texte à haute voix en vue de reproduire la voix humaine. La reproduction de la voix humaine est contrôlable à l'aide des boutons standards semblables à ceux de tout programme multimédia (« reproduire », « pause », « arrêter »).



TAP

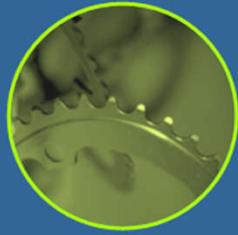
4. Logiciels, exemples et applications

Cas de la **synthèse vocale** :

Logiciels : exemples

- **Balabolka**

Le programme est capable de reproduire le contenu du presse-papiers, afficher le texte contenant dans les textes avec différentes extensions.



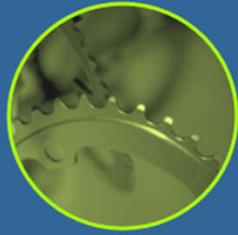
TAP

4. Logiciels, exemples et applications

Cas de la **synthèse vocale** :

Logiciels : applications

- La lecture d'e-mails et de SMS pour les personnes déficientes visuelles
- Les annonces dans les gares et les GPS



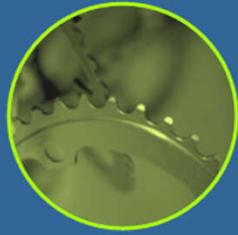
TAP

4. Logiciels, exemples et applications

Cas de la **synthèse vocale** :

Logiciels : applications

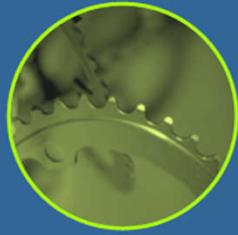
- Certains distributeurs de banque
- La vocalisation d'écrans informatiques pour les malvoyants
- L'aide aux troubles d'apprentissage



TAP
Sources

Définitions, historique, fondements théorique :

- <https://slideplayer.fr/slide/3092598/>
- https://fr.wikipedia.org/wiki/Traitement_de_la_parole
- <https://perso.telecom-paristech.fr/grichard/Enseignements/AIC/Bases-signal-pour-traitement-parole.pdf>
- <https://alexandre.alapetite.fr/dess-irr/tap/tap02/index.html>
- <http://dictionnaire.sensagent.leparisien.fr/Synth%C3%A8se%20vocale/fr-fr/>
- [http://dictionnaire.sensagent.leparisien.fr/Reconnaissance%20vocale/fr-fr/#Principe de base](http://dictionnaire.sensagent.leparisien.fr/Reconnaissance%20vocale/fr-fr/#Principe_de_base)



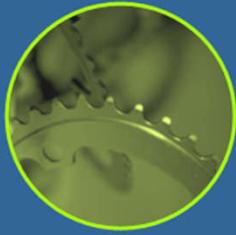
TAP
Sources

Vidéos :

- <https://www.youtube.com/watch?v=tEmXaHPE6Ys>
- <https://www.youtube.com/watch?v=se2kJKkR8nY>
- <https://www.youtube.com/watch?v=UEMxjbSNYIo>
- https://www.youtube.com/watch?v=YOetqn_1eW4&t=187s

Démo de logiciels en ligne

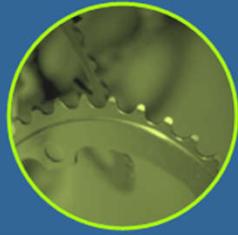
- <https://www.nuance.com/omni-channel-customer-engagement/voice-and-ivr/text-to-speech.html#!>
- [https://www.artonik.com/fr/outils~\\$c12/synthese-vocal-en-ligne/synthese-vocal-en-ligne~\\$a1047/synthese-vocal-en-ligne~infos~.aspx](https://www.artonik.com/fr/outils~$c12/synthese-vocal-en-ligne/synthese-vocal-en-ligne~$a1047/synthese-vocal-en-ligne~infos~.aspx)



TAP
Sources

Logiciels, exemples et applications :

- <http://asso-apaches.fr/faq/utilisation-de-logiciels-de-reconnaissance-vocale/>
- <https://ecolebranchee.com/logiciels-reconnaissance-synthese-vocale/>
- <https://www.nuance.com/fr-fr/dragon.html>
- <https://www.gralon.net/articles/materiel-et-consommables/materiel-informatique-et-consommable-informatique/article-la-synthese-vocale---presentation-et-fonctionnement-1910.htm>
- <https://www.atingo.be/tout-sexplique-la-synthese-vocale/>
- <http://coridys.fr/handicap-et-compensation/logiciels/liste-logiciels-adaptes-synthese-vocale-dictee-vocale-lenregistrement/>
- <https://www.nch.com.au/verbose/fr/index.html>



*InHM : TAP - Traitement
Automatique de la Parole*



Merci de votre aimable attention!