

# *Introduction au TAL*

## Table des matières

1. Introduction .....	1
2. Outils de base .....	1
2.1. Le traitement de la parole .....	1
2.2. Les niveaux d'analyse de la parole .....	2
2.2.1. Le niveau acoustique .....	2
2.2.1.1. Fréquence fondamentale .....	2
2.2.1.2. Spectrogramme .....	2
2.2.1.3. Audiogramme .....	3
3. Synthèse de la parole .....	4
3.1. Concepts de base .....	4
3.2. Applications .....	4
3.3. Synthèse par règles - Synthèse par concaténation .....	4
3.3.1. Synthèse par règles .....	4
3.3.2. Synthèse par concaténation .....	4
4. Conclusion .....	4

## Liste des figures

FIGURE 1 TRAITEMENT DE LA PAROLE .....	1
FIGURE 2 L'EVOLUTION TEMPORELLE DU SIGNAL VOCAL POUR LES MOTS 'PARENTHESE', ET 'EFFACER' .....	2
FIGURE 3 LES TRAITES ACOUSTIQUES DU SIGNAL .....	3
FIGURE 4 FREQUENCE FONDAMENTALE.....	3

## 1. Introduction

On regroupe sous le vocable de traitement automatique du langage naturel (TALN) l'ensemble des recherches et développements visant à modéliser et reproduire, à l'aide de machines, la capacité humaine à produire et à comprendre des énoncés linguistiques dans des buts de communication. Il sera donc question ici de langage humain, d'où l'adjectif naturel, et non pas de langage formel, tel que C ou encore ADA. Ce naturel fait d'ailleurs tout le sel de l'affaire : les langages formels sont précisément conçus et optimisés dans l'optique de manipulations algorithmiques. Il en va tout autrement pour le langage naturel, dont le traitement automatique pose des difficultés majeures (voir la section 0.1.3). Précision importante, nous nous limiterons quasiment exclusivement au traitement du langage sous forme écrite, le traitement de la parole étant encore, en dépit de convergences de plus en plus marquées avec le traitement de l'écrit, une question de traitement du signal.

## 2. Outils de base

### 2.1. Le traitement de la parole

Le traitement de la parole est aujourd'hui une composante fondamentale des sciences de l'ingénieur. Située au croisement du traitement du signal numérique et du traitement du langage (c'est-à-dire du traitement de données symboliques), cette discipline scientifique a connu depuis les années 60 une expansion fulgurante, liée au développement des moyens et des techniques de télécommunications.

S'il n'est pas en principe de parole sans cerveau humain pour la produire, l'entendre, et la comprendre, les techniques modernes de traitement de la parole tendent cependant à produire des systèmes automatiques qui se substituent à l'une ou l'autre de ces fonctions :

- Les **analyseurs** de parole cherchent à mettre en évidence les caractéristiques du signal vocal tel qu'il est produit, ou parfois tel qu'il est perçu (on parle alors d'analyseur perceptuel), mais jamais tel qu'il est compris, ce rôle étant réservé aux reconnaisseurs.
- Les **reconnaisseurs** ont pour mission de décoder l'information portée par le signal vocal à partir des données fournies par l'analyse.

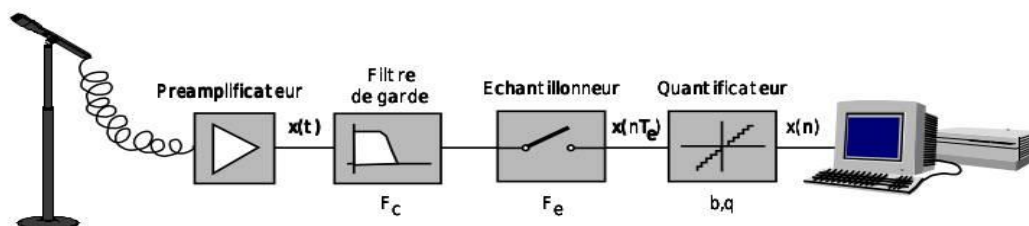


Figure 1 traitement de la parole

Les **synthétiseurs** ont quant à eux la fonction inverse de celle des analyseurs et des reconnaisseurs de parole : ils produisent de la parole artificielle. On distingue fondamentalement deux types de synthétiseurs : les synthétiseurs de parole à partir d'une représentation numérique, inverses des analyseurs, dont la mission est de produire de la parole à partir des caractéristiques numériques d'un signal vocal telles qu'obtenues par analyse, et les synthétiseurs de parole à partir d'une représentation symbolique, inverse des reconnaisseurs de parole et capables en principe de prononcer n'importe quelle phrase sans qu'il soit nécessaire de la faire prononcer par un locuteur humain au préalable.

## 2.2. Les niveaux d'analyse de la parole

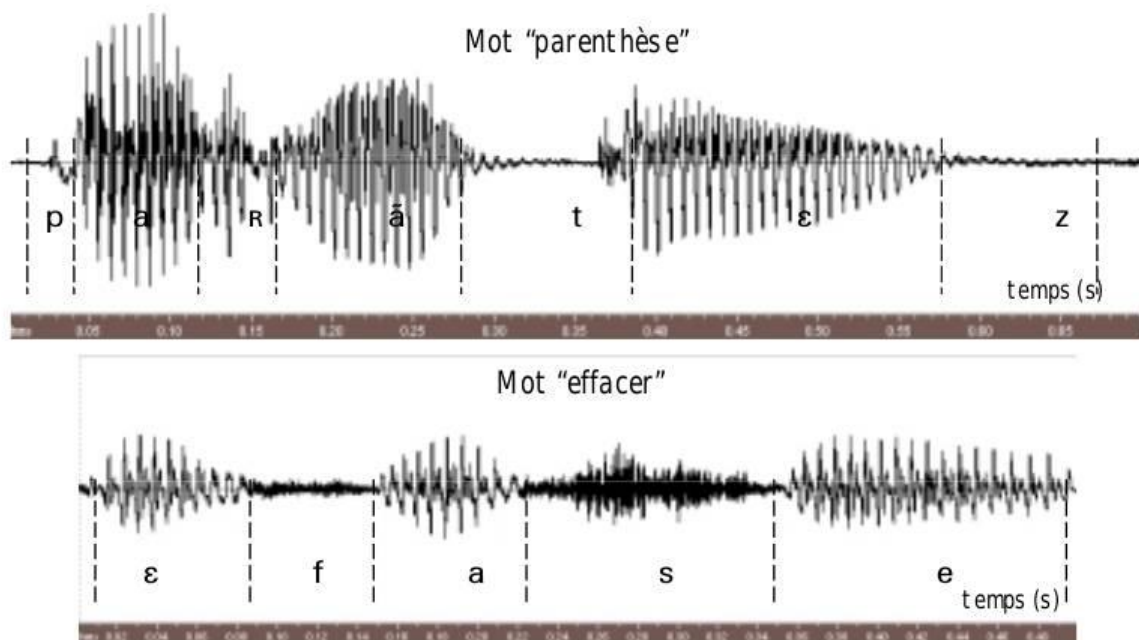
### 2.2.1. Le niveau acoustique

#### 2.2.1.1. Fréquence fondamentale

Une analyse d'un signal de parole n'est pas complète tant qu'on n'a pas mesuré l'évolution temporelle de la fréquence fondamentale  $f_0$  ou pitch.

#### 2.2.1.2. Spectrogramme

Il est souvent intéressant de représenter l'évolution temporelle du spectre à court terme d'un signal, sous la forme d'un spectrogramme. L'amplitude du spectre y apparaît sous la forme de niveaux de gris dans un diagramme en deux dimensions temps-fréquence.



**Figure 2** l'évolution temporelle du signal vocal pour les mots 'parenthèse', et 'effacer'

On parle de spectrogramme à large bande ou à bande étroite selon la durée de la fenêtre de pondération. Les spectrogrammes à bande large sont obtenus avec des fenêtres de pondération de faible durée (typiquement 10 ms); ils mettent en évidence l'enveloppe spectrale du signal,

et permettent par conséquent de visualiser l'évolution temporelle des formants. Les périodes voisées y apparaissent sous la forme de bandes verticales plus sombres. Les spectrogrammes à bande étroite sont moins utilisés. Ils mettent plutôt la structure fine du spectre en évidence : les harmoniques du signal dans les zones voisées y apparaissent sous la forme de bandes horizontales.

### 2.2.1.3. Audiogramme

L'opération de numérisation, schématisée à la figure 1.1, requiert successivement : un filtrage de garde, un échantillonnage, et une quantification.

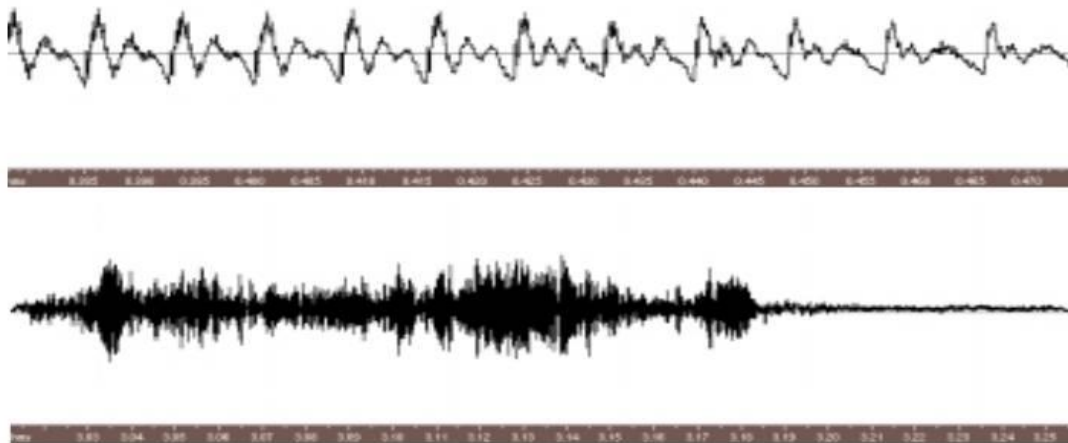


Figure 3 les traits acoustiques du signal

La figure 1.2 représente l'évolution temporelle, ou audiogramme, du signal vocal pour les mots 'parenthèse', et 'effacer'. On y constate une alternance de zones assez périodiques et de zones bruitées, appelées zones voisées et non-voisées. La figure suivante donne une représentation plus fine de tranches de signaux voisés et non voisés. L'évolution temporelle ne fournit cependant pas directement les traits acoustiques du signal. Il est nécessaire, pour les obtenir, de mener à bien un ensemble de calculs ad-hoc.

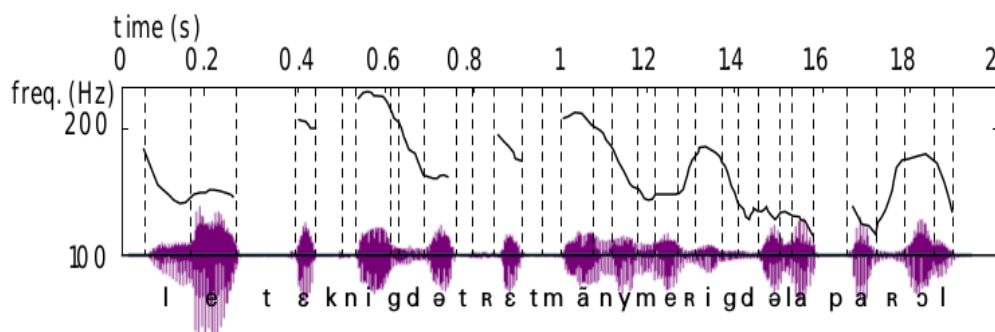


Figure 4 Fréquence fondamentale

### 3. Synthèse de la parole

#### 3.1. Concepts de base

Un système de synthèse à partir du texte (TTS : Text-To-Speech) est une machine capable de lire a priori n'importe quel texte à voix haute, que ce texte ait été directement introduit par un opérateur sur un clavier alpha-numérique, qu'il ait été scanné et reconnu par un système de reconnaissance optique des caractères (OCR : Optical Character Recognition),

#### 3.2. Applications

**Communication homme-machine**, multimédia. A plus long terme, le développement de synthétiseurs de haute qualité (ainsi que la mise au point de reconnaissanceurs fiables et robustes) permettra à l'homme de communiquer avec la machine de manière plus naturelle. L'explosion récente du marché du multimédia prouve bien l'intérêt du grand public en la matière.

**Recherche fondamentale et appliquée.** Enfin, les synthétiseurs possèdent aux yeux des phonéticiens une qualité qui nous fait défaut : ils peuvent répéter deux fois exactement la même chose. Ils sont par conséquent utiles pour la validation de théories relatives à la production, à la perception, ou à la compréhension de la parole.

**Monitoring vocal.** Dans certains cas, l'information orale est plus efficace qu'un message écrit. L'utilisation d'une voix de synthèse dans un centre de contrôle de site industriel, par exemple, permet d'attirer l'attention du personnel de surveillance sur un problème urgent.

#### 3.3. Synthèse par règles - Synthèse par concaténation

##### 3.3.1. Synthèse par règles

Les synthétiseurs par règles ont principalement la faveur des phonéticiens et des phonologistes. Ils permettent une approche cognitive, générative du mécanisme de la phonation. Ils sont basés sur l'idée que, si un phonéticien expérimenté est capable de «lire» un spectrogramme, il doit lui être possible de produire des règles permettant de créer un spectrogramme artificiel pour une suite de phonèmes donnés. Une fois le spectrogramme obtenu, il ne reste plus alors qu'à générer l'audiogramme correspondant.

##### 3.3.2. Synthèse par concaténation

Au contraire des synthétiseurs par règles, les synthétiseurs par concaténation ont une connaissance très limitée du signal qu'ils mettent en forme. La plupart de ces connaissances se trouve en effet stockée dans les unités de paroles mises en œuvre par le synthétiseur.

### 4. Conclusion

L'étude du langage naturel et des mécanismes nécessaires à la mise en œuvre à son traitement automatique par des machines est un domaine d'études foisonnant, et riche en applications potentielles ou émergentes. De nombreux progrès restent à accomplir pour mieux

comprendre cette faculté et pour bâtir des systèmes capables de soutenir la comparaison avec l'humain, mais l'état des connaissances en permet aujourd'hui de proposer de nombreuses solutions efficaces à des problèmes et des demandes réels.