



**Version: 1.0.0**



# KREYOSPEECH :

**SYNTHÈSE VOCALE DU CRÉOLE  
HAITIEN**

*-Vers des perspectives nouvelles*

Avril 2007

---



---

<b>Avant- propos.....</b>	<b>4</b>
<b>Première partie</b>	
<b>Présentation du créole, en tant qu’outil linguistique</b>	
<b>Chap. 1 Le créole: outil linguistique.....</b>	<b>7</b>
1.1 Définition.....	7
1.2 Origine du créole.....	7
1.3 La formation des langues créoles.....	8
1.4 Distribution géographique du créole .....	10
1.4.1 Les créoles des Antilles.....	10
1.4.2 Les créoles d'Afrique.....	10
1.4.3 Les créoles du SUD-EST asiatique et du Pacifique.....	10
1.5 Les créoles français.....	11
1.6 Le statut officiel des créoles.....	11
<b>Chap.2. Le créole haïtien.....</b>	<b>13</b>
2.1 Introduction.....	13
2.2 Définition et origine.....	13
2.3 Le créole introduit officiellement à l'école.....	14
2.4 Eléments constitutifs du créole haïtien.....	14
<b>Chap.3. Éléments de phonétique .....</b>	<b>17</b>
3.1 Introduction.....	17
3.2 Le son et l'audition humaine.....	17
3.3 La voix humaine.....	17
3.4 Phonétique.....	17
3.5 Phonologie et syntaxe du créole haïtien.....	18
3.5.1 Rencontre de deux voyelles.....	19
3.5.2 Sons Marginaux.....	19
3.6 Tableau phonétique du créole haïtien.....	20
3.6.1 Voyelles nasales.....	20

3.6.2 Les semi-consonnes.....	20
3.7 La graphie du créole haïtien.....	21

## Deuxième partie

<b>Chap. 4. Cadre méthodologique.....</b>	<b>25</b>
---	-----------

### Introduction

4.1 Les synthétiseurs de parole.....	25
4.2 La synthèse à canaux.....	26
4.2.1 Les paramètres du synthétiseur.....	26
4.2.2 Méthode utilisée.....	26
4.2.3 Fonctionnement de l'icophone.....	26
4.2.4 Fonctionnement de l'icolog.....	27
4.2.5.1 Avantage de cette technique.....	27
4.3 La synthèse par formant.....	28
4.3.1 Réalisation d'un synthétiseur à formant.....	29
4.3.1.1 Montage en série.....	29
4.3.1.1.1 Désavantage de ce montage.....	29
4.3.1.1.2 Comment l'éviter ?.....	29
4.3.1.1.3 Représentation des formants en série.....	30
4.3.1.2 Montage en parallèle.....	30
4.3.1.2.1 Avantage du montage en parallèle.....	30
4.3.1.2.2 Inconvénient du montage en parallèle.....	30
4.4 La synthèse raisonnante, dans le domaine temporel.....	31
4.5 La synthèse à prédiction linéaire.....	31
4.5.1 Techniques de sondage.....	32
4.5.2 Fréquence d'échantillonnage.....	32
4.6 La synthèse par mots ou petits bout de phrases.....	35
4.7 Synthèse par élément phonétique.....	35

<b>Chap. 5. Les circuits de parole du marché.....</b>	<b>37</b>
---	-----------

5.1 Les circuits de synthèse.....	38
-----------------------------------	----

<b>Chap. 6. Comparaison entre langages naturels et langages artificiels.....</b>	<b>42</b>
--	-----------

6.1 Introduction.....	42
-----------------------	----

6.2 Quelques critères de compréhension des langues naturelles.....	44
--	----

6.3 Aspects spécifiques de la compréhension du langage naturel parlé.....	45
6.4 Les premières tentatives de traduction automatique des langues	
Naturelles.....	45
6.5 Des techniques orientées automates.....	46

## Troisième partie

### Traduction de texte en son

Chap. 7. La conversion de texte en parole.....	50
Chap. 8. Analyse des systèmes existant.....	51

## Quatrième partie

### Présentation du logiciel ***KREYOSPEECH***

Chap. 9. Vue d'ensemble de l'application.....	54
9.1 Problématique.....	54
9.2 Objectifs.....	54
9.3 Limites.....	55
9.4 Digramme de conception du logiciel.....	55
9.5 Principales interfaces (écrans) de l'application.....	55

## Cinquième partie

A- Conclusion et perspectives d'avenir.....	57
B- Petit lexique, <i>pour une compréhension facilitée du document</i> .....	58
C- Bibliographie.....	59

## Avant- propos

Le créole est aujourd'hui considéré en Haïti comme la langue parlée par 97% de la population. Elle est en outre la seule et unique langue parlée, et vraiment comprise de bon nombre d'étudiants même à l'université. Ainsi, des efforts ont été déployés pour enseigner le créole dans les établissements scolaires, afin de permettre à plus d'un de mieux apprécier la richesse de la langue, et par-dessus tout d'avoir une plus large compréhension des cours qui, jusque-là ne se faisaient qu'en français - un français souvent mal exprimé.

La nécessité d'apprendre en créole se fait surtout sentir dans les nombreux programmes d'alphabétisation mis en place par la Secrétairerie d'Etat à l'alphabétisation.

Le créole en tant qu'outil d'apprentissage sera mieux compris à la lumière d'une enquête menée par Louis-Auguste JOINT de l'Institut de Recherche pour le Développement- Université des Antilles et de la Guyane, auprès de 852 élèves dans trois écoles secondaires (Saint Martial, Saint Louis de Bourdon et Juvénat du Sacré Coeur), réputées d'être des écoles d'élites à Port-au-prince, la capitale d'Haïti. L'étude révèle que, selon l'établissement, seulement 3% à 8% des élèves utilisent les deux langues pour communiquer en classe, à la maison ou pour écrire une lettre. La préférence est accordée au français par rapport au créole dans ces écoles privées congréganistes. Le français demeure la langue de l'écriture. Dans les trois établissements, presque la totalité des élèves (96,1% à 99,6%) l'utilisent pour écrire une lettre. Dans la majorité des cas, c'est aussi la langue de la transmission orale des connaissances scolaires : 71,8% des élèves de 7<sup>ème</sup> fondamentale de Saint Louis de Bourdon et 64,6% de ceux de Saint Martial l'utilisent dans ce cas; Mais au Juvénat, les élèves utilisent plus le créole (51,4%) que le français (44%) en classe. Cela correspond aux positions sociales des populations qui fréquentent chacun des établissements: plus les élèves sont issus des catégories sociales supérieures, plus ils utilisent le français dans les différentes sphères de leur vie sociale. (Réf. Louis-Auguste JOINT *Le Bilinguisme Français / Créole dans l'enseignement Haïtien Quel enjeu pour le développement social?*)

Il devient donc important que des efforts soient déployés dans le but de permettre une plus large vulgarisation (d'un point de vue linguistique) des technologies de l'information (TICs).

Considérant les limites d'outils de recherche en créole Haïtien, et tout l'intérêt que pourrait représenter un outil technologique élaboré en créole, Linguistes, chercheurs, développeurs ou autres pourraient être intéressés par un tel projet dans le but d'en exploiter au bénéfice d'un programme relevant de leurs domaines de compétences respectives.

Aussi devient-il important de mettre à leur disposition ce programme de synthèse présenté en créole haïtien- le premier du genre- doté de fonctionnalités permettant de l'étendre au profit d'autres réalisations...

AGENOR JEAN BERNARD  
BERNARD VENET  
CLERGER MARTIN LUTHER

## Première partie

### **Présentation du créole, en tant qu'outil linguistique**

Cette partie consistera en une étude détaillée du créole, et se penchera sur la phonologie et la phonétique du créole haïtien.

# Chap.1 *Qu'est-ce que le créole*

---

## **Aperçu**

- 1.1 Définition
- 1.2 Origine du créole
- 1.3 La formation des langues créoles
- 1.4 Distribution géographique du créole
- 1.5 Les créoles français
- 1.6 Les créoles anglais
- 1.7 Les créoles portugais
- 1.8 Le statut officiel des créoles

### 1.1 Définition

On appelle créole en linguistique, une série de langues mixtes nées, dans les colonies, du contact du blanc parlant une langue européenne avec l'indigène ou l'esclave importé.

### 1.2 Origine du créole

Le terme créole possède deux étymologies, l'une portugaise (crioulo), l'autre, espagnole (criollo), qui viennent du même mot latin criare, signifiant «nourrir» ou plus précisément «serviteur nourri dans la maison». Un «Créole» désigne d'abord quelqu'un qui a été «élevé sur place», c'est-à-dire «qui est du pays». Le mot a servi tout d'abord à désigner l'enfant blanc né et élevé dans les colonies européennes ou d'outre-mer: Haïti, Martinique, Guyane française, Réunion, Louisiane, etc.

Par la suite, le mot créole a été employé pour désigner la population noire; on parle alors de «Créoles de couleur» — et, par voie de conséquence, la langue de cette population, le créole. Le mot a même été étendu aux animaux et aux objets : les vaches, les poules et le café pouvaient être créoles, à la condition qu'ils proviennent des colonies. Être créole, c'était donc, avant tout, naître sur les terres des colonisateurs. Le mot créole a longtemps été utilisé en ce sens en Louisiane et l'est encore aujourd'hui: les «Créoles blancs», généralement des membres de riches familles propriétaires de plantations. Ce terme s'opposait à la notion d'«étranger à la culture locale». Évidemment, aujourd'hui,



le substantif créole est employé, avec des sens divers, pour des personnes, mais il s'applique surtout à des langues mixtes.

### 1.3 La formation des langues créoles

Le créole est un phénomène linguistique d'importance considérable. Un créole se forme au contact des langues. Tout créole est essentiellement le résultat du mixage de langues différentes. Il existe plusieurs créoles, mais ce sont tous des langues mixtes qui se sont formées aux XVI<sup>e</sup> et XVII<sup>e</sup> siècles à la faveur de la traite des Noirs organisée par les puissances coloniales de l'époque, particulièrement la Grande-Bretagne, la France, le Portugal, l'Espagne, les Pays-Bas et, plus tard, les États-Unis. Mais d'autres pays furent également en cause.

L'origine de la formation des créoles est relativement simple, bien que, dans ce domaine comme dans d'autres, nous n'ayons pas de réponses absolument sûres. Les créoles ne se sont pas formés dès l'apparition de l'esclavage, mais un peu plus tard, lorsque les plantations se développèrent et firent appel à une main-d'oeuvre massive. Les premiers esclaves, alors encore peu nombreux, durent apprendre, en une année tout au plus, le français, le portugais, l'espagnol ou l'anglais, selon le cas. Cette langue que les esclaves apprenaient ne devait pas être rigoureusement identique à celle de la Métropole, car les colons eux-mêmes n'employaient guère la langue de la cour dont ils dépendaient. De plus, il est certain que les Noirs apportèrent d'importantes modifications à la langue de leurs maîtres.

Puis le développement rapide de l'économie de plantation bouleversa tout. L'importation massive d'esclaves venus de plusieurs régions de l'Afrique changea la donne. Les nouveaux arrivants se trouvèrent dans l'impossibilité de communiquer entre eux dans leur langue maternelle. D'une part, les règlements en vigueur dans les entrepôts d'embarquement et dans les plantations interdisaient le recours aux langues des Noirs, d'autre part, les «voisins de labeur» parlaient rarement la même langue, car ils provenaient de régions ou d'ethnies différentes. Dans ces conditions, les esclaves durent utiliser la langue du maître, mais ils n'y avaient qu'un accès déformé et limité. Ils n'apprenaient que des mots transmis par des ordres donnés par des Blancs ou des esclaves arrivés les premiers et auxquels les colons avaient délégué des responsabilités. Lentement, d'approximations en interprétations et en appropriations de stratégies qui transformèrent la langue de base, la langue des Blancs subit de tels changements qu'elle se transforma en une autre pour devenir un créole.

Bien qu'ils n'aient pas été en mesure d'apprendre la langue des maîtres blancs (négriers ou planteurs), les esclaves s'approprièrent en général les mots et les sons de ceux-ci tout en recourant à la grammaire et aux notions qu'ils connaissaient dans leurs langues d'origine. Les créoles sont donc dérivés à la fois des langues des maîtres et des langues africaines (en général) parlées par les travailleurs asservis. Les créoles sont donc des langues, au sens linguistique du terme: ils servent à la communication au même titre que le français ou l'anglais.

Nous trouvons surtout des créoles à base française, anglaise, portugaise, espagnole et néerlandaise. Cependant, dans un recensement effectué en 1977, Ian Hancock, docteur en linguistique à l'Université du Texas, dénombrerait 127 créoles différents dans le monde:

35 à base d'anglais	6 à base d'allemand
15 à base de français	1 à base de slave
14 à base de portugais	6 à base amérindienne
7 à base d'espagnol	21 à base africaine
5 à base de néerlandais	10 à base non indo-européenne (asiatique)

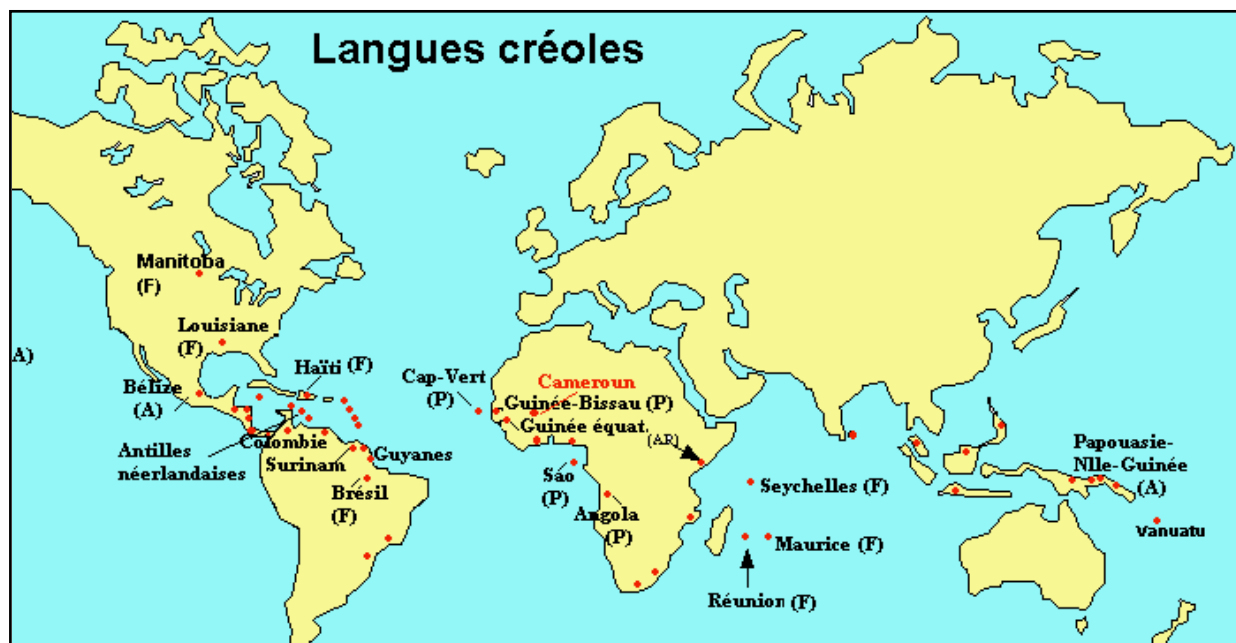
**(Réf. [www.tlfq.ulaval.ca/axl/amsusudant/creole.htm](http://www.tlfq.ulaval.ca/axl/amsusudant/creole.htm))**

Il existe aussi quelques créoles à plusieurs bases linguistiques. Le cas le plus connu est celui du papiamentu en usage dans les Antilles néerlandaises, surtout dans les îles d'Aruba, de Bonaire et de Curaçao, au nord du Venezuela. En effet, ce créole parlé par au moins 200 000 locuteurs a comme base première le portugais, puis le néerlandais, l'espagnol, ainsi qu'un peu d'anglais et de français.

Rappelons-le, du point de vue linguistique, les créoles sont des langues à part entière, dont la structure grammaticale est proche de celle des langues africaines et dont le lexique est en très grande majorité d'origine européenne. Cependant, le statut des langues créoles dans le monde est généralement inférieur au plan social, culturel et politique.

## 1.4 Distribution géographique du créole

Ces créoles se retrouvent dans de nombreuses parties du monde, à l'exception l'Europe.



\* Source: [www.tfq.ulaval.ca/axl/amsudant/creole.htm](http://www.tfq.ulaval.ca/axl/amsudant/creole.htm)

### 1.4.1 Les créoles des Antilles

La population créolophone la plus importante se trouve aux Antilles et dans les Guyanes: Haïti, Anguilla, Antigua, les Antilles néerlandaises (Aruba, Bonaire et Curaçao), les Bahamas, la Barbade, le Belize, la Grenade, la Jamaïque, la Guadeloupe, la Martinique, la Dominique, Sainte-Lucie, Saint-Christophe-et-Niévès (Kitts-et-Nevis), Saint Vincent et Grenadines, la Trinité-et-Tobago, le Suriname, la Guyana et la Guyane française. Il ne faudrait pas oublier le créole brésilien (à base de portugais).

### 1.4.2 Les créoles d'Afrique

Un second groupe de créolophones est situé en Afrique: les îles du Cap-Vert, la Guinée-Bissau, le Sierra Leone, les îles Saõ Tomé et Príncipe, la République centrafricaine, puis dans l'océan Indien, les Comores, les Seychelles, l'île Maurice et l'île de La Réunion.

### *1.4.3 Les créoles du SUD-EST asiatique et du Pacifique*

Le troisième groupe est localisé en Asie du Sud-est et en Océanie: les Philippines, Singapour, la Papouasie- Nouvelle-Guinée, le Vanuatu et la Nouvelle-Calédonie. Dans un grand nombre de cas, les créolophones correspondent à de petites communautés de quelques milliers de personnes chacune, parfois de quelques centaines. Les pays où l'on compte le plus de locuteurs sont: Haïti (Huit millions) et l'île Maurice (un million).

### 1.5 Les créoles français

Dans les anciennes colonies françaises (Haïti, Martinique, Guadeloupe, Sainte-Lucie, la Dominique, la Guyane française, La Réunion, les Seychelles, l'île Maurice, etc.), la base lexicale du créole s'est élaborée à partir du français. Aujourd'hui, on distingue le créole martiniquais, le créole guadeloupéen, le créole haïtien, le créole dominicain, le créole Saint-lucien, le créole réunionnais, le créole guyanais, le créole seychellois, le créole mauricien, etc.

La population créolophone à base française est estimée à environ 9,5 millions de locuteurs, dont huit millions en Haïti, environ un million à l'île Maurice, 600 000 à la Réunion, 380 000 à la Martinique, 425 000 à la Guadeloupe, 70 000 aux Seychelles, etc.

L'intercompréhension entre les créoles français est relativement aisée. Par exemple, un créolophone de la Martinique (Antilles) et un autre de La Réunion (océan Indien) peuvent se parler en utilisant chacun son créole sans trop de problème de communication; il reste que l'accent, l'intonation, un nombre plus ou moins important de termes inconnus, de même que certains éléments grammaticaux et des tournures syntaxiques, peuvent néanmoins entraver la compréhension, surtout lorsque les créolophones sont moins instruits. En raison de la proximité géographique, le créole martiniquais et le créole guadeloupéen présentent de nombreuses similitudes. Évidemment, un créolophone à base de français voit sa marge de compréhension rétrécir considérablement s'il parle à un créolophone à base d'anglais; la compréhension risque de se limiter à des messages très simples.

### 1.6 Le statut officiel des créoles

Parmi les rares États où l'on a reconnu officiellement un créole dans un texte constitutionnel, mentionnons Haïti (8 millions) avec le français au même titre que le créole (du moins juridiquement); l'archipel des Seychelles (70 000 habitants) avec l'anglais, le français et le créole comme langues co-officielles; le Vanuatu en Mélanésie (145 000 habitants), qui a reconnu son créole mélanésien, le bichlamar, comme sa «langue officielle parlée», avec l'anglais et le français. Toutefois, il ne faut pas

se faire trop d'illusion, car la reconnaissance officielle du créole par un État ne supprime pas nécessairement les préjugés dévalorisants.

Par ailleurs, le gouvernement des Antilles néerlandaises et celui d'Aruba projettent de présenter une loi qui ferait du papiamento et de l'anglais des langues officielles au même titre que le néerlandais. Le projet de loi n'a pas encore été adopté, mais les Antilles néerlandaises bénéficieraient ainsi de trois langues officielles. À l'instar du Vanuatu, le papiamento des Antilles néerlandaises deviendrait probablement la «langue officielle parlée».

Peu de créoles sont écrits, même lorsqu'ils bénéficient d'un statut de co-officialité comme en Haïti; quoi qu'il en soit, les créoles écrits sont submergés par la langue coloniale. Néanmoins, beaucoup d'associations de défense du créole prennent la relève des anciens conteurs oraux, afin de le retranscrire et lui donner une identité. Dans certains créoles, il existe des grammaires, des dictionnaires et une littérature.

# Chap.2 *Le créole haïtien*

---

## **Aperçu**

- 2.1 Introduction
- 2.2 Définition et origine
- 2.3 Le créole introduit officiellement à l'école
- 2.4 Eléments constitutifs du créole haïtien
  - 2.4.1 L'indien
  - 2.4.2 L'espagnol
  - 2.4.3 L'africain
  - 2.4.4 Le français
  - 2.4.5 L'anglais

### 2.1 Introduction

Le créole haïtien est un créole à base lexicale française. Cela veut dire que la majorité des mots du lexique du créole haïtien proviennent du français, contrairement à d'autres créoles dont les mots proviennent d'autres langues européennes, telles que l'anglais (créole jamaïcain) ou le portugais (le papiamentu). Il est important de souligner ici que, malgré la forte prédominance de mots directement empruntés du français, le créole haïtien fonctionne structurellement de manière autonome. La structure des phrases du créole haïtien ne révèle aucune dépendance particulière à l'égard de la syntaxe française. Le créole haïtien, comme toute langue naturelle, se présente d'abord comme un ensemble de règles, de principes et de contraintes implantés dans les cerveaux de ses locuteurs. Dans ce sens, le créole haïtien n'est absolument pas différent de l'anglais ou du chinois ou du Swahili. C'est qu'il existe des propriétés grammaticales et des principes fondamentaux pour la construction des phrases, communs à toutes les langues, faisant partie de notre héritage génétique et implantés en nous dès notre naissance. Le linguiste américain Noam Chomsky a désigné ces principes fondamentaux sous le nom de grammaire universelle. A l'intérieur de ce regroupement fondamental, chaque langue humaine possède certaines caractéristiques particulières. Sur les plans phonologique, sémantique et morphosyntaxique, le créole haïtien (CH) dégage donc certaines particularités.

## 2.2 Définition et origine

Le créole haïtien est la langue populaire parlée et familière de la République d'Haïti. L'homme du peuple des villes parle généralement le créole; cependant comme il a fréquenté l'école primaire de longues années et qu'il est en contact perpétuel avec des gens parlant tantôt créole, tantôt français, il peut s'exprimer en français à l'occasion.

Le paysan lui, ne parle pas français ou, s'il est un noble de son village, parle un français mutilé; l'enfant de 12-14 ans qui se rend encore à l'école possède un petit vocabulaire scolaire qu'il oubliera plus tard; pour lui comme pour ses parents le créole constituera un jour l'unique moyen d'expression.

D'après Suzanne Sylvain, le créole haïtien est probablement né dans l'île de la Tortue au cours du 17<sup>ème</sup> siècle, du jour où un esclave nègre s'est avisé, pour se faire comprendre du flibustier français, son maître, d'essayer de lui parler dans sa langue. Il y a eu un effort d'ajustement des deux côtés, ce qui explique la formation du vocabulaire haïtien et l'extraordinaire fortune de certains mots.

## 2.3 Le créole introduit officiellement à l'école

Il est important de souligner que le créole a été introduit officiellement à l'école le 18 Septembre 1979, par décret présidentiel. Depuis lors il est le moyen par lequel on apprend à l'enfant la rigueur et sert de principal outil linguistique d'alphabétisation jusqu'au aujourd'hui.

## 2.4 Eléments constitutifs du créole haïtien

Nous allons, dans cette partie, considérer certaines raisons qui ont contribué à enrichir le vocabulaire du créole haïtien.

### **2.4.1 L'indien**

Lorsqu'en 1492 débarquèrent les espagnols, l'île d'Haïti était occupée par deux tribus indiennes : Les Tainos et les Caraïbes. Les Tainos succédaient à des Aborigènes. Bien que réduits en esclaves et exterminés en grande partie par les espagnols, les derniers indiens s'étaient vus reléguer dans le territoire de Boya.

Cependant on ne peut leur dénier une certaine influence dans la formation du créole. Influence indirecte sur le vocabulaire grâce aux mots transmis par l'intermédiaire du Français ou de l'Espagnol.

### **2.4.2 L'espagnol**

Certains mots espagnols sont entrés dans le vocabulaire de notre créole dès ses premiers balbutiements, mais la plupart d'entre eux sont d'emprunts plus récents. Plusieurs datent de

l'occupation de l'île Saint-Domingue par la République d'Haïti (1822-1844), d'autres de l'établissement en Haïti de nombreux cubains. D'autres enfin de l'émigration annuelle des ouvriers agricoles haïtiens vers Cuba ou la République Dominicaine ou encore des rapports des populations des frontalières haïtiennes ou dominicaines.

#### **2.4.3 L'africain**

Les dialectes africains se sont très peu perpétués en Haïti. Cependant jusqu'à nos jours certaines chansons vodouesques se sont conservées tellement pures que l'une d'entre elles a pu être comprise au Dahomey par une société secrète.

#### **2.4.4 Le français**

Au 18<sup>eme</sup> siècle, lors de la grande postérité de Saint-Domingue, les provinces fournirent leurs petits contingents, si bien qu'on retrouve des mots provençaux ou gascons à côté des termes auvergnats, lorrains ou picards.

#### **2.4.5 L'anglais**

Il y avait des anglais à la tortue, parmi les flibustiers. D'autre part certaines régions ont été plus ou moins longtemps occupées par les Anglais au 18<sup>eme</sup> siècle.

Une grande partie des immigrants sous le président Géffrard étaient Jamaïcains ou américains. Enfin, le commerce d'exportation a eu lieu principalement avec les Etats-Unis.



# Chap.3 *Eléments de phonétique et de phonologie*

---

## **Aperçu**

- 3.1 Introduction
- 3.2 Le son et l'audition humaine
- 3.3 La voix humaine
- 3.4 Phonétique et phonologie
- 3.5 Morphologie et syntaxe du créole haïtien
  - 3.5.1 Introduction
  - 3.5.2 Rencontre de deux voyelles
  - 3.5.3 Sons marginaux
- 3.6 L'alphabet phonétique du créole haïtien
- 3.7 La graphie du créole haïtien

### 3.1 Introduction

Grammaticalement, la phonétique se définit comme étant la partie de la grammaire qui étudie l'ensemble de sons et des articulations propres à une langue.

En linguistique, la phonétique trouve une définition beaucoup plus large. Ainsi distingue-t-on deux principaux types de phonétiques. La *phonétique historique* qui compare les différents idiomes et suit leur évolution, et la phonétique expérimentale qui use d'instruments enregistreurs pour fixer les mouvements des organes de la voix.

La loi dominante de la phonétique est celle de la constance des lois phonétiques : à une même époque, dans un même dialecte, le même phonème, placé dans des conditions analogues varie de la même façon. La cause principale de cette uniformité de variation est dans l'immigration du langage des hommes entre eux. C'est la rigueur acquise par la phonétique qui a donné sa solidité à l'étymologie.

### 3.2 Le son et l'audition humaine

Le son est une onde qui se propage dans un milieu matériel sous la forme de faibles variations de pression. Il est perceptible à l'oreille humaine lorsque sa fréquence se situe globalement entre 20 Hz

et 20 kHz au maximum. Toutefois, on estime que l'information phonétique se trouve en dessous de 10 kHz. Dans le cas d'une communication téléphonique, les fréquences au-dessus de 3,5 kHz sont coupées, mais la conversation reste intelligible, certains phonèmes très aigus, comme [S], étant toutefois mal rendus. Une caractéristique assez importante de l'audition humaine, connue sous le nom de « loi d'Ohm » dans le domaine de l'acoustique, est que notre oreille n'est pas sensible à la phase du signal.

Il est admis en moyenne que la fréquence dite fondamentale ; celle qui fixe la hauteur de la voix, est de 120 Hz pour un homme et de 220 Hz pour une femme et encore plus pour les enfants considérant que sa voix est plus aigue.

(Réf. MARC, Ferretti, FRANÇOIS, Cinare, *SYNTHESE, RECONNAISSANCE de la PAROLE, Le conduit vocale*, Paris, éditeurs, 1983, p.82§1).

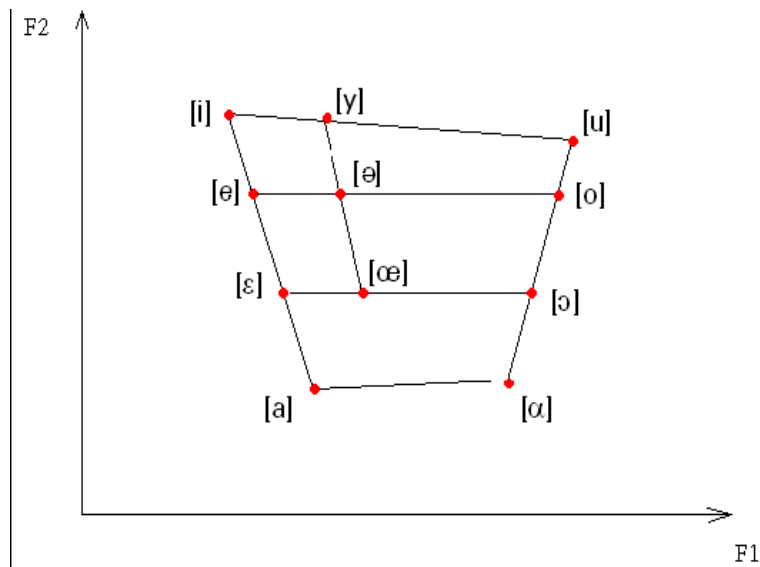
### 3.3 La voix humaine

La voix humaine résulte de l'interaction entre l'air expiré et les divers organes phonatoires sur son parcours. Dans le cas de phonèmes voisés, le son est initialement produit par la vibration des cordes vocales. Il est ensuite traité de manière différente suivant les cavités par lesquelles il passe à savoir le pharynx et la bouche principalement. Ces cavités agissent comme des résonateurs qui renforcent pour certains phonèmes les fréquences correspondantes à leur fréquence de résonance. Ces fréquences renforcées sont appelées formants, et ce sont eux que les phonologistes essaient de repérer dans un spectrogramme pour reconnaître les phonèmes prononcés. Il est intéressant de noter que la cavité nasale, elle, atténue les formants des phonèmes dits nasalisés, et c'est pourquoi parle-t-on de phénomène d'anti-résonance. Ces organes phonatoires n'étant pas figés, leur forme varie en fonction du mode d'articulation propre à un phonème particulier, et leurs fréquences de résonance aussi, d'où l'apparition de formants différents pour chaque phonème. En pratique, on considère les trois ou quatre premiers formants, les deux premiers étant les plus liés au mode d'articulation.

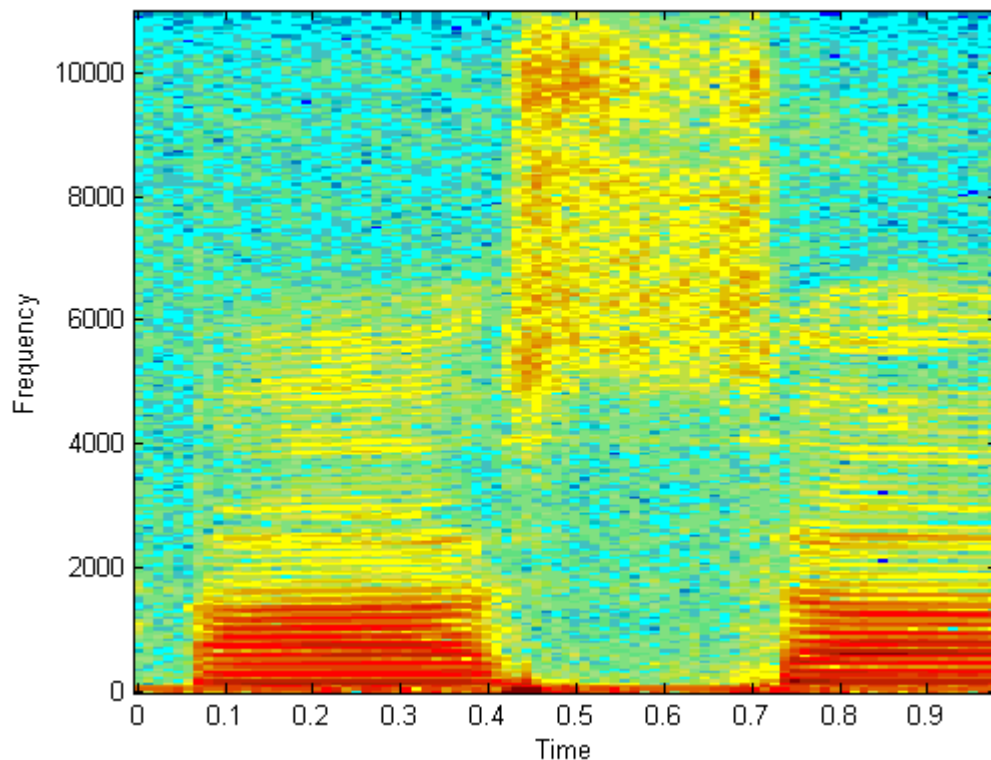
### 3.4 Phonétique

La phonétique est l'étude des phonèmes, ces derniers étant souvent définis comme « la plus petite unité phonique distinctive ». Par exemple, l'opposition des mots créoles « *pen* » et « *ben* » montre que [p] et [b] sont des phonèmes. On peut les classer en plusieurs classes et sous-classes, le premier niveau étant celui des voyelles et des consonnes. Les voyelles sont des sons qui ont une durée assez longue, caractérisés par un libre écoulement et une grande constance de leurs propriétés fréquentielles dans le temps: en l'absence d'éléments prosodiques trop marqués, les formants sont horizontaux sur un spectrogramme. On classe souvent les voyelles au moyen du trapèze vocalique, reproduit ci-dessous, qui représente leur position dans le plan défini par les deux

premiers formants et traduit la position de la langue lors de l'articulation (voyelles postérieures ou antérieures, ouvertes ou fermées):



Les consonnes sont des phonèmes qui, lors de leur articulation, rencontrent un obstacle. Elles sont généralement beaucoup plus courtes que les voyelles et beaucoup plus variables dans le temps. Elles peuvent être bruyantes ou sonantes. Ce n'est que dans ce dernier cas qu'elles présentent des formants. La figure ci-dessous représente le spectrogramme de la syllabe [asa] et illustre quelques-uns des concepts formulés précédemment. On perçoit très clairement l'horizontalité des formants de la voyelle [a]. La consonne [s] est un bruit très aigu situé au-dessus de 5 kHz, clairement exempt de formants.



### 3.5 Phonologie et syntaxe du créole haïtien

Phonologiquement, le créole haïtien se distingue du français par la nasalisation des voyelles orales / i /, / u / devant une consonne nasale, comme dans "moun" (gens), "youn" (un, une), "larim" (mucus), ou l'existence de variantes syllabiques formées par une consonne unique avec les nasales / m / et / n /, comme dans la phrase "m te wè li" (je l'ai vu). D'une manière générale, le phénomène de la nasalisation semble jouer un rôle déterminant dans la phonologie du créole haïtien.

Telle qu'elle apparaît maintenant, l'étude sémantique du créole haïtien se manifeste surtout à travers la fabrication de dictionnaires bilingues (créole anglais et anglais-créole). Comparativement, il existe très peu de dictionnaires bilingues créole-français et français-créole. Pourtant, avec l'établissement aux Etats-Unis de programmes bilingues scolaires anglais-créole dans les communautés à forte population d'élèves haïtiens (NY, Mass., Florida) où la langue créole haïtienne est utilisée comme langue-objet.

#### **3.5.1 Rencontre de deux voyelles**

En créole, deux voyelles ont du mal à se faire suivre. En général pour éviter cette rencontre, le locuteur introduit un autre graphème soit le "y", soit le "w".

Par conséquent on ne dit pas :

*Koulyea*

Mais

*Koulyeya*

*Mai*

Mais

*Mayi*

<i>Pei</i>	“	<i>Peyi</i>
<i>Kaotchou</i>	“	<i>Kawotchou</i>
<i>Aeopò</i>	“	<i>Ayewopò</i>

### 3.5.2 Sons Marginaux

Certains sons sont marginaux parce qu'ils ne sont prononcés que par une très faible partie de la population. On peut les transcrire au moyen des graphèmes suivants : u, eu, èu, eun.

Prenons quelques exemples :

- ❖ *Vann mwen on goud suk*
- ❖ *Meulon an vèt*
- ❖ *Bèu a rans*
- ❖ *Leundi ak madi pa menm*

On trouve dans ce même groupe : ue, uè, ua.

Exemples :

- ❖ *Sue, suè, nuaj.*
- ❖ *Li Sue tankou pitit bouki.*
- ❖ *Tout do chemiz la se mak Sue.*
- ❖ *Kite m fè on nuaj sigarèt.*

### 3.6 Tableau phonétique du créole haïtien

Ci-dessous sont représentés les différents phonèmes et graphèmes qui constituent L'alphabet phonétique du créole haïtien.

<u>Voyelles orales</u> - vwayèl bouch				
<b>E</b>	<b>I</b>	<b>o</b>	<b>ò</b>	<b>ou</b>
[ɛ]	[i]	[o]	[ɔ]	[u]

<u>Consonnes</u> – konsòn								
<b>B</b>	<b>Ch</b>	<b>D</b>	<b>F</b>	<b>G</b>	<b>H</b>	<b>j</b>	<b>k</b>	<b>l</b>

[b]	[f ]	[d]	[f]	[g]	[h]	[ʒ]	[k]	[l]
<b>M</b>	<b>N</b>	<b>Ng</b>	<b>P</b>	<b>R</b>	<b>S</b>	<b>t</b>	<b>v</b>	<b>z</b>
[m]	[n]	[ŋ]	[p]	[ʁ], [w]	[s]	[t]	[v]	[z]

<u>Voyelles nasales</u> - vwayèl bouch-nen		
En	On	oun
[ɛ̃]	[ɔ̃]	[ũ]

### 3.6.1 Voyelles nasales

Quand une voyelle orale devient nasale, on la coiffe d'un petit chapeau en phonétique. En orthographe créole, on la fait accompagner d'un "n".

### 3.6.2 Les semi-consonnes

Voici les principaux semi-consonnes : W, y, ui.

Nous allons donner quelques exemples montrant leur utilisation.

#### 1 W

<i>Graphèmes</i>	<i>Exemples</i>
Wa	watè, lakwa
Wan	wanga, lakwann
We	wete, dwe
Wen	wench, kwen
Wi	witi, kwi
Wo	wotè, gwo
Wò	wòl, bwòs
Won	wont, twont
Wou	woule, kwout

#### 2) y

<i>Graphèmes</i>	<i>Exemples</i>
Ya	gaya, pyas
Yan	yanvalou, pyan
Ye	pie, katye

Yen	anyen
Yè	ayè
Yon	bouyon, pyon
Yo	koyo, konfyolo
You	pyout
yò	miyò

### 3) ui

*Graphèmes*

Ui

*Exemples*

zuit, nuit, luil, kuit, suif, uit

### 3.7 La graphie du créole haïtien

Voici les 32 graphèmes permettant d'écrire le créole haïtien moderne :

**a, an, b, ch, d, e, è, en, f, g, h, l, j, k, l, m, n, ng, o, ò, on, ou, oun, p, r, s, t, ui, v, w, y,z.**

## Deuxième partie

# Synthèse de la parole

### Introduction

#### Domaines du traitement automatique de la parole

La reconnaissance et la synthèse de la parole sont deux domaines du *traitement automatique de la parole*. Le fonctionnement d'un système de *reconnaissance de la parole* consiste à employer des techniques d'appariement afin de comparer une onde sonore à un ensemble d'échantillons, composés généralement de mots mais aussi, plus récemment, de phonèmes. En contrepartie, le système de *synthèse de la parole* permet la prononciation d'un texte écrit numérisé, qui est produite

par un générateur de phonèmes. Ces deux domaines font appel aux *connaissances de plusieurs sciences* : l'anatomie et 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, les statistiques, etc.

### **Perspectives et liens avec le langage naturel et les systèmes d'information**

Le traitement automatique de la parole ouvre des *perspectives* nouvelles compte tenu de la différence considérable qui existe entre la commande manuelle et vocale. L'utilisation du *langage naturel* dans le dialogue personne/machine met la technologie à la portée de tous et entraîne sa vulgarisation, en réduisant les contraintes de l'usage des claviers, souris et codes de commandes à maîtriser. En simplifiant le protocole de dialogue personne/machine, le traitement automatique de la parole vise donc aussi un gain de productivité puisque c'est la machine qui s'adapte à l'homme pour communiquer, et non l'inverse. De plus, il rend possible l'utilisation simultanée des yeux ou des mains à une autre tâche. Il permet d'humaniser les systèmes informatiques de *gestion de l'information*.

### **Principaux problèmes à résoudre et défis à relever**

- Le *mode d'élocution* : le discours continu (langage naturel) est beaucoup plus complexe à modéliser que le discours isolé (pause après chaque mot)
- La *dépendance au locuteur*: les systèmes multilocuteurs qui sont indépendants du locuteur sont beaucoup plus complexes à modéliser que les systèmes monolocuteurs qui nécessitent un apprentissage préalable
- l'*adaptation*: les systèmes les plus récents sont fondés sur la reconnaissance d'unités plus petites que le mot, correspondant le plus souvent à ce qu'on appelle un **triphone**, qui représente la forme acoustique d'un phonème dans le contexte de ses phonèmes voisins immédiats. Il est alors demandé à tout nouvel utilisateur de prononcer un ensemble de phrases comportant l'ensemble des phonèmes d'une langue, afin d'*adapter* les références à sa voix. Le défi est de réduire au maximum la durée d'une telle *adaptation*, et même d'éliminer la nécessité d'une phase d'*adaptation préalable* pour permettre d'obtenir des performances satisfaisantes dès la première utilisation.
- la *taille du vocabulaire* reconnu : on vise le plus grand nombre de mots mais, dans ce cas, il y a plus de risques d'erreurs d'identification et le temps de traitement est plus long. Un vocabulaire volumineux est constitué de quelques milliers de mots, un moyen de quelques centaines et un petit de quelques dizaines.
- le *bruit* environnant : son filtrage est essentiel mais difficile à faire pour une machine



- les *types de locuteurs* : prononciations différentes d'un seul ou de plusieurs locuteurs
- les *ambiguïtés sur les homonymes* : nécessité de reconnaître les contextes, par la syntaxe grammaticale et/ou sémantique, pour différencier des syllabes identiques ayant des prononciations différentes (ex: les poules du *couvent couvent*).

## Chap.4 *Cadre Méthodologique*

---

### **Aperçu**

#### 4.1 Les synthétiseurs de parole

#### 4.2 La synthèse à canaux

##### 4.2.1 Les paramètres du synthétiseur

##### 4.2.2 Méthode utilisée

##### 4.2.3 Fonctionnement de l'icophone

##### 4.2.4 Fonctionnement de l'icolog.

##### 4.2.5.1 Avantage de cette technique

#### 4.3 La synthèse par formant

##### 4.3.1 Réalisation d'un synthétiseur à formant

##### 4.3.1.1 Montage en série

##### 4.3.1.1.1 Désavantage de ce montage

##### 4.3.1.1.2 Comment l'éviter ?

##### 4.3.1.1.3 Représentation des formants en série

##### 4.3.1.2 Montage en parallèle

- 4.3.1.2.1 Avantage du montage en parallèle
- 4.3.1.2.2 Inconvénient du montage en parallèle
- 4.4 La synthèse résonante, dans le domaine temporel
- 4.5 La synthèse à prédiction linéaire
  - 4.5.1 Techniques de sondage
  - 4.5.2 Fréquence d'échantillonnage
- 4.6 La synthèse par mots ou petits bouts de phrases
- 4.7 Synthèse par élément phonétique
- 4.8 La synthèse articulatoire
- 4.9 Méthode choisie

## Introduction

La **synthèse vocale** est la technique de synthèse sonore qui vise à la création artificielle de paroles, via un système de traitement du signal. Elle peut s'inscrire dans un système d'interaction vocale.

Il y a eu plusieurs techniques de synthèse vocale jusqu'à nos jours.

La première, populaire entre 1965 et 1985, est appelée **synthèse vocale par règles**. Celle-ci est basée sur la modélisation de la parole à partir d'un spectre sonore. Des règles peuvent être écrites pour générer un spectre sonore artificiel. Cette technique permet des économies de mémoire considérables. La seconde technique, dite **par concaténation de diphones**, n'est pas purement artificielle. Les sons synthétisés sont en fait des segments d'enregistrement de parole artificiellement attachés les uns à la suite des autres. Cette technique peut être produite avec moins de 10 Méga octets de données sonores. La parole synthétisée semble plus naturelle que celle produite par règles, mais des problèmes persistent lorsque la longueur de phonème, l'intonation et l'accent tonique sont pris en compte. Pour remédier à ces problèmes, il est possible d'augmenter la quantité d'extrait sonore à utiliser pour la concaténation. Plusieurs extraits pour un même diphone peuvent être utilisés dans un même contexte, et plusieurs types de contextes (intonation, accent, type de phrase) peuvent avoir chacun leurs diphones. On parle alors de synthèse par **sélection de diphones** ou **sélection**

**d'unités** (Unit Selection). Ceux-ci sont ensuite utilisés lors de la synthèse vocale de façon à réduire les mauvaises transitions. Cette amélioration peut demander des bases de données de diphtonges de plusieurs Méga octets, voire de plusieurs Giga octets.

#### 4.1 Les synthétiseurs de parole

De manière générale, un synthétiseur c'est la possibilité qu'a une machine de produire la parole en puisant des trois sources suivantes :

- 1) des sources d'impulsion pour les sons voisés
- 2) d'une source de bruit pour les sons non voisés
- 3) d'un simulateur de conduit vocal.

On crée un « modèle » qui est supposé stable sur des intervalles de temps de l'ordre de la vingtaine de millisecondes et qui (logiquement) reçoit des paramètres de commandes toutes les 20 ms environ. Ce modèle peut être réalisé par plusieurs types de synthétiseurs. On parlera de :

- 1) la synthèse à canaux
- 2) la synthèse par formants
- 3) la synthèse résonante dans le domaine temporel
- 4) la synthèse à prédiction linéaire
- 5) la synthèse par élément phonétique
- 6) l'assemblage des éléments de parole synthétique
- 7) la synthèse par mots ou petits bouts de phrase
- 8) la synthèse articulatoire

#### 4.2 La synthèse à canaux

Le modèle qui est réalisé par un nombre plus ou moins important de filtres électriques dont les caractéristiques produisent les propriétés résonnantes du conduit vocal dont les valeurs de gain sont adaptées aux spectres de la parole est appelé «synthèse par canaux ».

Le synthétiseur à canaux est réparti sur la bande audio fréquence à l'aide de l'une ou l'autre des sources d'excitations produisant des impulsions ou bruit blanc, selon que les sons sont voisés ou non. On dose ensuite l'énergie dans chaque canal lors de l'analyse de la parole (que l'on souhaite faire ensuite dire par la machine). Et, on fait correspondre la fréquence des impulsions à celle de la mélodie.

##### **4.2.1 Les paramètres du synthétiseur**

Les paramètres de gouverne du synthétiseur pendant 20 ms (à réactualiser) sont :

- 1) la commande « voisée/non-voisée »

- 2) La fréquence du fondamental
- 3) L'énergie dans chaque canal.

#### ***4.2.2 Méthode utilisée***

Chaque portion de l'onde finale est construite par sommation des ondes partielles synthétisées par chacun des canaux.

Cette technique a été utilisée initialement en musique synthétique, et a été appliquée dans le domaine de la parole.

Ex. Le Vocodeur ciphon, l'icophone, et l'icolog (l'icophone logiciel) du Limsi.

#### ***4.2.3 Fonctionnement de l'icophone***

On réalise la synthèse sur un banc de 44 oscillateurs analogiques sinusoïdaux, de fréquence fixe en tout ou rien.

Les fréquences vont de 100hz à 4400 Hz, avec une différence constante de 100 Hz entre un oscillateur et le suivant. Ce synthétiseur a pour avantage d'être flexible; il est en effet possible d'effectuer des variations temporaires qui permettront en particulier l'introduction de la mélodie, c'est-à-dire la variabilité de hauteur des sons successifs, la programmabilité des formes d'ondes et des fréquences des oscillations.

#### ***4.2.4 Fonctionnement de l'icolog.***

La méthode choisie est de multiplier le signal par une fenêtre temporelle fixe reproduisant les impulsions du signal glottique, se répétant à la fréquence fondamentale recherchée. De cette façon, la mélodie est reconstituée dans le domaine temporel.

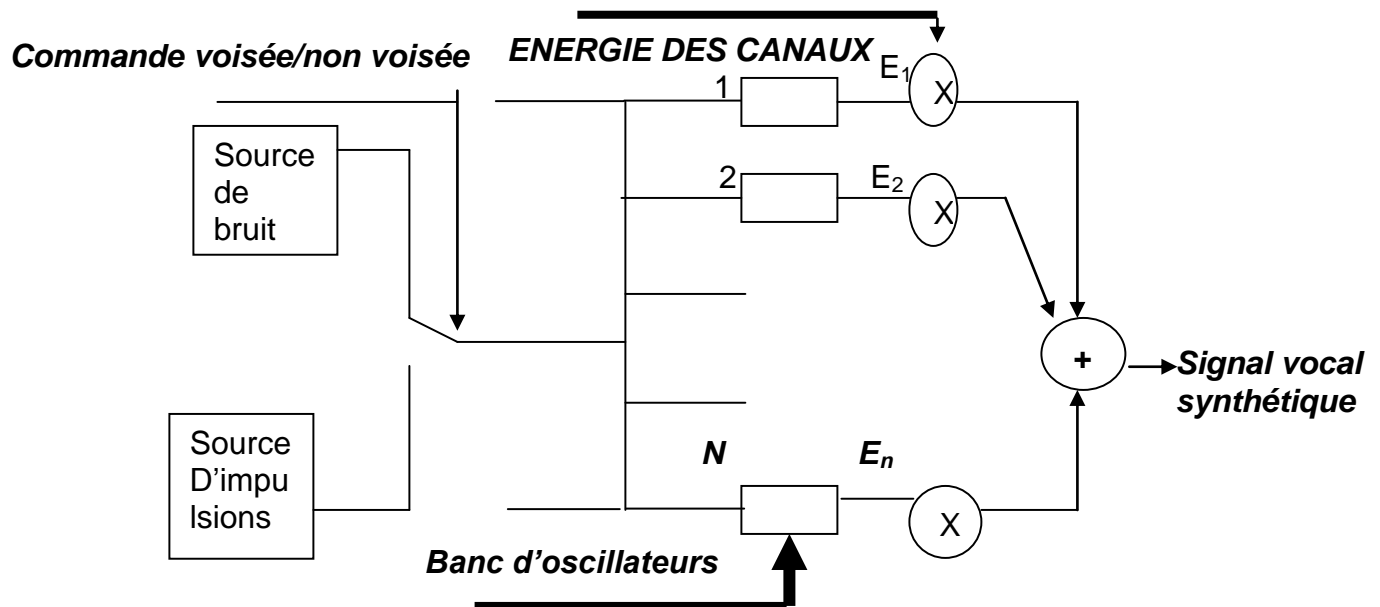
L'icolog suit la technique la plus économique et la plus souple. Cette technique est appelée « table Look-up » employé surtout dans les systèmes digitaux de synthèse musicale de la génération numérique de la forme d'onde d'un oscillateur sinusoïdale. D'après cette technique, la forme d'onde, une fois calculée, est stockée dans une mémoire tampon ; cette mémoire est lue ensuite circulairement avec un pas variable dépendamment de la fréquence recherchée. Il utilise 40 oscillateurs à raison d'une table de forme d'onde par oscillateur.

#### **4.2.4.1 Avantage de cette technique**

Cette technique a pour avantage la réduction des temps de calcul et l'obtention d'un banc d'oscillation avec un nombre de circuits très réduit. Mais cette technique n'utilise pas la forme d'onde sinusoïdale pour chaque oscillateur. Elle contient plusieurs fréquences pures autour de la fréquence

nominale.

Le tableau suivant est la représentation d'un synthétiseur à canaux :



#### 4.3 La synthèse par formant

Les caractéristiques du modèle réalisé par un nombre important de filtres électriques produisent des propriétés résonantes du conduit vocal que l'on appelle « synthèse par formants ».

Les paramètres du synthétiseur sont :

- 1) La commande voisée/ non –voisée
- 2) La fréquence centrale
- 3) La largeur de bande
- 4) L'énergie de chacun des formants

##### 4.3.1 Réalisation d'un synthétiseur à formants

De par son principe, le synthétiseur à formant simule plus fidèlement le fonctionnement de l'appareil vocal que le synthétiseur à canaux ; il y a en fait une véritable analogie entre le montage électrique et la machine vocale de l'homme.

Si le synthétiseur à canaux a pour seul rôle de fournir un signal sonore, et le spectre en fréquence se rapproche au mieux d'un certain modèle fourni par l'analyseur, sans aucune contrainte quant au conduit vocal, Le synthétiseur à formants possède un spectre se rapprochant de celui du débit d'air au niveau des cordes vocales.

La qualité de parole synthétisée est par contre beaucoup moins sensible aux caractéristiques de

l'excitation non-voisée. un bruit blanc à large bande convient parfaitement : on sait le générer par un circuit numérique tel qu'un registre à décalage de 18 bits à bouclage pseudo-aléatoire travaillant au rythme d'une horloge à 20KHz.

Le conduit vocal, de la glotte aux lèvres est modélisable par un tube de section variable, dont la fonction de transfert s'écrit :

$$H(P) = \frac{w_1^2}{P^2 + b_1 \cdot p + w_1^2} * \frac{w_2^2}{p^2 + b_2 \cdot p + w_2^2} \dots$$

Nous allons maintenant considérer la modélisation d'un formant par la fonction de transfert :

$$H(P) = \frac{w_c^2}{P^2 + b \cdot p + w_c^2} \quad (2)$$

Avec  $w_c$  : fréquence nominale de résonance en rad/s

b : la bande passante de la courbe de résonance

$P = j\omega$  et appliquons donc à 2 formants et l'on a :

$$\left| H(j\omega) \right|^2 = \frac{w_c^4}{[\omega^2 - (w_c^2 - b^2/2)]^2 + b^2(w_c^2 - b^2/4)} \quad (2)$$

et est max quand  $\omega$  varie et prend la valeur  $(w_c^2 - b^2/2)^{1/2}$  pour laquelle le premier terme du dénominateur s'annule.

NB1- La largeur de bande  $b$  est généralement plus petite que la valeur de la fréquence centrale  $w_c$  de sorte qu'une bonne approximation du pic est :  $\omega = w_c$  et son amplitude est alors égale à  $w_c/b$ .  
 Considérons deux sons dont l'un est le double de l'autre. Dans ce cas la fréquence est supérieure au pic, et l'énergie varie comme  $1/\omega^4$  le niveau est divisé par un facteur 16, ce qui s'exprime en décibels par  $10 \log 16 = 16 \text{ dB/octave}$ .  
 A  $3 \text{ dB}$  au dessus du pic, l'énergie devient :  $H(j\omega_{3\text{dB}})^2 = H(j\omega_{\text{max}})^2 = w_c^2 / 2b^2$  et avec une bonne approximation, la fréquence est :  $\omega_{3\text{dB}} = w_c \pm b/2$ .

NB2- La largeur de bande d'un formant est quasiment constante quelque soit la fréquence du formant.

### 4.3.1 Réalisation d'un synthétiseur à formants

Un formant, est par définition un ensemble de filtres électriques. Un formant est réalisable de deux manières différentes :

#### 4.3.1.1) En série

La façon la plus simple de réaliser un synthétiseur à formants est d'associer en série plusieurs filtres

électriques.

#### 4.3.1.1.1 Désavantage de ce montage

Quand il y a deux fréquences de formants très proches l'une de l'autre successivement et que les largeurs de bandes sont minimales, on atteint des acuités de résonance extrêmement élevée. Par exemple pour une fréquence de 2500Hz, l'acuité est de l'ordre de 1800, soit 65dB ; ce qui entraîne une charge trop importante pour la plupart des filtres analogiques. Dans ce cas, les filtres vont être altérés. Il en résulte une dégradation de la qualité de la parole.

#### 4.3.1.1.2 Comment l'éviter ?

On l'évite en choisissant des bandes passantes les plus larges et en implémentant le filtre du formant1 entre deux des formants 2 et 3 de sorte qu'il y a une atténuation du signal du formant2 avant d'atteindre le circuit du formant 3; bien qu'en fait dans la parole naturelle, deux formants n'ont jamais exactement la même fréquence.

On peut également simplifier les formants d'ordre supérieur. Vu que la bande passante  $b_n$  plus lente que la fréquence  $w_n$ ,  $w_n/b_n$  est devenue négligeable pour le  $n^{\text{ième}}$  formant.

#### Spectre vocal en décibels :

Posons :  $-10 \text{ Log } [(1-w^2/w_n)^2 + b_n^2 w^2/w_n^4]$ , soit approximativement  $-10 \text{ Log } (1 - w^2/w_n^2)^2$ .

Il vient après développement limité :

$20 \cdot w^2/w_n^2 \cdot \text{Log } e$  avec  $e = 2.71828...$  (soit  $\text{Log } e = 0.434$ )

$f_n = w_n/2\pi = (2_{n-1})^* 500 \text{ Hz}$

Considérons les fréquences suivantes des formants 4, 5, 6, etc....3500Hz, 4000Hz, 5500Hz.

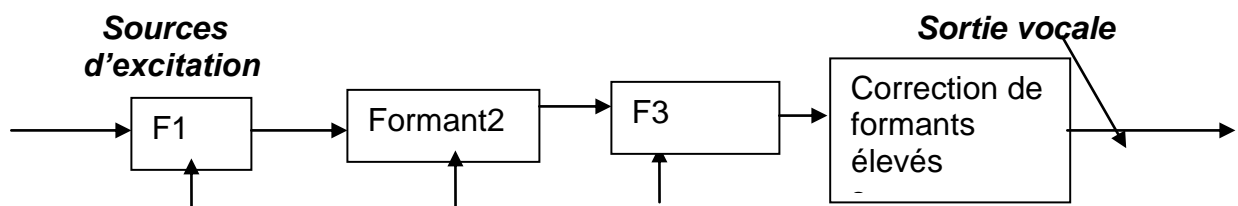
Posons  $f = w/2\pi$ , la contribution de tous les formants d'ordre supérieur à 3 s'obtient en sommant les contributions. Il vient donc :

$$20 \cdot \text{Log } e \sum_{n=4}^{\infty} f^2 / (2_{n-1})^2 \cdot (500)^2$$

Soit finalement :  $20 \cdot \text{Log } f^2/500^2(\pi^2/8 - \sum 1/(2n-1)^2)$ ; (n va de 1 à 3) =  $2.87 \cdot 10^{-6} \cdot f^2$  Décibels. Ce qui donne une correction très élevée de 72dB pour produire un spectre. Cette fois la correction de formants supérieurs est de 43dB pour bien représenter le signal de la parole jusqu'à  $f = 5\text{kHz}$ .

Si l'on accepte de réduire à 3500Hz la bonne représentation du signal de la parole, la correction de formants supérieurs ne sera plus que 35dB.

#### 4.3.1.1.3 Représentation des formants en série



### 4.3.1.2 Montage en parallèle

Pour le montage en parallèle, la fonction transfert peut être donnée par :

$$H(p) = \frac{w_1^2}{p^2 + b_1 p + w_1^2} + \frac{w_2^2}{p^2 + b_2 p + w_2^2}$$

#### 4.3.1.2.1 Avantage du montage en parallèle

Ce montage présente un avantage certain; qui permet qu'à chaque filtre de formant, associer un filtre de correction. Dans ce cas, le filtre doit avoir de faible gain pour éviter de grandes variations d'amplitude.

#### 4.3.1.2.2 Inconvénient du montage en parallèle

Contrairement au montage en série, ce dernier exige plus d'informations. Il faut préciser l'amplitude et la fréquence des formants. Mais, il permet d'avoir un meilleur contrôle du synthétiseur.

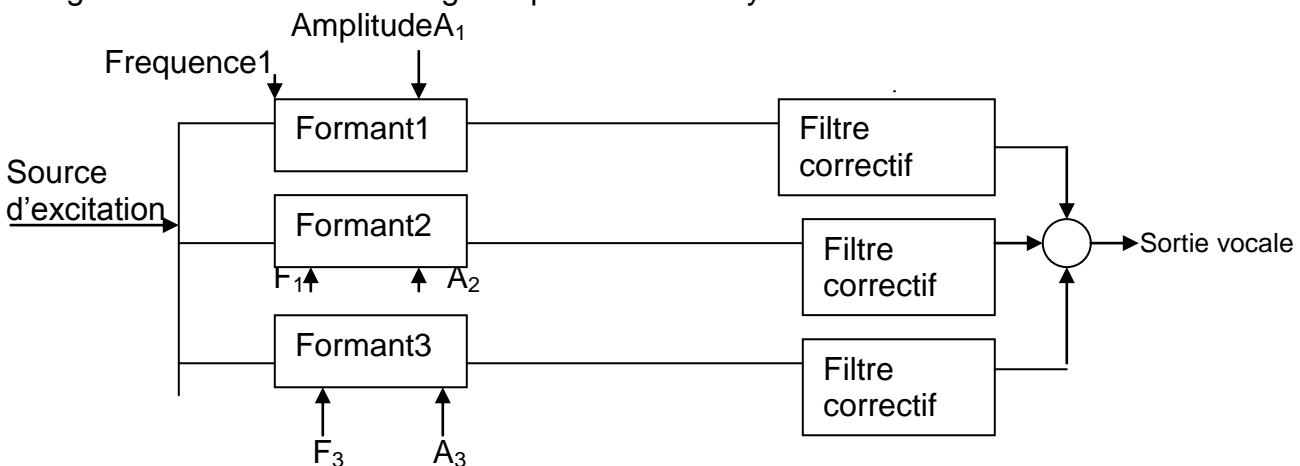
Un exemple d'un synthétiseur en parallèle de sons est le **nez** et la **bouche**. Donc, la mise en

$$H(p) = \frac{A_1 w_b^2}{p^2 + b_b p + w_b^2} + \frac{A_2 w_n^2}{p^2 + b_n p + w_n^2}$$

parallèle de la cavité nasale à la cavité buccale donne la nasalisation. Et sa fonction de transfert s'écrit donc :

; Avec b: mis pour buccale et n: pour nasale

La figure suivante est le montage en parallèle d'un synthétiseur :



### 4.4 La synthèse résonante dans le domaine temporel

Soit la fonction transfert dans le domaine fréquentiel :  $H(p) = w_n^2 / p^2 + bp + w_n^2$ . Pour la transformer



dans le domaine temporel, on obtient une réponse sinusoïdale amortie de la forme :  $X_n(t) = w_n e^{-bt/2} \sin(w_n t)$ , en négligeant les termes  $b^2/w_n^2$ . Il faut noter que, en associant plusieurs de ses formes d'onde, excitées régulièrement par une impulsion à la fréquence (fondamentale) du signal glottique, on engendre aussi un son voisé, et, puis la fonction devient:

$X_n(t) = e^{-bt/2}(A_1 \sin w_1 t + A_2 \sin w_2 t + A_3 \sin w_3 t)$  avec  $A_1, A_2, A_3$  servant à la commande d'amplitude de chaque formant. Pour implémenter un synthétiseur sous la forme numérique, on stocke dans une mémoire morte, un cycle entier d'une onde sinusoïdale ; Les fréquences de formant sont contrôlées en lisant cette mémoire à différentes vitesses. Chaque mot numérique ainsi obtenu est multiplié par un coefficient d'amplitude ou d'atténuation, converti en information analogique. Celle-ci est amortie à l'aide d'un circuit capacitif. L'ensemble peut fonctionner en temps réel sous la forme d'un microprocesseur couplé à un convertisseur numérique analogique.

#### 4.5 La synthèse à prédiction linéaire

Les vitesses lentes du conduit vocal permettent de prédire approximativement un échantillon de parole  $s$  (NT) par une somme pondérée linéairement d'un certain nombre d'échantillons :

$$s'_n = \sum_{k=1}^P a_k s_{n-k}$$

Où  $s_n$  représente  $s(nT)$ , avec  $T$ : période d'échantillonnage du signal de la parole  $a_k$ , et,  $k$  le  $K^{ième}$  coefficient de prédiction,  $P$  l'ordre de prédiction (ou nombre de pôles),  $s_n$ : échantillon prédit.

Notons qu'il peut y avoir d'erreurs ; laquelle est alors définie par la relation :

$$e_n = s_n - s'_n = s_n - \sum_{k=1}^P a_k s_{n-k}$$

Il faut dire que les coefficients  $a_k$  ont été choisis pour minimiser cette erreur; ils sont réactualisés toutes les 10 à 25 ms

En prenant la transformée en  $z$  des deux membres de cette relation, on obtient :

$$E(z) = S(z) \left( 1 - \sum_{k=1}^P a_k z^{-k} \right)$$

$$S(z) = E(z) / \left( 1 - \sum_{k=1}^P a_k z^{-k} \right)$$

Et, le conduit vocal est susceptible d'être représenté par une succession de circuits numériques de formants ayant chacun pour fonction de transfert :

$1/(1-b_1z^{-1}+b_2z^{-2})$  où  $b_1$  et  $b_2$  contrôlent la position et la largeur de bande de la résonance d'un formant. Les lèvres sont modélisées en général par un filtre du premier ordre:

$$1/(1-b_1z^{-1}).$$

Le produit de tous les termes donne: 
$$\frac{1}{1-c_1z^{-1}-c_2z^{-2}-\dots-c_qz^{-q}} \frac{1}{1-\sum_{k=1}^p c_k z^{-k}}$$

Et  $q$  vaut deux fois le nombre de formant +1 (pour tenir compte de la présence des lèvres), la transformée en  $z$  de l'échantillon de parole est :

$$S(z) = I(z) / (1 - \sum_{k=1}^q a_k z^{-k})$$

On retrouve une formulation similaire à celle obtenue précédemment par la prédiction linéaire : il suffit d'établir l'analogie entre l'ordre de prédiction et le nombre de formants, entre les coefficients de prédiction  $a_k$  et  $c_k$  donnant la position et la largeur de bande des formants, enfin entre l'erreur de prédiction  $E(z)$  et l'excitation par impulsion  $I(z)$ .

Des lors, il est possible de paramétrer le signal d'erreur (donc le calculer) par sa fréquence et son amplitude (deux quantités variant lentement), plutôt que de le transmettre d'un échantillon au suivant.

#### 4.5.1 Techniques de sondage

En général, les techniques de sondage de la parole par prédiction linéaire fonctionnent avec des valeurs de  $p$  pris entre 10 et 15, ce qui correspond à la synthèse respectivement par 4 et 7 formants. Les paramètres de la synthèse sont recalculés toutes les 10 à 25 ms : ce laps de temps correspond à la durée minimale pendant laquelle des événements acoustiques peuvent survenir dans la production naturelle de parole.

#### 4.5.2 Fréquence d'échantillonnage

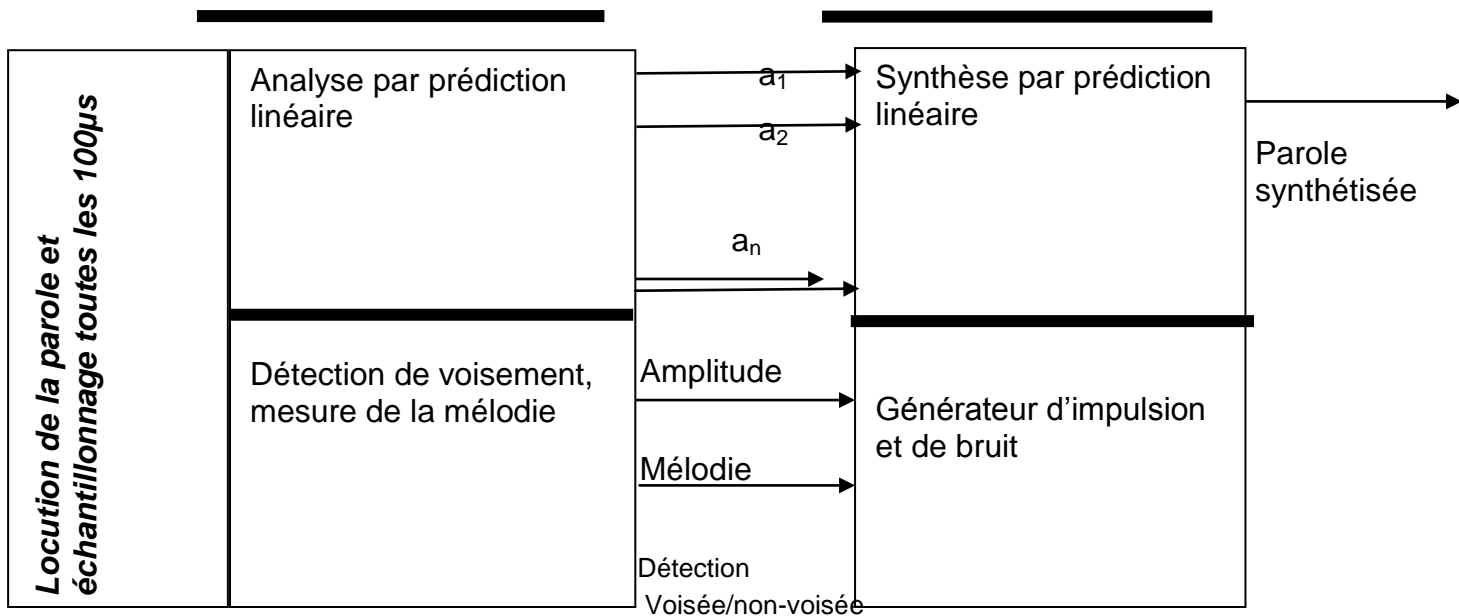
La parole naturelle a elle-même été échantillonnée à la fréquence de 8 à 24 kHz. Pour produire un timbre plus riche, le nombre  $p$  de pôles est augmenté: ce qui le porte à  $p=55$ , avec un échantillonnage à 22,5 KHz.

A la réception, la forme d'onde est reconstituée au moyen d'impulsions pour les sons voisés et d'un bruit blanc pour les sons non-voisés.

Plusieurs types de filtres ont été utilisés pour la synthèse. La structure directe présente des inconvénients du fait que les coefficients  $a_k$  de prédiction ne sont pas bornés en particulier si elle est utilisée pour le codage à faible débit : il y a un risque d'instabilité du filtre de synthèse. En cas d'instabilité du filtre, le signal de parole synthétique oscille fortement. Ce phénomène se présente lorsqu'on apporte des modifications même très légères aux coefficients  $a_k$  du prédicteur.

On résout le problème en faisant usage d'un autre ensemble de paramètres: les coefficients de réflexions qui eux sont bornés et compris entre -1 et +1 pour les filtres stables; il y a un moyen simple de contrôler la stabilité d'un filtre.

La figure suivante montre que dans la technique de synthèse par prédiction linéaire, on sépare bien les paramètres en provenance de la source de sons (liés au signal d'erreur de prédiction) de ceux fournis par le conduit vocal (représentés par les coefficients de prédiction linéaire).



Le conduit vocal peut-être assimilé à un tube acoustique non uniforme constitué par une succession de N sections cylindrique d'égale longueur, moyennent 2 hypothèses: la surface transversale de chaque section est très petite par rapport à la longueur d'onde (on peut alors traiter l'onde se propageant à travers une section comme une onde plane) et les dimensions du type acoustique sont fixes pendant une période d'analyse. Sa fonction de transfert a pour expression:

$G(z) = k/E(z) / 1 - \sum_{k=1}^p a_k z^{-k}$ . Cette fonction a deux pôles.

Et les coefficients  $K_n$  de réflexion de l'onde à la jonction entre deux sections  $n$  et  $n+1$  du conduit vocal sont liés de manière réversible aux coefficients de prédiction par trois relations :

$$A_k^{(n+1)} = a_k^n + k_n * a_{n+1-k}^{(n)}, \text{ (pour } k=1,2,3,\dots,n)$$

$$A_{N+1}^{(n+1)} = -k_n$$

$$K_n = \frac{R_{n+1} - \sum_{k=1}^n a_k^{(n)} R_{n+1-k}}{R_0 - \sum_{k=1}^n a_k R_k}$$

Les Coefficients  $a_k^{(n+1)}$  sont ceux d'un filtre modélisant un conduit vocal à «  $n+1$  » sections terminés par une charge résistive (simulation des lèvres); les coefficients  $R_n$  sont des échantillons de la fonction d'autocorrection du signal de parole.

Les aires  $A_n$  des sections du conduit vocal sont, quant à elles liées aux coefficients de réflexion par la relation apparemment très simple :

$K_n = \frac{A_n - A_{n+1}}{A_n + A_{n+1}}$
---

Les coefficients  $k_n$  sont des fonctions d'aire du conduit vocal.

#### 4.6 La synthèse par mots ou petits bouts de phrases

Selon la méthode de codage adoptée (synthèse par phrases, par mots, par fraction de mots, par phonatomes, par allophones ou par phonèmes), on disposera de plus ou moins de parole synthétisée: un disque de 100Mo contiendra trois heures et demie de parole code MIC à 64Kbits/s. Ce même disque pourra par contre stocker huit jours pleins de parole continue codée à 1150 bits/s par prédiction linéaire

NB.- le codage prédictif peut-être appliqué à toutes les méthodes de concaténation

La synthèse par mots (ou petits bouts de phrases), la reconstitution de message se fait par

juxtaposition des mots les uns à la suite des autres: “ou jwenn” “lyon an”. “men” “kay la wouj la”; “chèche ankò” “se” “kamyon ble a”.

Cette synthèse est plus adaptée au synthétiseur par formants.

#### 4.6.1 Désavantage

La difficulté principale est de résoudre le problème de liaison d'un mot à un autre pour obtenir la parole continue et de créer une intonation artificielle pour conserver le naturel de la voix. Le problème est complexe et l'on se contente de nombreuses applications de synthétiser les mots isolément à moins, bien entendu, d'avoir recours à des techniques temporelles (codage MIC), qui néanmoins, malgré leur fort débit de parole, ne peuvent éviter la présence d'interférences ou de clics (c'est-à-dire deux phénomènes à double occlusions) ressemblant à un bruit de succion) entre les mots : les informations, si elles sont numérisées sont alors éditées afin d'éliminer ces défauts. Ces techniques s'appliquent dans des systèmes de renseignement automatiques, ou des bribes de phrases sont séparées par des blancs que l'ordinateur remplit de mots selon les circonstances.

#### 4.7 Synthèse par éléments phonétiques

Pour certaines langues il est possible d'augmenter le vocabulaire tout en diminuant l'encombrement en mémoire, en constituant artificiellement des mots à partir de morceaux de mots. Par exemple, à partir des cinq éléments suivants : « auto », « Mathé », « Mati », « Cien » et « que », il est possible, par juxtaposition de ceux-ci, de créer de nombreux mots : « automatique », « automaticien », « mathématicien », « mathématique », ... Le découpage peut se faire plus finement au niveau des syllabes. Ce procédé est utilisé quand les lois de concaténation des syllabes ne sont pas trop compliquées.

On peut aller plus loin dans la segmentation d'un langage, jusqu'à la plus petite unité phonétique : le phonème, qui n'a de sens que par lui-même; seul l'assemblage des phonèmes, selon un ensemble de lois particulières à chaque langage permet de reconstituer les mots parlés.

NB- L'association de plusieurs phonèmes ne donne pas nécessairement un mot compréhensible dans le langage parlé; ainsi un « cha » « ...a » n'est pas un « chat ». Donc, l'association de plusieurs phonèmes conduit à la perception de deux sons séparés par un temps mort ; c'est la transition d'un phonème vers un second qui rend un mot intelligible. Pour reconstituer un mot ou une phrase il suffit d'assembler les diphonèmes nécessaires par leur extrémité commune comme on assemble les dominos.

Une analyse statistique réalisée au Limsi à Orsay a montré qu'à partir d'un répertoire de 627 diphonèmes, il est possible de reconstituer n'importe quelle phrase en français.

(Réf. MARC, Ferretti, FRANÇOIS, Cinare, *SYNTHESE, RECONNAISSANCE de la PAROLE*, éditeurs, Paris, 1983, p.142 §4).

La synthèse par phonèmes ou phonatomes a donné lieu à de nombreuses réalisations. A Orsay, Limsi a développé L'icophone en collaboration avec le laboratoire d'acoustique de l'Université Paris expliqué plus haut.

#### 4.8 Synthèse articulatoire

Dans la synthèse articulatoire, on tient compte par exemple des inerties de la mâchoire et de la langue. On devrait obtenir une parole de très haute qualité. La réalisation pratique d'un tel synthétiseur est encore à démontrer.

#### 4.9 Méthode choisie

Parmi les méthodes précitées, nous utilisons une méthode assez proche de la méthode générique d'il y a quelques années. Elle coupe le texte en phonèmes avant de les convertir en propos audibles en suivant un algorithme finalement bien connu.

#### Les différentes étapes

Pour y parvenir, nous avons utilisé la même méthodologie que *FreeTTS*. Des opérations successives sur le texte ont été effectuées, les résultats cumulatifs de ces opérations sont stockés dans une structure d'expression qui permet finalement l'analyse complète du texte. C'est donc une méthode de synthèse par *concaténation*, dans le style de ce qui a été implémenté dans *Mbrola* vers la fin des années 90 :

- ❖ *La normalisation du texte* : C'est la première étape du traitement. Elle permet de transformer le texte saisi en un "flux" de mot prononçables en détectant et en transformant tous les termes dont la prononciation pourraient poser des problèmes
- ❖ *L'analyse lexicale* : ce processus permet de déterminer les éléments de prononciation, l'identification des syllabes et le stress à mettre sur chaque mot et/ou expression. FreeTTS utilise un lexique pour déterminer ces éléments.
- ❖ *La Génération de la prosodie* : c'est ce processus qui détermine les « meilleurs » paramètres pour les pauses, la durée, la tonalité, le stress, etc. Pour l'expression. On sait que plusieurs méthodes ou algorithmes peuvent être utilisés à cet effet. FreeTTS utilise typiquement les arbres de classification et de régression (CARTS [brieman84]) pour produire l'information de prosodie
- ❖ *La synthèse* : Cette étape permet de générer le son comparable à la parole en faisant la concaténation des unités obtenues sur la base de diphones ou d'autres unités de paroles.

Décomposé en ces étapes successives, le processus de synthèse par concaténation utilisé par

*FreeTTS* devrait être relativement fidèle, et donc le rendu assez proche de la réalité.

De plus la méthode utilise des bases de données qu'on peut rendre suffisamment grande pour avoir des performances optimales. Cette description, succincte mais complète de la forme de la méthode par concaténation choisie par les concepteurs de FreeTTS, en illustre le caractère complexe mais néanmoins relativement incomplet. En effet, cette méthode fait le choix (délibéré ?) de laisser plusieurs étapes, (par exemple l'analyse morphologique, ...) de la synthèse par concaténation classique. Les conséquences se font remarquer nettement à l'audition. Il est vrai que la synthèse par concaténation (même moins que les algorithmes de la génération suivante) demande une grande complexité de calcul à la machine. Ajouter à cela, le langage (Java) qui installe une surcouche au dessus du système d'exploitation pour fonctionner, on pourrait imaginer une explication plausible : le souci de performance a justifié le choix de la simplification de la procédure de synthèse.

## Chap.5 *Les circuits de parole du marché*

---

### **Aperçu**

#### 5.1 Introduction

#### 5.2 Les circuits de synthèse

##### 5.2.1 Truqueur de voix Kit MK171

##### 5.2.2 Voix de robot (Kit B107)

##### 5.2.3 Chien de garde électronique (Kit K2655)

##### 5.2.4 Module d'enregistrement (Kit K8030)

##### 5.2.5 American Microsystems, inc. (Ami, une filiale de Gould Inc.)

##### 5.2.6 Prosit 400 de AML

##### 5.2.7 Robot de reconnaissance orale chez BFA (en collaboration avec Valoris)

##### 5.2.8 CENTIGRAM : une compagnie vocale

Dans cette partie, sans prétendre à l'exhaustivité, nous présentons différents circuits de synthèse qui ont été commercialisés.

## 5.2 Les circuits de synthèse

On accorde un certain intérêt à l'activité de conversion de texte en parole. Ceci par le fait du développement rapide des nouvelles technologies, les ordinateurs personnels devenus plus puissants ont permis de réaliser jusque-là des applications intéressantes en ce sens. En témoignent les nombreux produits de ce type existant sur le marché.

### 5.2.1 Truqueur de voix Kit MK171



Ce petit montage permet de faire ressembler la voix humaine à celle d'un robot (avec un effet vibrato). Bouton 'pitch' prévu pour élever ou baisser le ton de votre voix. Equipé d'un microphone et d'un amplificateur avec réglage de volume, il suffit de connecter un haut-parleur de 4 ou 8 ohms. Applications: fête, théâtre, enregistrements, jeux, blagues téléphoniques etc. Sortie haut-parleur/ligne. Interrupteur on/off. Puissance de sortie maxi: 1.5 W (4 ohms). Alimentation à prévoir: 9 Vcc/300mA. Dimensions: 77 x 46 mm. Kit à souder soi-même.

### 5.2.2 Voix de robot (Kit B107)



Cet appareil est tout simplement intercalé entre le microphone et l'amplificateur final. Il provoque une déformation de la voix, qui ressemblera à celle d'un robot. On peut régler la manière et l'intensité de la déformation. Ce module est intéressant pour les discothèques, les prises de son, etc. Alimentation: 9...12 Vcc. Entrée: env. 5 mV pour microphone (non fourni). Sortie: > 50 mV pour entrée amplificateur. Dimensions platine: 55 x 45 mm. Kit à souder soi-même.

### 5.2.3 Chien de garde électronique (Kit K2655)





Circuit qui aboie comme un chien et qui a l'oreille sensible à tout ce qui se passe dans les environs. Choix entre 2 différents chiens. Sensibilité réglable. Sortie de haut-parleur (2 W sous 4 ohms).

#### 5.2.4 Module d'enregistrement



Enregistrement et reproduction de la voix sans intervention mécanique. Applications: sonnette d'entrée, jeux, etc. excellente qualité sonore. Durée d'enregistrement maximale: 4 à 20 secondes. Micro et HP livrés avec le kit. Deux possibilités de lecture. Possibilité de sélectionner jusqu'à 5 messages différents (à l'aide d'un commutateur non fourni - code 07204). Alim. à prévoir: 8-12 Vcc ou piles 6 V (non fournies). Dimensions: 94 x 73 x 25 mm. Kit à souder soi-même.

#### 5.2.5 American microsystems, inc. (Ami, une filiale de Gould inc.)

Ami produit deux synthétiseurs à codage prédictif à dix coefficients de réflexion pour les sons voisés et quatre pour les sons non-voisés. Ce sont le S3610 et le S3620. À partir du signal de la parole échantillonné à 8kHz, on extrait les coefficients représentatifs de train de 20 ms.

La valeur de la fréquence fondamentale est calculée par un algorithme fondé sur l'analyse cepstrale, tandis que les coefficients de réflexion sont obtenus par corrélation.

Les paramètres de commerce du synthétiseur S3610 se trouvent archivés dans une mémoire de 2560 octets; celle-ci peut contenir 32 mots, soit approximativement 17s de parole.

Au modèle 3620, doit nécessairement être associée une mémoire morte externe pour le stockage du vocabulaire.

Le vocabulaire est généré par AMI grâce à un logiciel exécuté sur VAX 11/780 ou par tout utilisateur possédant des moyens similaires.

#### 5.2.6 Prosit 400 de AML

C'est un synthétiseur de parole à partir du texte conçu et réalisé à partir du savoir-faire du CNET et son développement a bénéficié d'une aide à l'innovation de l'ANVAR.

Le mater comporte un microprocesseur (de 8 bits) 8085; la synthèse de la parole est réalisée par le synthétiseur SDP du CENT.

Le logiciel fonctionne en deux phases principales:

- une transcription orthographique- phonétique du français, produisant une suite de phonèmes;
- La synthèse proprement dite utilise un appariement des phonèmes en diphones avec calcul de la mélodie et du rythme en fonction d'un schéma mélodique pré-établi, de marqueur et du nombre de syllabes de mots.

#### 5.2.6 Robot de reconnaissance orale chez BFA (en collaboration avec Valoris)

BFA (Basse Fréquence Appliquée), commercialise des produits qui parlent. La carte de synthèse de parole Oratoire est un modèle utilisant la technologie de synthèse par mots comprenant :

- Un générateur de mots ;
- Un codeur ;
- Un système d'amplification.

#### 5.2.7 CENTIGRAM: une compagnie vocale

Cette compagnie établie à Californie a connu le succès en 1977 avec Mike, une interface vocale d'entrées-sorties pour ordinateurs. La première version de Mike pouvait reconnaître 16 mots (avec l'aide d'un ordinateur) et de dire à haute voix 16 ms de parole synthétique.

Chez CENTIGRAM, la synthèse est assurée par LISA (Logical Integrated Annunciator); la technique employée est un codage dénommé PWC (Parameterized Waveform Coding) par Centigram, mais qui en fait dérivé du codage par prédiction linéaire. On obtient un débit de parole de 4800bits/s.

## Chap.6 *Comparaison entre langages naturels et artificiels*

---

### **Aperçu**

#### 6.1 Introduction

6.1.1 Pourquoi l'intelligence artificielle ?

6.1.2 Pourquoi l'intelligence artificielle ?

#### 6.1.3 L'apprentissage

#### 6.2 Quelques critères de compréhension des langues naturelles

#### 6.3 Aspects spécifiques de la compréhension du langage naturel parlé

#### 6.4 Les premières tentatives de traduction automatique des langues naturelles

#### 6.5 Des techniques orientées automates

## 6.1 Introduction

L'intelligence artificielle est donc la discipline visant à comprendre la nature de l'intelligence en construisant un programme d'ordinateur imitant l'intelligence humaine.

Donc, un programme d'intelligence artificielle est un programme capable de prendre des décisions. C'est le cas par exemple d'un programme qui résout un problème ou prend une décision en vue de la description de certaines situations trouve lui-même la méthode de résolution à employer face à un vaste gamme de mécanismes de raisonnement qui lui ont été incorporés.

### 6.1.1 Pourquoi l'intelligence artificielle?

Pour écrire ce programme, nous avons exploité le potentiel des technologies de l'information et des télécommunications (tics). Et, le but du programme c'est la conception et l'implémentation d'un logiciel suffisamment intelligent pour lire un texte écrit en créole et le transforme en son. D'où, la nécessité du recours à l'intelligence artificielle.

#### **6.1.1.1 L'apprentissage**

La faculté d'apprendre est l'une des composantes fondamentales de l'intelligence. Elle permet d'augmenter ses connaissances, d'améliorer ses aptitudes et ses mécanismes de raisonnement.

Il faut noter qu'un bon programme d'apprentissage doit s'appuyer sur une bonne représentation des connaissances qu'il a pour but d'améliorer. Pour la compréhension de la notion d'apprentissage, Rosenblatt est reconnu comme pionnier avec les perceptions dites « perceptions de Rosenblatt ». A cette époque, l'apprentissage était synonyme d'adaptation puisqu'il s'agissait simplement d'ajuster la valeur de paramètres inconnus par des approximations successives. Ces paramètres étant censés caractériser certaines structures ou certains concepts dont on essaie d'apprendre le comportement ou la définition.

Cette forme était numérique. Son but : acquérir une fonction discriminante linéaire ou polynomiale à partir d'exemples.

Cette approche a été abandonnée dans les années 1960 au profit de celle dite symbolique fondée sur des relations conceptuelles.

Il existe 5 types d'apprentissage :

- 1) L'apprentissage par cœur et implémentation directe de nouvelles connaissances.
- 2) L'apprentissage par instruction

- 3) L'apprentissage par analogie
- 4) L'apprentissage à partir d'exemples
- 5) L'apprentissage par observation et découverte

1) Par cœur et implémentation de nouvelles connaissances.

Aucune inférence n'est nécessaire de la part du programme qui mémorise les nouveaux faits ou de nouveaux cas sans leur faire subir de transformation ; c'est le type de la programmation habituelle.

2) Celui dit par instruction.

Dans ce type d'apprentissage les nouvelles connaissances ou conseils proviennent de l'extérieur du programme sous une forme à simuler et à intégrer afin d'améliorer le raisonnement

3) Par analogie

On a des connaissances existantes et l'on acquiert de nouveaux faits avec de fortes similitudes. L'apprentissage par analogie permet de s'adapter à de nouvelles situations analogues à des situations déjà rencontrées.

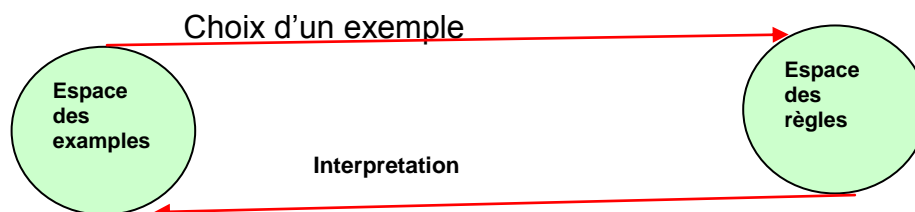
4) À partir d'exemples

Cette forme d'apprentissage veut qu'un programme puisse apprendre à partir d'exemples. Dès lors que des situations spécifiques lui sont présentées, le programme doit induire des règles générales. Pour enseigner le concept « garçon » à un programme on lui donnera des exemples de personnes (et autres) en lui précisant celles qui sont des garçons. Le programme doit découvrir progressivement quelles sont les règles pour reconnaître un garçon. Dans ce cas, nous avons le modèle bi-spacial de Simon et de Léa. On y trouve :

- a) l'espace des exemples
- b) l'espace des règles.

NB. L'espace des règles pourrait avoir une autre forme quand la description est de haut niveau à la manière de la description d'un objet ou de concept.

Considérons la description suivante :



❖ L'espace des exemples

Ce modèle veut qu'on dispose des exemples de bonne qualité. Ces exemples choisis doivent permettre de facteurs discriminants dans la définition des concepts à apprendre.

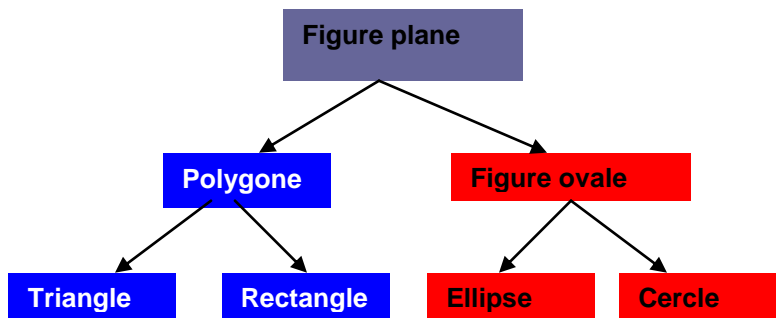
Un aspect très important est l'ordre dans lequel sont présentes les exemples c'est-à-dire la qualité du pédagogue éduquant le programme : un bon pédagogue accroît progressivement la difficulté des exemples.

❖ L'espace des règles

Le processus fondamental de l'apprentissage à partir d'exemples est la « généralisation ». Considérons 'X' et 'Y'. 'X' est plus général que 'Y' si 'X' s'applique dans tous les cas où 'Y' s'applique au moins.

**Exemple de généralisation par ascension dans un arbre**

Elle décrit la notion de forme en géométrie plane.



Une généralisation par ascension dans un arbre peut-être :

Forme (X, Rectangle) ::> Bleu(X)

Forme (X, Triangle) ::> Bleu(X)

Se généralise en :

Forme (X, Polygone) ::> Bleu(X)

Alors on sait que les objets d'une classe X qui sont triangulaires sont bleus.

De même pour 'Y'

Forme (Y, Ellipse) ::> Rouge(Y)

Forme (Y, Cercle) ::> Rouge(Y)

Se généralise en :

Forme (Y, Ovale) ::> Rouge(Y)

Le programme doit savoir qu'à chaque qu'il trouve **Figure plane**, il y a soit un rectangle, Triangle, Ellipse ou Cercle.

Deux approches peuvent être considérées pour parler du langage naturel et artificiel :

- a) Communication homme-homme
- b) Communication homme-machine

Abordons la première approche (Communication homme-homme) : le langage naturel est considéré ici, comme étant simplement l'utilisation d'une langue pour communiquer (Anglais, Français, Créole, etc..) ; Dans ce même cas le langage artificiel est considéré, comme étant des langues construites c'est-à-dire inventées pour une utilisation pratique.

La deuxième approche dite, communication homme-machine est celle que nous allons considérer

dans ce document : cette communication implique que l'individu s'adresse directement à l'ordinateur dans un langage parlé (il y a deux contextes ici : le langage naturel parlé ou langage naturel écrit c'est-à-dire tapé sur un clavier) tels que : Créole, Espagnol, etc. Et on attend de l'ordinateur qu'il soit capable en retour d'interpréter la langue humaine et donner des réponses appropriées.

Pour le langage artificiel dans le contexte "homme-machine", cela signifie que l'individu communique à la machine d'une manière indirecte, par le biais des langages inventés (intermédiaires) dits : « langage de programmation ». On passe certaines instructions à l'ordinateur et il les exécute.

### 6.2- Quelques critères de compréhension des langues naturelles

Pour qu'on communique à l'ordinateur il y a certains critères de compréhension qui entrent en jeu.

1<sup>er</sup> Critère : la capacité de répondre de façon appropriée à des questions.

2<sup>ème</sup> Critère : la capacité de paraphraser, consistant à exprimer autrement la même idée.

3<sup>ème</sup> Critère : la capacité de faire des inférences, c'est-à-dire de tirer des conséquences possibles ou probables de ce qui vient d'être dit.

4<sup>ème</sup> Critère : la capacité de traduire d'une langue à une autre.

5<sup>ème</sup> Critère : la capacité de résoudre des problèmes de références (quand une même chose se présente sous une autre forme).

6<sup>ème</sup> Critère : Réussir le test de Turing (critère d'intelligence) : Si on écoute un programme parlant et une personne simultanément que l'auditeur ne peut les différencier, c'est que le programme est intelligent.

### 6.3- Aspects spécifiques de la compréhension du langage naturel parlé

Le problème de compréhension de la parole est plus ardu que celui de la compréhension du langage naturel écrit (tapé sur un clavier); ceci est dû à plusieurs facteurs dont les principaux sont :

- α) Le signal résultant de l'émission de la parole comportant des bruits, ce sont des non-informations à éliminer.
- β) La prononciation des locuteurs n'est pas parfaite et varie d'un individu à un autre.
- γ) La prononciation d'un locuteur donné est variable suivant son état (psychologique, physiologique).
- δ) La prononciation d'un phonème voire, d'un mot, n'est pas la même, suivant que ce dernier est isolé ou en contexte.
- ε) La frontière entre les mots successifs d'une phrase n'apparaît pas sur le signal vocal, il peut très bien y avoir des intervalles de silence au milieu d'un mot et aucun intervalle entre deux mots successifs.

Il résulte de tout cela qu'un système de compréhension de parole reçoit en entrée un treillis d'éléments phonétiques correspondant à la transcription de la phrase prononcée. Ce treillis contient de nombreuses erreurs ; actuellement on estime le taux d'erreurs à 30% dans le meilleur des cas.

#### 6.4- Les premières tentatives de traduction automatique des langues naturelles

Les premières idées visant à faire des programmes de traduction d'une langue à une autre ont vu le jour en 1946 au cours de discussions entre Warren Weaver et Donald Booth; ces tenants pensaient que les méthodes utilisées pour découvrir les codes secrets (basées sur des tables de fréquence de lettres et de mots), pouvaient s'appliquer à la traduction d'une langue à une autre. Les difficultés qui les imposaient à l'époque, étaient l'incorporation de dictionnaires complets des deux langues, ainsi que le sens multiple des mots (polysémie) et l'ordre des mots est différent d'une langue à une autre. Outre qu'incorporer des dictionnaires, les chercheurs estimaient qu'ils n'étaient pas encore en mesure de faire la traduction mot à mot, car le premier programme de traduction russe en anglais donnait en sortie une liste de réponses possibles sans décider quelle était la traduction exacte du mot demandé. Ce qui provoquait un abaissement d'intérêt pour la traduction automatique vers les années 1960 après une quinzaine d'années d'efforts infructueux.

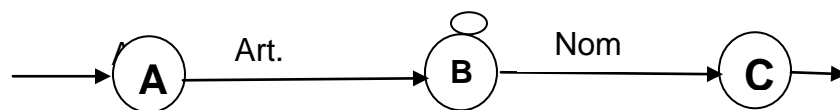
#### 6.5 - Des techniques orientées automates

Dans ce même ordre d'idées, nous allons considérer un chercheur appelé « Schank » qui a élaboré des processus cognitifs en vue de la compréhension d'une phrase. Alors, on utilise des structures en interne représentables dans un ordinateur en vue de traduire des phrases du langage naturel.

Considérons les phrases suivantes :

1) Le joli petit cheval 2) Le cheval 3) La jolie petite fille 4) La belle femme

Nous allons représenter un automate fini pour un groupe nominal simple ; lequel permettra de reconnaître les phrases susmentionnées.



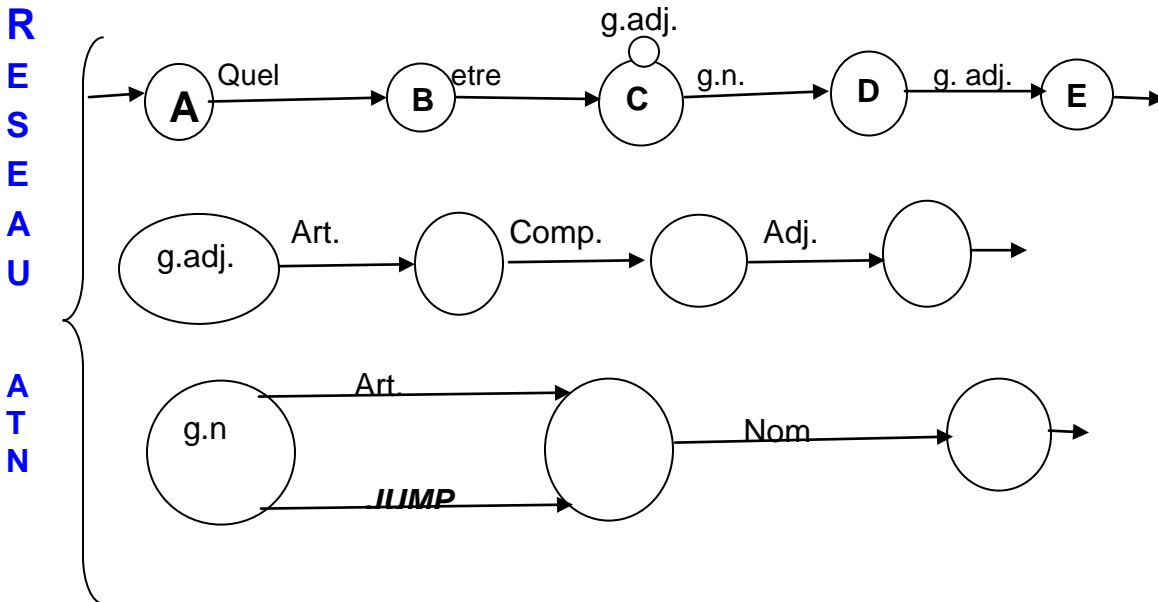
A noter que cet automate ne permet pas de reconnaître les phrases complexes comme la suivante :

“La sœur du père du Directeur Général des succursales de Pétion ville a été à la faculté d'Infotronique d'Haïti.”

William Wood a élaboré des réseaux (ATN= réseau de transition augmentée) permettant non seulement de pouvoir représenter la structure d'une telle phrase, mais également de considérer l'aspect contextuel.

L'ATN suivant permet de spécifier sur les arcs non seulement des symboles terminaux (qui donnent accès direct aux mots du langage à savoir : Verbes, Nom, adj. Art., comp.), mais aussi des symboles non-terminaux (soit : g.n, g.adj.) définis eux-mêmes par un autre automate pouvant faire référence à l'automate dont il fait partie. Il permet ensuite de sauter d'un état à un autre sans consommer le mot d'entrée et quand on a besoin de récupérer le début d'une phrase après interruption ou par souci grammatical (pour les accords grammaticaux), on utilise l'arc de type « JUMP » également de William Wood.

La figure suivante présente un ATN qui reconnaît certaines phrases :



Ce réseau permet de comprendre les questions suivantes :

- 1) Quel est le bateau le plus long ? 2) Quel est le plus long bateau ?
- 3) Quels sont les bateaux les plus longs? 4) Quels sont les plus longs bateaux ?

(Réf. ALAIN Bonnet, *L'intelligence artificielle, Promesse et Réalités, Quelques techniques d'analyse des langues naturelles*, Paris, interEditions, 1984, pp.52-56)



## Troisième partie

# Traduction de texte en son

### Introduction

Nous vivons dans un environnement où le langage est omniprésent sous la forme de stimuli acoustiques ou visuels. Chaque jour, nous ressentons dans l'histoire de l'homme, que le savoir lire et écrire pour tous est devenu un objectif il y a seulement quelques générations, et que cet objectif est loin d'être entièrement atteint. Même quand nous nous souvenons du fait que l'écriture est récente et que le savoir lire et écrire ne s'est répandu que tardivement, nous ne doutons pas qu'ils sont là pour de bon. Traitons sans doute plus de texte écrit que de parole. Nous avons tendance à donner autant de valeur à notre capacité de lire et d'écrire qu'à des capacités perceptives et motrices plus

fondamentales. Nous concevons l'écriture comme essentielle à la réalisation de soi. Nous oublions facilement que l'écriture est une invention.

## **Langage et écriture**

Il y a un contraste frappant entre l'acquisition du langage et celle de l'écriture. Dans des conditions ordinaires, le langage est spontanément acquis par le très jeune enfant. Le soutien pédagogique des adultes (qui fait presque entièrement défaut dans certaines sociétés) joue au mieux un rôle marginal. L'acquisition de l'écriture et de la lecture en revanche consiste en un processus long et intensif d'entraînement délibéré en interaction avec un maître. Est-ce parce que les systèmes d'écriture sont plus complexes que les langues parlées ? C'est le contraire. Des langues comme l'anglais, l'amharique ou le chinois sont des objets bien plus complexes qu'un système d'écriture alphabétique, syllabique ou même logo graphique. En fait, les linguistes ne sont pas encore parvenus à donner une grammaire complètement explicite d'aucune langue humaine, tandis que les systèmes d'écriture comportent, eux, des règles entièrement explicites. Cette remarquable différence entre l'acquisition du langage d'une part et de l'écriture d'autre part tient à des prédispositions psychologiques : les humains sont prédisposés à acquérir spontanément la langue de leur communauté. Ils n'ont aucune prédisposition à acquérir l'écriture. Ce sont les systèmes d'écriture qui ont dû s'adapter plutôt à des dispositions perceptives et motrices qui avaient émergé bien avant l'invention de l'écriture. Comment se fait-il, dans ces conditions que des systèmes d'écriture soient apparus et se soient répandus et stabilisés ?

L'accumulation et la diversification de textes écrits, rendues possibles par le développement de l'écriture et les transformations économiques et politiques associées, ont rendu les coûts impliqués dans l'acquisition de l'écriture et de la lecture inférieurs aux bénéfices de cette acquisition pour une proportion croissante de la population. Dans les sociétés modernes, les bénéfices sont plus grands que les coûts pour la majorité, et l'analphabétisme est devenu quelque chose de honteux et donc un coût en lui-même.

Il y a un facteur important qui contribue à expliquer la généralisation de l'écriture et de la lecture. Une fois le savoir-faire véritablement acquis, écrire devient un genre d'automatisme : on peut écrire sans prêter aucune attention consciente au mouvement de sa main (et cela est vrai aussi de la dactylographie). De même, pour le vrai lecteur, la lecture est un genre de reconnaissance visuelle automatique de formes parmi d'autres. Les premières formes d'écriture, comme l'écriture cunéiforme des Sumériens, avec ses matériaux et ses instruments encombrants, ne permettaient pas le même degré d'automatisme.

Deux faits expliquent ainsi le succès de l'écriture : le fait que les bénéfices soient devenus, pour un

nombre croissant de personnes, plus grands que les coûts, et le fait que, une fois payé les coûts initiaux d'acquisition du savoir-faire, les coûts d'utilisation du savoir-faire acquis soient comparativement négligeables. Ces deux faits sont liés. Si la distribution des coûts et des bénéfices était plus également répartie dans le cours de la vie ou, en d'autres termes, si les coûts marginaux de l'écriture et de la lecture ne s'effondraient pas une fois la compétence vraiment acquise, on lirait et on écrirait beaucoup moins (telle était d'ailleurs la situation quand on écrivait sur la pierre ou la glaise). Avec un usage moins fréquent de la lecture et de l'écriture, il y aurait moins d'écrits à lire et moins de personnes disposées à les lire. Les bénéfices de l'écriture et de la lecture en seraient d'autant diminués, et pourraient ne plus en justifier les coûts, si ce n'est pour un petit groupe de scribes professionnels. Ce qui se passe en fait, c'est qu'une fois maîtrisées, l'écriture et la lecture sont généralement profitables. Plus grand est le nombre de gens qui lisent et écrivent, plus il est avantageux de le faire soi-même, et plus grande est la motivation pour faire apprendre le savoir-faire à ses enfants. Dans ces conditions, comment l'avenir de l'écriture et de la lecture pourrait-il être incertain ? C'est, qu'écrire n'est pas la seule façon de produire des écrits.

Il peut être avantageux de dicter pour des raisons de vitesse, ou ce peut être une nécessité comme dans le cas de personnes ayant des problèmes visuels. Cependant, même si nous avons le choix, la plupart d'entre nous préféreraient écrire que dicter. La raison principale de cette préférence est le fait qu'on a un bien moindre contrôle de son texte en le dictant qu'en l'écrivant. De toute façon la dictée traditionnelle était une forme de division du travail scriptural, et non une façon de rendre l'écriture obsolète. Aujourd'hui, en revanche, les nouvelles technologies de l'information sont en passe de permettre une nouvelle forme de dictée ne souffrant pas des mêmes défauts que l'ancienne, et telle que la division du travail ne se fera plus entre employeur et employé mais entre humains et machines.

## Chap.7 *La conversion du texte à la parole*

La conversion du texte à la parole est un moyen de se faire lire un texte à voix haute plutôt que de le lire soi-même. De même que des personnes riches ou puissantes (ou des personnes handicapées) se sont servies de secrétaires pour dicter plutôt qu'écrire, de même des personnes riches (ou

handicapées) ont-elles fait appel à des lecteurs ou des lectrices payés. Le ton de la voix du lecteur contribue à la façon dont on interprète le texte. Dans certains cas - un acteur lisant un poème ou une mère lisant une histoire à son enfant, cela peut être merveilleux. Mais, en général, nous préférons interpréter ce que nous lisons dans notre propre voix silencieuse. En outre, nous ne nous éprendrons peut-être jamais vraiment du ton de voix d'un ordinateur, et nous resterons sans doute justement réticents à être influencés par lui dans notre interprétation d'un texte.

Dans le processus d'écoute de la parole (qu'il s'agisse d'une conversation ou d'une lecture à voix haute), l'information fournie par chaque parole doit être retenue en mémoire à court terme assez longtemps pour permettre le décodage linguistique (même si une partie de cette information peut être reconstruite à partir du contexte). Il n'en va pas de même avec la lecture. Le texte écrit fournit une mémoire à court terme externe efficace qui peut être parcourue dans les deux sens. C'est ce qui permet au lecteur de suivre un texte à son propre rythme par opposition à une écoute qui se fait au rythme du locuteur. Le lecteur peut d'abord parcourir le texte pour ensuite le lire attentivement. Il peut revenir en arrière sur un passage dont la pertinence lui est apparue rétrospectivement. Il peut vérifier la cohérence du texte. Quand on lit, on perd le supplément d'information donné par la voix et les gestes, mais on est à même de tirer bien plus d'un texte et de le comprendre plus à fond.

D'un point de vue pratique, l'écoute d'un texte est bien plus lente que la lecture. Elle est aussi plus bruyante (mais ceci se corrige facilement au moyen d'écouteurs). Le plus grand obstacle à l'abandon de la lecture, c'est sans doute le rôle qu'elle joue, non dans l'utilisation des textes mais dans leur production. Ce que nous apprécions le plus, et à juste titre, dans l'activité d'écriture, ce ne sont pas les mouvements de la main (sinon la dactylographie n'aurait pas remplacé l'écriture manuelle à ce point), mais c'est le fait que nous pouvons lire au fur et à mesure ce que nous écrivons.

Tout ceci rend peu plausible que l'effet cumulatif de décisions individuelles en la matière entraîne le remplacement, à l'échelle de la société, de l'activité de lecture par l'emploi des techniques de conversion du texte à la parole.

Néanmoins on accorde un certain intérêt à l'activité de conversion de texte en parole. Ceci par le fait du développement rapide des nouvelles technologies, les ordinateurs personnels devenus plus puissants ont permis de réaliser jusque-là des applications intéressantes en ce sens. En témoignent les nombreux produits de ce type existant sur le marché.

## Chap.8 *Analyse des systèmes existants*

Les programmes de reconnaissance de la parole qui permettent la conversion de la parole en texte se sont rapidement améliorés au cours des dernières années. Ils permettent désormais de parler de façon continue et naturelle au rythme de la conversation et de voir ce que l'on dit s'inscrire au fur et à mesure sur l'écran. Actuellement, le taux d'erreurs est encore trop élevé, les programmes doivent eux-mêmes être entraînés et, bien des utilisateurs qui ont un besoin pressant de ce genre de programme se découragent. Il est évident cependant que ces défauts seront surmontés et que, dans quelques années, il sera possible de parler normalement et que la machine transcrira ce qui est dit avec très peu d'erreurs, et en distinguant, dans le flot de la parole, les instructions qui doivent être exécutées (par exemple : « souligner ! ») du texte proprement dit qui doit être transcrit. Il deviendra bien plus facile de dicter à une machine. De façon plus générale il sera plus facile de donner des ordres à un ordinateur (et à divers genres d'appareils domestiques, de véhicules, et d'autres machines) en leur parlant qu'en manipulant claviers, souris et autres boutons. Les machines seront capables de donner de l'information oralement plutôt que sur un écran. Grâce aux progrès de la technologie de conversion de texte en parole, les machines seront capables de lire à voix haute un texte écrit avec des intonations naturelles. Les interactions orales en langue naturelle seront la règle plutôt que l'exception.

Aussi imparfaites que soient à présent ces technologies de conversion de la parole au texte et du texte à la parole, elles transforment déjà la vie de personnes qui, souffrant d'un handicap visuel, auditif ou moteur, ou de dyslexie, ne peuvent pas lire et écrire normalement. Si les millions d'analphabètes dans le monde ne profitent pas de ces mêmes technologies, c'est bien sûr à cause de leur pauvreté, laquelle explique, pour commencer, qu'ils sont analphabètes.

Dans peu de temps, les coûts et les bénéfices de l'écriture et de la lecture seront comparés non seulement avec ceux de l'analphabétisme, mais aussi avec ceux d'autres modes de création et de consultation de textes issus des nouvelles technologies. Avenir prometteur ...



## Quatrième partie

# Présentation du logiciel *KREYOSPEECH*

Le logiciel est totalement réalisé en *java*. Il est supporté par une couche logicielle qui est venue se greffer sur les packages déjà existant du compilateur *J2SDK1.4.2* fourni par la société *SUN MICROSYSTEMS*.

La couche logicielle comprenant le synthétiseur (*JSAPI 1.0*) permettant la réalisation du travail, est obtenue sous forme de licence par la société *GLOUD GARDEN*.

## Chap. 9 *Vue d'ensemble de l'application*

### *Aperçu*

9.1 Problématique

9.2 Objectifs

9.3 Limites

9.4 Digramme de conception

9.5 Principales interfaces (écrans) de l'application

### 9.1 Problématique

Dans le but d'améliorer la communication homme-machine, on s'est assez vite intéressé, dans la deuxième moitié du vingtième siècle, au problème de la reconnaissance vocale. En 1959 est mis au point un système monocuteurs de reconnaissance des chiffres. Mais la difficulté du problème dans sa généralité a été sous-estimée: si un programme ARPA avait été lancé dans les années soixante-dix aux États-Unis, il faut attendre 1997 pour qu'IBM propose le premier logiciel de dictée vocale, et nombreux sont encore les écueils inhérents à la reconnaissance vocale: un processus complexe de

la production de la parole à la détermination de l'orthographe, le bruit de fond, des différences parfois importantes suivant les accents et les dialectes au sein d'une même langue, l'incapacité de traiter le bilinguisme, situation commune dans certains pays, ainsi que des besoins très importants en calcul, rendant difficile la reconnaissance (Synthèse vocale ou reconnaissance proprement dite) en temps réel.

## 9.2 -Objectifs

Bien qu'Haïti accuse d'un certain retard pour ce qui est de l'acquisition et de la maîtrise des TICs, ces dernières n'en sont pas moins présentes et jouent comme dans tous les pays où elles s'implantent, un rôle de plus en plus prépondérant dans la vie des Haïtiens.

Des technologies de tout genre apparaissent sous des formes diversement variées, et touchent presque toutes les catégories sociales améliorant bien souvent leurs moyens de communication.

Le créole étant, d'après le dictionnaire *Wikipedia* du site Internet du même nom, la seule langue parlée par 90% de la population d'Haïti, Le besoin de leur fournir des produits présentés en créole se fait vite sentir.

Le projet *KREYOSPEECH* consiste à développer, non un synthétiseur du créole haïtien, mais une plateforme de synthèse vocale pour le créole haïtien.

*KREYOSPEECH* entend, par sa version initiale, présenter une plateforme de synthèse vocale dotée de fonctionnalités permettant aux utilisateurs de convertir un texte créole écrit en sons. Il donnera ainsi aux utilisateurs la possibilité d'écouter le texte en "version" audio. Il permettra donc de convertir le texte saisi sous format de fichier WAV, AIF ou AU.

Les versions futures se pencheront sur l'opération inverse : *reconnaissance vocale adaptée au créole haïtien*. Ce qui permettra de produire un texte écrit à partir de la voix.

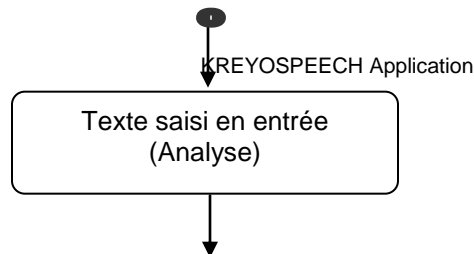
Il est intéressant de noter qu'à partir de *KREYOSPEECH* on pourra créer des fichiers sous format audio ou texte.

## 9.3- Limites

*KREYOSPEECH*, ne reconnaîtra sûrement pas tous les mots du créole, car seul un échantillon limité a servi dans la base d'exemples utilisée.

Cette limite n'est en aucune manière un obstacle, car la base d'exemples peut toujours être enrichie.

## 9.4- Diagramme de conception du logiciel



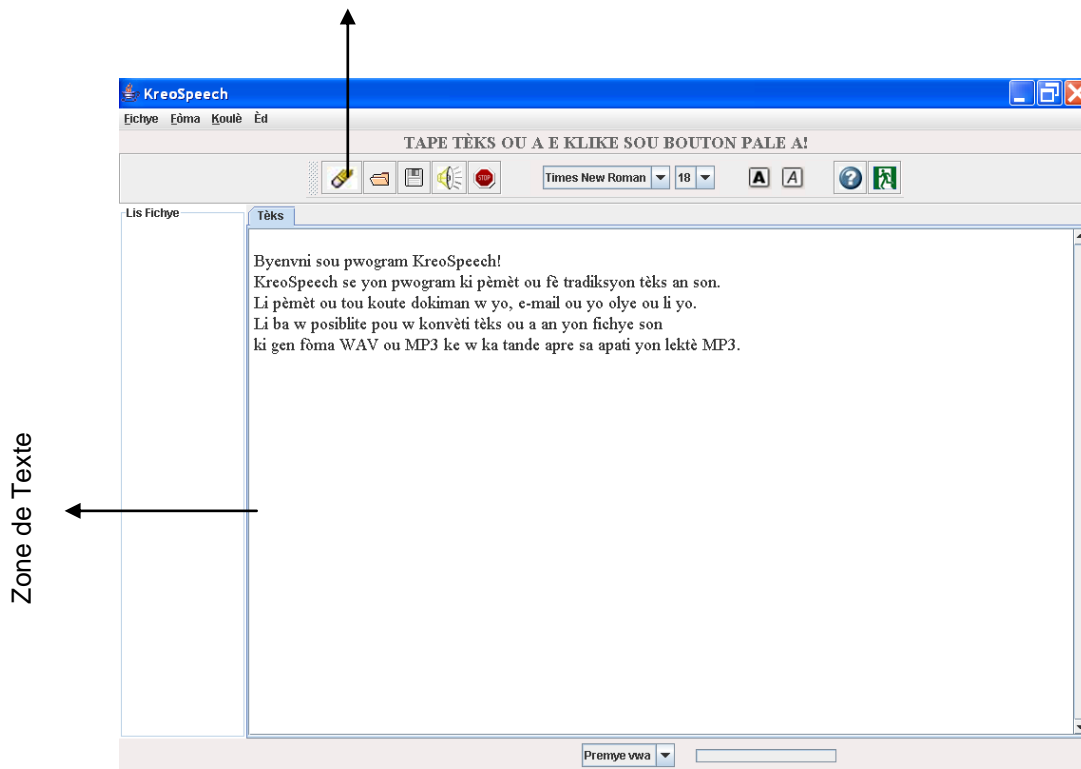


end

### 9.5- Principales interfaces (écrans) de l'application

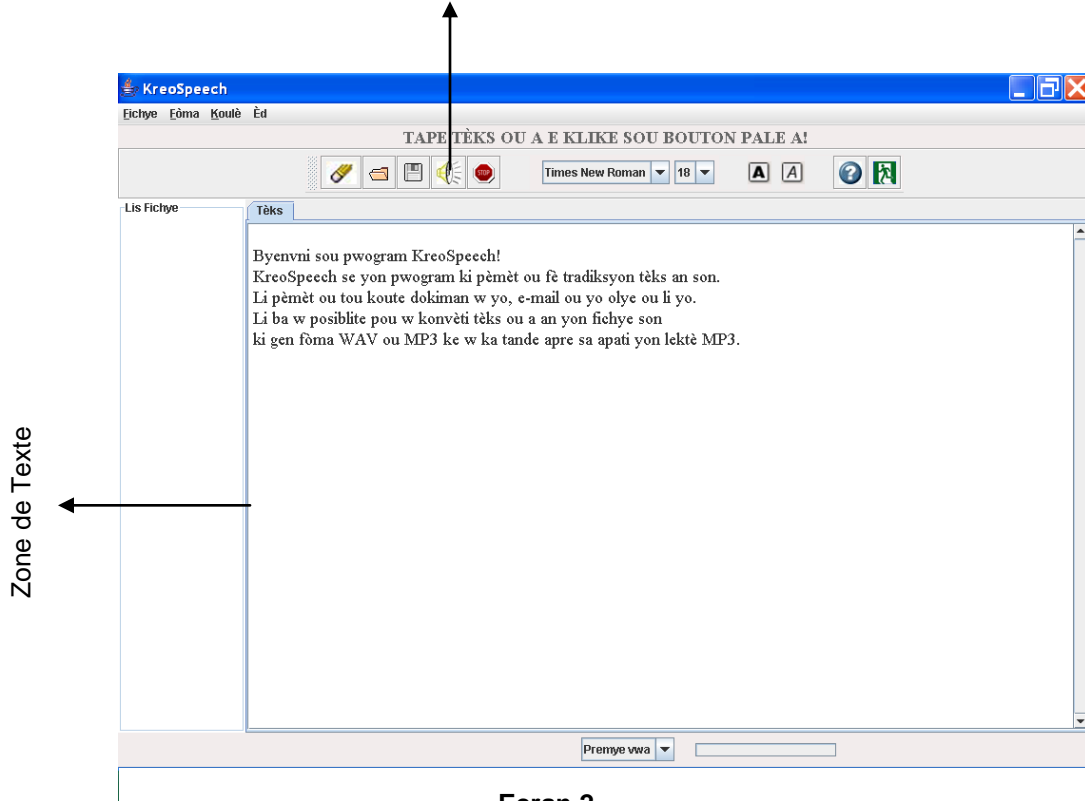
La suivante présente l'interface de travail. Celle-ci est obtenue une fois que l'application soit lancée.

Ce bouton, appelé « **Netwaye** » permet de nettoyer la zone de texte.

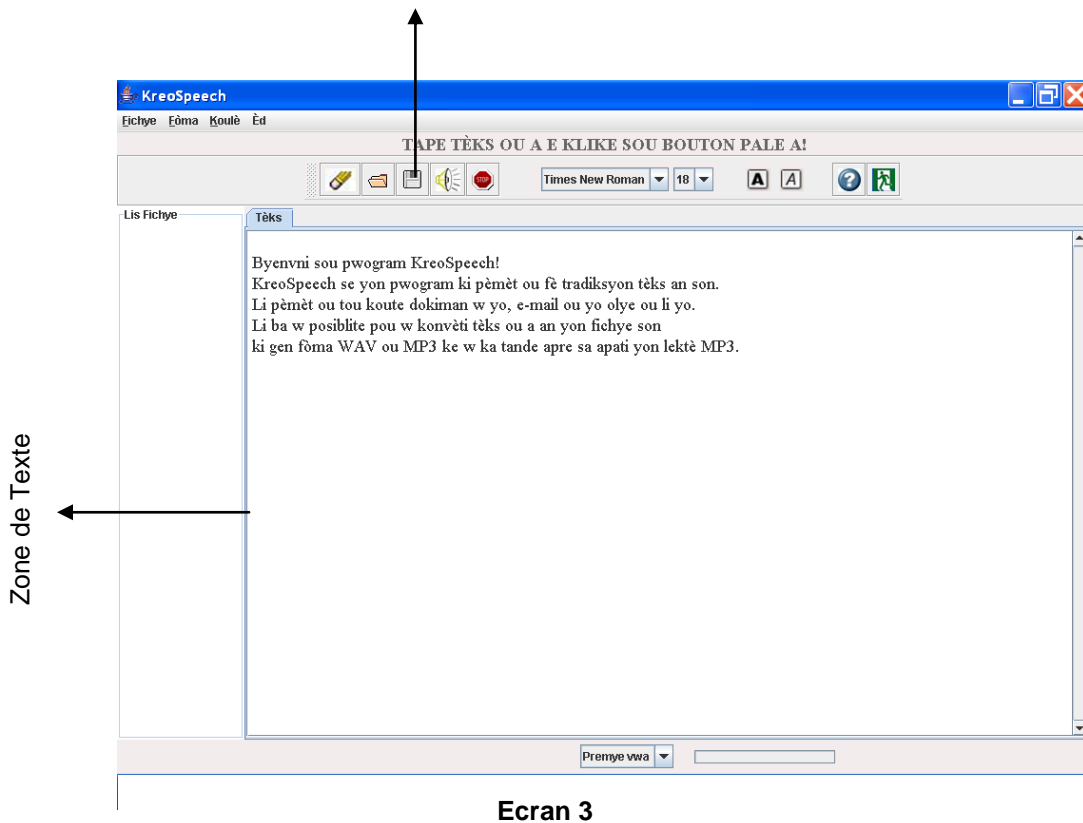


Ecran 1

Le bouton **“Pale”** permet à l’ordinateur de lire le texte en créole. Il permet aussi la création d’un fichier Audio (format AIF, WAV, AU).



Le bouton **“anrejistre”** permet de faire la sauvegarde du texte en créole.



Ceci permet l’ouverture d’un fichier texte.



## Cinquième partie

### **Conclusion et perspectives d'avenir**

L'objectif de notre travail a été de mener une étude dans le domaine de la synthèse de la parole à partir du texte du créole haïtien standard. Cela consistait en l'élaboration d'une plateforme de synthèse vocale pour le créole haïtien. Nous avons donc été amené à faire, dans un premier temps une étude phonologique du créole haïtien- Étude qui nous a permis de bien comprendre le rapprochement linguistique (du point de vue lexical) étroit qui existe entre le créole d'Haïti et le français.

La deuxième partie du travail s'était portée sur une analyse de quelques synthétiseurs de parole et des principales techniques de synthèse de la parole, dont la méthode de synthèse par règles et la méthode de synthèse par concaténation. Cette dernière analyse s'était révélée très utile, en ce sens qu'elle nous a permis de comprendre que des adaptations pourraient être apportées en vue de permettre à un synthétiseur donné de produire des résultats pour une autre langue que celle pour laquelle il a été développé. De ce fait, FREETS nous a beaucoup aidé et nous a permis de parvenir aux résultats escomptés ; il faut aussi mentionner que le langage de programmation Java nous a un peu facilité la tâche, compte tenu du fait qu'avec ce dernier les adaptations algorithmiques sont plus faciles à faire. ..

KREYOSPEECH, étant ce synthétiseur de parole du créole haïtien, la voie dans ce domaine est donc restée ouverte aux chercheurs qui sont intéressés par la production de logiciels en créole haïtien. Il peut aussi être utilisé en vue de permettre la réalisation de véritable synthétiseur pour le créole haïtien (qui n'a aucune dépendance par rapport à un autre). Et KREYOSPEECH peut même trouver une application intéressante dans le domaine de la réalité virtuelle. Il pourra finalement être utilisé dans l'élaboration d'applications automatiques. Avenir prometteur !

### **Petit lexique, pour une compréhension facilitée du document**

**Algorithme:** Ensemble de règles opératoires propres à un calcul.

**Articulateur:** Partie mobile de la bouche: Langue, mâchoire, pharynx.

**Bits** : Débit d'information élémentaire.

**Corpus** : Ensemble limité des éléments sur lesquels se base l'étude d'un Phénomène linguistique

**Diphonème** : (ou Diphone ou phonatome): Segment de parole qui s'étend de zone stable d'un autre phonème et qui contient en son centre la zone de transition entre deux sons successifs.

**Formant** : maximum de la courbe de réponse en fréquence du conduit vocal pour son déterminé.

**Logatome:** mot sans signification constitué d'une consonante, d'une voyelle et d'une consonante par exemple " STIC, TOK, BIF " etc.

**MIC:** modulation par impulsions codées: technique permettant de convertir un signal électrique analogique en une suite de bits en vue de sa transmission sur le réseau numérique téléphonique.

**N-MOS:** Technologies Metal-Oxyde Semi-conducteur canal N utilisée dans la conception des circuits intégrés.

**Phonème:** unité élémentaire des sons du langage (voyelles et consonnes).

**Prédiction linéaire:** la majeure partie des sons de la parole est de type quasi-stationnaire, leur évolution n'est donc pas de nature purement aléatoire. Par conséquent il est possible de prévoir la valeur du signal à partir de la combinaison linéaire d'un certain nombre d'échantillons précédents.

**Prosodie:** Phénomènes accentués, rythmiques et intonatifs, qui se manifestent par l'intermédiaire de trois paramètres: durée, intensité et fréquence fondamentale.

## Bibliographie

- .....
- 1 SUZANNE, Sylvain, *Le créole, Morphologie et Syntaxe*, Port-au-Prince, 1936.
  - 2 *The Coming Age of Talking Computers, the Futurist*, December 1999.
  - 3 Cote Part-Dieu006.454REC, *Reconnaissance automatique de la parole*, pp. 11, 12, 40-55.
  4. MacGraw-Hill encyclopedia cote603/ANGMcG (XVII), *les Techniques de l'ingénieur*, pp. 4-5, Vol. H1.940
  - 5 cote Part-Dieu 006.454 REC ; Science et Vie Micro n°128.
  - 6 *Reconnaissance automatique de la parole " cote Part-Dieu 006.454 REC p.34- 36*
  - 7 **Encyclopedia** cote 603 / ANG McG (XVII) p. 233, 234 ;( ref page)
  - 8 CHAUDENSON, Robert, *La créolisation, théorie, applications, implications*, Paris, Éditions L'Harmattan, 2003, p. ...
  - 9 MANESSY, Gabriel, *Créoles pidgins, variétés véhiculaires*, Paris, Éditions CNRS, 1995.
  10. YACOUB, Joseph, «Les minorités en Amérique latine et aux Caraïbes», IN *Les minorités dans le monde*, Paris, Desclée de Brouwer, 1998, (pp. 781-805) P.900.
  - 11 BICKERTON, Derek, *Language\_and Species*, Chicago, University of Chicago press, 1990.
  - 12 LASSAGNE, François, «Créole, la naissance d'une langue» dans *Science et Vie*, Paris, no 227, hors série juin 2004, p. 78-85.
  - 13 ALAIN Bonnet, *L'intelligence artificielle, Promesse et Réalités*, Paris, interEditions, 1984, p.271
  - 14 [festival2001] Black, A., Taylor, P., and Caley, R., 2001. "The Festival Speech Synthesis System, Version 1.4.2." Document disponible à <http://www.cstr.ed.ac.uk/projects/festival.html>.
  - 15 [flite2001] Black, A., and Lenzo, K., 2001. "Flite,"  
Version 0.91. Unpublished source code available via <http://www.cmufllite.org/>.

- 16** [brieman84] Breiman, L., Friedman, J. H., Olshen, R.A., and C. J. Stone. 1984. "Classification and Regression Trees." Wadsworth, Belmont, CA.
- 17.** [T. Dutoit ] T. Dutoit : "« Je parle, donc je suis ? »: un bilan des développements récents en traitement automatique de la parole. " Faculté Polytechnique démons

## URLs

- 1** [http://WWW.vie\\_artificielle.com](http://WWW.vie_artificielle.com)
- 2** <http://dkc1.digikey.com/FR/FR/PDF/FR061/M6.html>
- 3** <http://jacbayle.club.fr/livres/archCreol/Valdman.html>
- 4** [http://www.gotronic.fr/catalog/kits/kitsframe.php?page\\_cible= parole.htm](http://www.gotronic.fr/catalog/kits/kitsframe.php?page_cible= parole.htm)
- 5** [http://www.cavi.univ-paris3.fr/ilpga/ilpga/tal/sitespp/slmd2-0304/SiteTal\\_Waterlot\\_Amandine/recvoc.html](http://www.cavi.univ-paris3.fr/ilpga/ilpga/tal/sitespp/slmd2-0304/SiteTal_Waterlot_Amandine/recvoc.html)
- 6** <http://www.speechware.be/french/Tips.php#b>
- 7** <http://freetts.sourceforge.net/docs/index.php>
- 8** [http://sourceforge.net/tracker/?group\\_id=42080&atid=432074](http://sourceforge.net/tracker/?group_id=42080&atid=432074)
- 9** <http://www.scansoft.com/viavoice/>
- 10** <http://www.speech.cs.cmu.edu/flite/>
- 11** <http://cepstral.com/>
- 12** <http://freetts.sourceforge.net/docs/ProgrammerGuide.html>