HAL
open science

# Dictée automatique de textes à haute voix : analyse de corpus et modélisation des stratégies d'énonciation

Mejdi Slim

▶ To cite this version:

Mejdi Slim. Dictée automatique de textes à haute voix : analyse de corpus et modélisation des stratégies d'énonciation. Sciences de l'Homme et Société. 2014. ⟨dumas-01062033⟩

HAL Id: dumas-01062033

https://dumas.ccsd.cnrs.fr/dumas-01062033v1

Submitted on 9 Sep 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Autorisation HAL

HAL Authorization

# Dictée automatique de textes à haute voix : Analyse de corpus et modélisation des stratégies d'énonciation

**Nom : SLIM**

**Prénom : Mejdi**

---

Mémoire de master 2 recherche – 30 crédits – Mention Sciences du langage

Spécialité : Industrie de la Langue

Parcours : Traitement Automatique de la Langue Écrite et de la Parole

Sous la direction de M. **Gérard BAILLY** & M. **Georges ANTONIADIS**

Effectué au GIPSA-Lab

Année universitaire : 2013-2014

*« Tout début est difficile mais sans début, pas de fin. »*
*[Proverbe allemand]*

# DÉCLARATION ANTIPLAGIAT

## DÉCLARATION

1. Ce travail est le fruit d'un travail personnel et constitue un document original.

2. Je sais que prétendre être l'auteur d'un travail écrit par une autre personne est une pratique sévèrement sanctionnée par la loi.

3. Personne d'autre que moi n'a le droit de faire valoir ce travail, en totalité ou en partie, comme le sien.

4. Les propos repris mot à mot à d'autres auteurs figurent entre guillemets (citations).

5. Les écrits sur lesquels je m'appuie dans ce mémoire sont systématiquement référencés selon un système de renvoi bibliographique clair et précis.

Nom : SLIM Prénom : Mejdi

Date : 26/08/2014 Signature :

*« Nous pensons que le langage a une essence, et c'est cette essence que nous essayons de définir, de décrire, de modeler. Mais comme Dieu ou le soleil de Platon, l'essence linguistique n'est pas quelque chose que nous pouvons regarder droit dans les yeux. Nous n'avons pas d'accès direct à cette essence, mais nous pouvons l'approcher en observant la grammaire des langues naturelles du monde. »*
*[Goldsmith, 1996]*

**Mots-clés :** synthèse de dictées, synthèse de parole, contours prosodiques, apprentissage automatique, réseaux de neurones.

# RÉSUME

Les travaux présentés dans ce mémoire se situent dans le cadre de la synthèse de la parole à partir du texte. Notre travail vise à concevoir un système permettant de dicter automatiquement à voix haute un texte orthographié à des élèves de niveau CM1-CM2. Ce système comporte quatre modules : 1) un module d'analyse morphosyntaxique, qui lève une structure arborescente à partir de la chaîne orthographique ; 2) un module de marquage prosodique qui découpe cet énoncé en groupes prosodiques et les articule entre eux ; 3) un module de génération de prosodie, qui calcule automatiquement la prosodie (contour de fréquence fondamentale et contour rythmique) à appliquer sur ces groupes en fonction des diverses marques ; 4) un module de synthèse qui convertit l'ensemble de ces données segmentales et suprasegmentales en signal acoustique. Partant de la sortie d'un analyseur syntaxique à l'état de l'art, notre travail a consisté tout d'abord à projeter toute ou partie de la structure syntaxique de chaque texte, afin de marquer le texte et les reprises de ses groupes de mots par un ensemble restreint de marqueurs. Nous avons ensuite paramétré un modèle de génération de la prosodie SFC « **S**uperposition of **F**unctional **C**ontours » développé par Bailly et Holm (2003) à l'aide du corpus de dictées. Nous avons finalement évalué la qualité de la synthèse produite par un système de synthèse disponible au GIPSA-Lab. Nous montrons que les générateurs de contours indexés par la version d'énonciation arrivent à capturer le ralentissement du débit, l'accentuation du découpage des mots et des syntagmes ainsi que les modifications des contours mélodiques des quatre versions de chaque groupe de mots.

**Keywords :** automatic dictation, speech synthesis, prosodic contours, machine learning, neural networks.

# ABSTRACT

The work presented in this paper is in the context of Text-To-Speech synthesis (TTS). This work aims to develop an automatic dictation system that converts written text into spoken words, for use by CM1 and CM2 students (4th and 5th grade in French primary school). This system includes four modules : 1) a morpho-syntactic analysis module which extracts tree data structure from orthographic strings; 2) a prosodic segmentation module which divides the utterances into prosodic groups and relates them to their propositional content; 3) a prosody generation module which calculates prosody automatically (fundamental frequency contour and speech rhythm), to be applied to those groups based on phonotactic features; 4) a synthesis module which converts this segmental and supra-segmental data into an acoustic signal. The syntactic parsing is first performed. The syntactic tree is then further projected onto the syntagmatic axis : sentence-internal prosodic markers are supposed to cue the dependency relations between adjacent constituents. Prosody generation is then performed by associating these markers with multiparametric contours via so-called contour generators. These contour generators are implemented as feed-forward neural networks and trained thanks to an iterative analysis-by-synthesis process (see description of the SFC "Superposition of Functional Contours" model in Bailly and Holm *(2003)* using a corpus of dictations. We finally assessed the quality of the modeling with a synthesis system available at GIPSA-Lab. We showed that strategy-specific contour generators are able to capture the slow rate of speech flow, the increased rate of word and phrase segmentation as well as changes in the melodic contours of the four versions for each phrase.

*« Le meilleur ami de « merci » est « beaucoup ». »*
*[Michel Bouthot]*

# REMERCIEMENTS

Le rapport, que je vous propose ici, est le fruit d'innombrables rencontres et de nombreux échanges. L'aventure n'aurait pas été aussi passionnante, étrange et constructive sans toutes les propositions qui m'ont été apportées par tant de chercheurs, femmes ou hommes, que je veux remercier, très sincèrement, ici :

Gérard BAILLY, mon directeur du mémoire, principal interlocuteur pour tout ce qui se trouve sur ces pages. Je lui suis reconnaissant de m'avoir apporté son enthousiasme et sa curiosité durant les conversations scientifiques. Il a été le seul à me soutenir dans les moments difficiles de l'enfantement de ce mémoire, je n'oublierai jamais la phrase qu'il prononçait si souvent : *« Mejdi, on va y arriver… »*. Je souhaite que ce rapport puisse refléter l'expression de sa finesse d'appréciation et de sa capacité de réflexion, mais aussi sa grande rigueur. Après ces six mois de « vie commune », finalement, l'histoire s'achève. L'expérience a été si enrichissante, tant au niveau scientifique qu'au niveau relationnel. Merci d'avoir accepté de me prendre en stage de master 2.

J’ai aussi ici à cœur de remercier Georges ANTONIADIS, le responsable de la spécialité de mon master, d’avoir accepté d’être le co-encadreur du ce mémoire.

Je tiens à remercier Véronique AUBERGÉ qui a accepté la charge de rapporter mon travail de mémoire.

Je souhaite également remercier mes professeurs : ANTONIADIS Georges, AUBERGÉ Véronique, DEGACHE Christian, DINI Luca, HATIER Sylvain, JANDARD Marie-Christine, KAUFMAN Fabrice, KRAIF Olivier, LEBARBÉ Thomas, LOPEZ Cédric, MEUNIER Jean-Luc, PONTON Claude, PORTET François, ROSSATO Solange, SEGOND Frédérique, et enfin TUTIN Agnès, qui m’ont démontré, lors de mes études de master, l’apport que pouvait avoir le TAL sur la Linguistique.

Merci à l'équipe administrative de GIPSA-lab constituée de GENIN Elsa, BOUFFARD-TOCAT Lucia, NOGUEIRA Sonia et BIOUD Nadine, toujours disponibles et serviables.

Merci aux stagiaires et aux doctorants, surtout ceux qui ont participé aux tests d’évaluation de notre système, pour toutes les soirées, les Laser games et les Barbecues. Je ne me lance pas dans une liste de noms, je risque d'en oublier et ça serait dommage.

Enfin, un grand merci à ma famille en Tunisie, notamment ma fille Rawene & mon fils Slah Eddine, chez qui je ne suis pas souvent allé, depuis le début de ce mémoire et je m’en excuse. Une bise à chacun d’entre vous.

*« Un ami est une route ... un ennemi est un mur. »*
*[Proverbe chinois]*

# DÉDICACES

Je tiens simplement à dédicacer ce mémoire à mes parents, Slah & Fathia, qui m'ont nommé ***« Mejdi»***. En effet, c'est une responsabilité que je dois assumer sciemment, vu que ***« Mejdi »*** en arabe « مجـــدي » signifie ***« ma gloire »***. J'espère bien que j'ai réussi à réaliser un minimum de cette « ***gloire*** » rêvée… J'avoue que ce prénom est mon seul ami, il ne trahit jamais !

.

# CITATIONS

*« La douleur d'hier fait notre force d'aujourd'hui. »*
*[Paulo Coelho]*

*« La chute n'est pas un échec. L'échec c'est de rester là où on est tombé. »*
*[Socrate]*

*« Le succès n'est pas la clé du bonheur. Le bonheur est la clé du succès. Si vous aimez ce que vous faites, vous réussirez. »*
*[Albert Schweitzer]*

*« La vie est trop courte pour se réveiller avec des regrets. Aime les gens qui te traitent bien. Oublie ceux qui n'en font rien. Crois que tout arrive pour une raison. Si tu as une seconde chance attrape-la à deux mains. Personne n'a dit que la vie serait facile, ils ont juste promis qu'elle en valait la peine ! Alors n'abandonne surtout pas…! »*
*[David Cloutier]*

# SIGLES & ABRÉVIATIONS UTILISÉS

## -Par ordre alphabétique-

Nous listons ici les abréviations utilisées dans ce mémoire. Nous utilisons les abréviations Anglaises lorsqu'elles sont l'usage courant dans la communauté francophone.

**ALAO :** Apprentissage des Langues Assisté par Ordinateur.

**CENTAL :** Centre de Traitement Automatique du langage de l'université Catholique de Louvain

**EVA :** Empan Visuo-Attentionnel

**GAMA :** Gestes phonatoires, Analyse et Modélisation Acoustique

**GIPC :** Group Inter Perception Center

**GIPSA :** Grenoble Images Parole Signal Automatique

**HTS :** HMM-based Speech Synthesis System ou H Triple S

**IRISA :** L'Institut de recherche en informatique et systèmes aléatoires

**IRIT :** Institut de Recherche en Informatique de Toulouse

**LPC :** Linear Predictive Coding

**MAGIC :** Machines parlantes, Gestes oro-faciaux, Interaction face-à-face, Communication augmentée

**MBROLA :** Multi Band Resynthesis Overlap and Add

**PCMD :** Parole Cerveau Multimodalité Développement

**PSOLA :** Pitch Synchronous Overlap-Add

**SFC :** Superposition of Functional Contours

**SLD :** Systèmes linguistiques et dialectologie

**PoS :** *Part-of-Speech*

**SPC :** Synthèse Par Corpus

**TALN :** Traitement Automatique de la Langue Naturelle

**TD-PSOLA :** Time Domain Pitch Synchronous Overlap-Add

**TTS :** Text-To-Speech synthesis

# Table des matières

RÉSUME ..... 6
ABSTRACT ..... 7
REMERCIEMENTS ..... 8
DÉDICACES ..... 10
CITATIONS ..... 11
SIGLES & ABRÉVIATIONS UTILISÉS ..... 12
TABLE DES ILLUSTRATIONS ..... 17
PRÉAMBULE ..... 21
INTRODUCTION ..... 23
PLAN DU MÉMOIRE ..... 25
PARTIE 1 : ÉTUDE THÉORIQUE ..... 26
1 ÉTAT DE L'ART ..... 27
1.1 La dictée ..... 27
1.1.1 Caractéristiques propres à la langue ..... 27
1.1.2 Stratégies existantes ..... 29
1.1.3 Les difficultés ..... 29
1.2 Logiciels pour la dictée ..... 31
1.2.1 Ordictée ..... 31
1.2.2 DICTOR ..... 31
1.2.3 PlatON ..... 32
1.2.4 La dictée interactive ..... 35
1.2.5 Commentaires ..... 35
1.2.6 Conclusions ..... 35
2 PROBLÉMATIQUE ..... 37
2.1 Introduction ..... 37
2.2 Répondre à un besoin ..... 37
2.3 Analyser et modéliser les stratégies d'énonciation mises en œuvre en dictée orale 38
2.4 Objectif de ce mémoire ..... 38
2.5 Attentes, hypothèses ..... 39

2.6 Conclusions ..... 40
3 SYNTHÈSE AUDIO DU TEXTE AU SIGNAL ACOUSTIQUE ..... 41
3.1 Introduction ..... 41
3.2 Architecture d'un système de synthèse de parole ..... 41
3.2.1 Les traitements linguistiques ..... 43
3.2.1 Le transfert phonologique ..... 46
3.2.2 La génération ..... 48
3.2.3 Techniques de synthèse ..... 58
3.3 Conclusions ..... 61
4 MODÈLE PROSODIQUE UTILISÉ : SFC ..... 62
4.1 Introduction ..... 62
4.2 Analyse ascendante (bottom up) ..... 62
4.3 Analyse descendante (top down) ..... 63
4.4 Motivation du SFC ..... 65
4.4.1 Un lexique structuré de formes prosodiques globales ..... 65
4.4.2 Contours dynamiques et coarticulés ..... 66
4.4.3 Contours multiparamétriques ..... 67
4.4.4 Modèle SFC ..... 67
4.5 Commentaires ..... 74
4.6 Conclusions ..... 74
PARTIE 2 : ÉTUDE EMPIRIQUE ..... 75
5 CORPUS & MÉTHODOLOGIE ..... 76
5.1 Introduction ..... 76
5.2 Corpus ..... 76
5.3 Alignement et étiquetages des signaux ..... 77
5.3.1 Remarque sur la ponctuation ..... 80
5.4 Analyse des corpus ..... 80
5.5 Analyse morphosyntaxique ..... 86

5.5.1 Format du fichier de sortie ..... 86
5.5.2 La négation ..... 87
5.5.3 Épellation ..... 88
5.6 Prétraitement ..... 89
5.7 Corpus et caractérisations ..... 90
5.8 Conclusions ..... 92
6 ÉVALUATIONS & RÉSULTATS ..... 93
6.1 Introduction ..... 93
6.2 Évaluation ..... 93
6.2.1 Pourquoi évalue-t-on ? ..... 93
6.2.2 Qu'évalue-t-on ? ..... 94
6.2.3 Comment évalue-t-on ? ..... 95
6.2.4 Évaluation objective ..... 95
6.2.5 Évaluation subjective ..... 96
6.3 Résultats ..... 97
6.4 Problèmes rencontrés ..... 99
6.5 Solution appliquée ..... 103
6.6 Conclusion ..... 103
7 CONCLUSIONS & PERSPECTIVES ..... 104
7.1 Conclusions ..... 104
7.2 Perspectives ..... 105
SITOGRAPHIE ..... 106
RÉFÉRENCES BIBLIOGRAPHIQUES ..... 110
8 TABLE DES ANNEXES ..... 114
*ANNEXE A :* Jeu de description proposé & Catégorisation des descripteurs ..... 114
*ANNEXE B :* Le Modèle de marquage de Bailly (ceci n’est qu’une interprétation possible) 115
*ANNEXE C :* Transcriptions des dictées ..... 117
*ANNEXE D :* Les différentes versions des dictées ..... 122
*ANNEXE E :* Quelques expressions régulières utilisées ..... 144

*ANNEXE F :* Systéme de marquage de Bailly appliqué aux dictées ..... 146
*ANNEXE G :* Questionnaire ..... 170
*ANNEXE H :* Tableaux présentant les résultats d'évaluation des différentes versions « originale », « contrôle » et « multi-style » ..... 175
*ANNEXE I :* Le schwa et son élision ..... 177
*ANNEXE J:* Erreurs commises en fonction de chaque système ..... 178

# TABLE DES ILLUSTRATIONS

**Figure 1.** Le site Ampère du laboratoire GIPSA-Lab ..... 21

**Figure 2**. Schéma représentant les difficultés rencontrées par les différents acteurs. ..... 30

**Figure 3**. Synthétiseur de la parole à partir du texte. ..... 33

**Figure 4.** Schéma synoptique du logiciel PlatON. ..... 34

**Figure 5.** Superposition des trajectoires prosodiques (haut : F0 ; bas : débit) d'une phrase dictée d'une seule traite à l'initiation de l'exercice pour familiariser l'apprenant au contexte sémantique ( a), par groupes (b), par mot (c) puis d'une seule traite pour vérification. ..... 39

**Figure 6**. Architecture d'un système de synthèse de parole. ..... 42

**Figure 7.** Modèle de Fujisaki (d'excitation de source vocale). ..... 49

**Figure 8.** Stylisation (école hollandaise) : ligne de déclination et stylisation en contours standards. ..... 50

**Figure 9.** Synthétiseur par concaténation : les opérations qui dépendent de la langue sont indiquées par un drapeau. *(Dutoit, T.)* ..... 53

**Figure 10.**Illustration d'une sélection statique (diphone), le symbole # est une indication de pause. ..... 54

**Figure 11.** Illustration de l'algorithme de sélection ; L'algorithme de Viterbi est en général utilisé pour optimiser cette recherche du meilleur chemin. ..... 55

**Figure 12.** « Principe de la synthèse par corpus. Dans un premier, en se basant sur une consigne définie et un corpus de parole annoté, un graphe d'unités candidates est constitué (A). En utilisant un algorithme de Viterbi, les unités optimales (au sens des coûts définis dans le système de synthèse) sont sélectionnées puis concaténées (B). » *(Le Maguer, S., 2013)*.... 55

**Figure 13**. Le système HTS : le corpus est composé du signal paramétrisé, en appliquant les outils STRAIGHT *(Kawahara 1999 ; Kawahara 2001)* et SPTK *(Fukada 1992)*, avec les descripteurs qui permettent d'identifier un segment. En s'appuyant sur ce corpus, les HMM sont appris. La séquence de paramètres qui correspond à la séquence de descripteurs définis par les outils d'analyse linguistique, est générée lors de la phase de synthèse. Les outils STRAIGHT et SPTK permettent d'avoir le signal acoustique à partir des paramètres générés par HTS. *(Le Maguer, S., 2013)* ..... 57

**Figure 14**. Principe de fonctionnement de MBROLA. ..... 59

**Figure 15**. Modification de la durée et du pitch avec TD-PSOLA. Partant de cinq fenêtres d'OLA sur l'axe temporel à l'analyse (t), on obtient quatre marqueurs de pitch sur l'axe du

temps de synthèse (t'). Dans cet exemple, la deuxième fenêtre d'analyse a été éliminée ; la troisième a été dupliquée. ..... 60

**Figure 16.** Analyse ascendante. ..... 63

**Figure 17**. Analyse descendante ..... 64

**Figure 18**. Contour multiparamétrique. L'ensemble des deux courbes représente le contour multiparamétrique de la phrase. Les traits sur l'abscisse indiquent les GIPCs. Les mots de la transcription orthographique de la phrase sont alignés avec la fin de leur dernier GIPC *(Bailly, G. & Holm, B., 2003).* ..... 67

**Figure 19**. Hypothèses du SFC ..... 68

**Figure 20.** « Exemple de superposition. Les résultats des prédictions du SFC seront, par la suite, souvent représentés sous cette forme. On y trouve tout en haut la transcription orthographique (les mots étant alignés avec la fin de leur dernier GIPC). La courbe en tiret représente le contour global observé ; celle en trait plein le contour prédit par superposition des contours fonctionnels qui, à leur tour, se trouve dans la partie inférieure des figures. Les étiquettes des fonctions associées aux contours fonctionnels sont notées tout en bas à l'endroit du point d'ancrage central du contour. » *(Bailly, G. & Holm, B., 2003)* ..... 70

**Figure 21**. Cycle d'analyse par la synthèse. ..... 73

**Figure 22.** Marquage d'une version (a) d'un échantillon de parole « On entendait parfois des soupirs de satisfaction et les bras se croisaient tout seuls sur les tables. » à l'aide de Praat. Le [:FF] balise délimitant le début et la fin des unités. D'autres tags sont utilisés pour marquer les fonctions métalinguistiques : [:DC] pour la phrase déclarative, [:IT] entre clauses, [:DG] et [:DD] entre les groupes nominaux et verbaux..... 81

**Figure 23**. Attention portée aux contours intonatifs, aux signes de ponctuation et aux pauses. Un échantillon de parole « Le maître ne lisait pas : il parlait, il racontait.». ..... 81

**Figure 24.** Utilisation du marqueur ID entre deux propositions connectées par une coordination. Un échantillon de parole « On entendait parfois des soupirs de satisfaction et les bras se croisaient tout seuls sur les tables.» ..... 82

**Figure 25.** Le marqueur ID dans la version (a) devient marqueur CT dans la version (c). Un échantillon de parole « Des soupirs de satisfaction.»..... 83

**Figure 26**. Forme du contour intonatif à la fin des différents groupes prosodiques dans chacune des versions (b) et (c) : le cas d'un marqueur de continuation..... 83

**Figure 27.** Forme du contour intonatif à la fin des différents groupes prosodiques dans chacune des versions (b) et (c) : le cas d'un marqueur d'interdépendance. ..... 84

**Figure 28.** Forme du contour intonatif à la fin des différents groupes prosodiques dans chacune des versions (b) et (c) : le cas d'une phrase déclarative. ..... 84
**Figure 29.** Découpage d'un bout de phrase sans signe de ponctuation avec le marqueur :CT. ..... 85
**Figure 30.** Découpage d'un bout de phrase sans signe de ponctuation avec le marqueur :CT. ..... 85
**Figure 31.** Sortie de Cordial avec l'extension .COR ..... 87
**Figure 32**. Sortie de Cordial : la négation. ..... 88
**Figure 33.** Marquage prosodique de l'épellation. ..... 89
**Figure 34** . Superposition des ellipses de dispersion des caractéristiques de réalisations des groupes de souffle pour les 4 versions ((a) : compréhension ; (b) : groupes syntaxiques ; (c): mots de contenus ; (d) : relecture finale). En abscisse, la durée de phonation du groupe et en ordonnée, la durée de pause subséquente. ..... 92
**Figure 35.** Tableau présentant la description des sujets. ..... 97
**Figure 36.** Histogramme présentant les résultats d'ordre de préférence des trois styles. ..... 98
**Figure 37.** Histogramme présentant les résultats d'évaluation du niveau, intelligibilité, agrément, manière de lecture, vitesse générale et rendu des ponctuations des trois styles. ..... 98
**Figure 38.** Histogramme présentant les résultats d'évaluation entre les versions (a), (b), (c) et (d) des trois styles. ..... 99
**Figure 39.** Schéma présentant les deux cas de la rime ..... 100
**Figure 40.** Rime voisée : la montée finale est réalisée sur le [r] de « bord » alors que le contour est plat sur la voyelle. ..... 100
**Figure 41**. Rime voisée: la montée finale est réalisée sur le [l] de « virgule ». ..... 101
**Figure 42**. Rime voisée. : La montée finale est réalisée sur le schwa (étiqueté comme une voyelle pleine ici) de « virgule » alors que le contour est plat sur la voyelle. ..... 101
**Figure 43**. Rime non voisée: la montée finale est réalisée sur la voyelle [□] de « bord ». 102
**Figure 44**. Rime non voisée: la montée finale est réalisée sur la voyelle [i] de « magique ». ..... 102
**Figure 45.** Rime non voisée: la montée finale est réalisée sur la voyelle [i] de « ensuite ». ..... 103
**Figure 46.** Bruits respiratoires ..... 105
**Figure 47.** Exemple d'arbre de dépendance et de marquage prosodique. D'après Bailly (1989, p. 140) ..... 116
**Figure 48.** Les différentes questions d'évaluation. ..... 175

**Figure 49.** Tableau présentant les résultats d'évaluation de la version « originale ». ..... 175
**Figure 50.** Tableau présentant les résultat d'évaluation de la version « contrôle ». ..... 176
**Figure 51.** Tableau présentant les résultat d'évaluation de la version « multi-style » ..... 176

*« Qui veut faire quelque chose trouve un moyen, qui ne veut rien faire trouve une excuse. »*
*[Proverbe Arabe]*

# PRÉAMBULE

## L'équipe MAGIC du laboratoire GIPSA

Le laboratoire d'accueil pour ce stage est le GIPSA-Lab, **Grenoble Images Parole Signal Automatique**. Le GIPSA-lab est une unité de recherche mixte du CNRS, de Grenoble-INP, de l'université Joseph Fourier et de l'université Stendhal, elle est conventionnée avec l'INRIA, l'Observatoire de Grenoble et l'université Pierre Mendès France. Fort de 300 personnes dont plus d'une centaine de doctorants, GIPSA-lab est un laboratoire pluridisciplinaire développant des recherches fondamentales et finalisées sur les signaux et systèmes complexes. Il est reconnu internationalement pour ses recherches en Automatique, Signal et Images, Parole et Cognition et développe des projets dans les domaines stratégiques de l'énergie, de l'environnement, de la communication, des systèmes intelligents, de la santé et de l'ingénierie linguistique. »

**Figure 1. Le site Ampère du laboratoire GIPSA-Lab**

Le laboratoire GIPSA est organisé en trois appartements AUTOMATIQUE, IMAGES-SIGNAL et PAROLE-COGNITION.

Le département PAROLE-COGNITION contient quatre équipes :

**GAMA :** Gestes phonatoires, Analyse et Modélisation Acoustique, travaille sur la caractérisation physiologique, acoustique et aéroacoustique et perceptive de la phonation humaine.

**PCMD :** Parole Cerveau Multimodalité Développement, l'objectif est l'étude de la parole dans le creuset des interactions perceptivo-motrices.

**SLD :** Systèmes linguistiques et dialectologie, Les recherches de l'équipe sont consacrées à l'observation, la compréhension, l'explication et la modélisation de la diversité et de la variation des systèmes linguistiques, dans leurs dimensions sonores et physiologiques.

**MAGIC :** Machines parlantes, Gestes oro-faciaux, Interaction face-à-face, Communication augmentée, l'ambition de cette équipe est d'étudier et de modéliser les signaux de la communication humaine face-à-face, de construire et de contrôler des têtes parlantes virtuelles qui soient capables de reproduire au mieux les structures des signaux de la communication humaine.

L'équipe d'accueil pour ce stage est MAGIC qui travaille, notamment, sur les systèmes de traitement interactifs de la parole.

## Le projet ORTHOLEARN

Mon travail sera financé par le projet ORTHOLEARN.

Ce projet, financé par l'ANR sur trois ans (2012-2015) a l'ambition d'apporter des éléments de réponse à des questions cruciales sur l'acquisition de l'orthographe, notamment sur le rôle de l'empan visuo-attentionnel (EVA) sur les stratégies de décodage de l'écrit. ORTHOLEARN vise une meilleure appréhension des dimensions cognitives et neurobiologiques de l'acquisition de l'orthographe lexicale.

*« En un mot, j'ai mis mon orgueil à prouver, ici, mon ignorance, aux admirables savants qui honorent notre espèce. »*
*[villiers, 1993]*

# INTRODUCTION

Durant les dernières décennies, les techniques de traitement de la parole et de traitement automatique de la langue ont évolué d'une manière importante. Il est aujourd'hui techniquement possible de réaliser des interfaces homme-machine fiables utilisant la voix comme moyen d'interaction principale. Néanmoins, la percée des interfaces vocales est moins impressionnante que celle qui nous était prédite, notamment dans le domaine de la synthèse vocale, en particulier dans les systèmes d'apprentissage des langues assisté par ordinateur (ALAO).

Alors que l'orthographe des élèves est de plus en plus catastrophique, des équipes de recherche universitaires puis des sociétés commerciales ont mis au point des systèmes de dictées automatiques, permettant de faire une dictée en ligne, les dictées étant gérées (choix du contenu, dictée à voix haute, évaluation des performances et correction) par un professeur virtuel. Dans de tels environnements, le système lit le texte au rythme de frappe de l'élève et prend en charge l'explication des fautes commises.

Dans notre revue de la littérature, nous avons identifié quelques systèmes de dictée : la dictée interactive, DICTOR, Ordictée et PlatON. Si ces systèmes proposent tous une correction et une explication des erreurs à l'apprenant, ils utilisent tous un système de synthèse de textes standard sans une adaptation spécifique des stratégies d'énonciation lors des diverses répétitions des mots lors de la dictée.

L'objectif de ce travail est d'étudier les stratégies d'énonciation et ceci de manière empirique, notamment les multiples passages opérés sur le texte, à partir d'un corpus de dictées à voix haute réalisées par des lecteurs experts. Un corpus de 15 dictées orales effectuées par une unique lectrice, qui s'appelle Aude, a été pour ceci collecté et analysé.

Nos dictées sont puisées sur un site www.ladictee.fr mettant à disposition gratuitement des dictées adaptées à divers niveaux d'études. Nous nous concentrerons sur le niveau CM1-CM2.

La plupart des travaux sur la dictée sont francophones et se concentrent sur l'analyse des productions des scripteurs. Seul un travail très récent s'est intéressé à l'étude des diverses stratégies de découpage des lecteurs. Dans nos dictées, la lectrice effectue ainsi 4 passes sur le texte : une lecture (a) dite de compréhension, où le texte est lu une première fois de manière intégrale ; deux répétitions de groupes de mots, une version (b) lente suivie d'une version (c) où pratiquement tous les mots non grammaticaux sont prononcés de manière isolée ; puis une version (d) dite de vérification, permettant au scripteur de vérifier l'orthographe de chaque phrase.

Aucun travail de fond n'a été effectué – à notre connaissance – sur l'étude des modifications des stratégies d'énonciation, notamment des changements de contours prosodiques, de ces divers découpages et relectures du texte. Nous avons pour ceci analysé de manière fine le corpus de 15 dictées orales. Un ensemble restreint de marques prosodiques a été utilisé : notre hypothèse forte est que l'énonciation à voix haute consiste à faire écouter des fragments syntaxiques prélevés sur un arbre syntaxique complet. Nous faisons l'hypothèse que des diverses versions d'énonciation des groupes de mot ne modifient que la forme des contours mélodiques (c'est-à-dire une courbe mélodique la plus naturelle et cohérente possible) et rythmiques (de manière plus concrète, la durée de chaque phonème et des pauses) et non pas leur fonction.

# PLAN DU MÉMOIRE

Ce mémoire présente un système de dictée automatique de textes à haute voix en se basant sur une analyse de corpus et une modélisation des stratégies d'énonciation. Nous décrirons, ainsi, dans le ***premier chapitre*** l'état de l'art. En effet, il existe différents systèmes de dictée automatique qui se penchent sur la détection et la correction des erreurs. Nonobstant, les stratégies d'énonciation restent orphelines à ce jour, hormis les rendez-vous donnés par les signes de ponctuation, qui, énoncés par le tuteur, aident à la structuration de l'énoncé. La problématique, les attentes et les hypothèses seront présentés dans le ***deuxième chapitre***. Dans le ***troisième chapitre***, nous introduirons des modules permettant de modéliser les composantes acoustiques du signal de parole à partir d'un texte. Nous présenterons, dans la première partie du ***quatrième chapitre***, les deux approches concurrentes, l'analyse ascendante versus l'analyse descendante, opposées dans leurs aboutissements, pour mettre en correspondance fonctions linguistiques et substance prosodique. Si de nos jours, les méthodes actuelles utilisées en synthèse de la parole opèrent par analyse descendante, en s'affranchissant de tout modèle prosodique intermédiaire, elles requièrent de grosses masses de données d'apprentissage. Nous introduirons, dans la deuxième partie du ce chapitre, le modèle prosodique SFC « **S**uperposition of **F**unctional **C**ontours », que nous avons utilisé. Il propose un encodage direct de fonctions linguistiques ou paralinguistiques par des contours multiparamétriques chevauchants. Le ***cinquième chapitre*** sera consacré à la description des corpus et les méthodologies de recherche. En conséquence, nous traiterons les différents problèmes rencontrés et les solutions envisagées. Le ***sixième chapitre*** introduira l'évaluation de notre système, ainsi que les résultats obtenus. Enfin, nous présenterons dans le ***septième chapitre*** nos conclusions et nos perspectives.

# PARTIE 1 : ÉTUDE THÉORIQUE

*« Écrire c'est tenter de savoir ce que l'on écrirait si on écrivait – on ne le sait qu'après – avant, c'est la question la plus dangereuse que l'on puisse se poser. Mais c'est la plus courante aussi »*
*[Marguerite DURAS, Écrire]*

# 1 ÉTAT DE L'ART

## 1.1 La dictée

### 1.1.1 Caractéristiques propres à la langue

La langue française est une langue complexe. En effet elle a une orthographe opaque car la correspondance entre phonèmes et graphèmes [1] est très variable. La complexité de l'orthographe du français apparaît notamment dans les problèmes d'accord, d'homophones, de lettres qui ne sont pas prononcées. Selon Nina Catach *(1989)*, la complexité de la langue française est due à sa composition de trois systèmes :

- **Le système phonogrammique :** Un phonogramme est un graphème qui transcrit un phonème. Quand un son s'écrit comme il se prononce, c'est qu'il est transcrit par un phonogramme (exemple : ro**s**e, f**e**mme, etc.).
- **Le système morphogrammique :** Un morphogramme est un graphème qui va ajouter une information autre que celles qui sont relatives aux sons. C'est donc un graphème qui se trouve généralement à la fin d'un mot et qui ne va pas être transcrit à l'oral, mais qui renseigne le lecteur sur un plan grammatical ou sur un plan lexical (exemple : du pluriel en S ou en ENT qui n'est audible que lorsqu'une liaison est possible).
- **Le système logogrammique :** Un logogramme est un graphème qui permet de différencier des homonymes sans faire référence à la famille de mots. Ce peut être une lettre ou un accent. Tout comme il existe deux types de morphogrammes, il existe deux types de logogrammes : les logogrammes grammaticaux (à/a ; où/ou ; son/sont…) et les logogrammes lexicaux (tente/tante ; repère/repaire).»

[1] Est l'unité de l'écrit correspondant à l'unité orale qui est le phonème.

Cela explique bien le fait que beaucoup d'élèves français ont donc une représentation négative de la dictée. Pendant la conversion phonétique-orthographique, il est important de poser des questions sur la nature et la fonction de chaque mot, ainsi que sur les liens existants entre ces différents mots. S'agit-il d'un mot qui a une catégorie syntaxique fermée comme un déterminant (le, cette, quelques), une préposition (sur, à l'aide de, de), une conjonction (et, mais, parce que), un pronom (celui-ci, elle, chacun), ou bien une catégorie syntaxique ouverte comme un adverbe (ici, très, demain, agréablement), un adjectif (bleu, grand, universitaire), un nom (porte, gentillesse, linguistique), un verbe (être, commencer, chanter) ? Dans le cas d'un verbe, il faut se demander où se trouve son sujet, si ce dernier est au singulier ou au pluriel, quelles sont les terminaisons qui correspondent à ce temps, à quel temps est ce verbe, quel est son mode, etc.

Au cours d'une dictée, un élève doit transformer un matériau phonique en un matériau graphique, cela signifie faire une réalisation d'une correspondance correcte entre sons et lettres. La dictée est donc une évaluation sommative de ce que les élèves ont acquis pendant le cours. Elle contribue à améliorer les connaissances orthographiques sans l'utilisation d'un dictionnaire. Elle vise à faire évoluer les concepts orthographiques en y incluant des activités de concentration et de réflexion.

Le rôle primordial de la dictée est de faciliter la compréhension et la mémorisation de l'énoncé qui se manifeste par le biais d'une production écrite, ce qui représente un savoir-faire difficile à maîtriser. Il existe différents types de dictées classées selon leur but :

<u>L'évaluation :</u>

- ✓ La dictée de contrôle : les apprenants ne peuvent poser aucune question et ils n'ont à leur disposition aucun document de référence.

<u>L'apprentissage :</u>

- ✓ La dictée préparée : cette dictée est précédée par une phase de préparation pendant laquelle l'apprenant sait d'avance la dictée ou bien certains mots.
- ✓ La dictée dirigée : le tuteur attire l'attention de l'apprenant sur les lacunes et les mots de blocage : des accords, de la conjugaison, des lettres muettes, etc.
- ✓ La dictée consultation (ou dictée assistée ou dictée sans faute) : l'apprenant a accès à des documents et à des ouvrages.
- ✓ L'autodictée : l'apprenant se dicte à lui-même ce qu'il doit écrire.

Les dictées peuvent être aussi classées selon la taille des unités linguistique à retranscrire :

- ✓ La dictée de mots : est appelée aussi dictée trouée. Le tuteur distribue aux apprenants un texte à trous dans lequel ces derniers doivent saisir les mots dictés.
- ✓ La dictée de phrases : l'apprenant doit transcrire des phrases indépendantes les uns des autres.
- ✓ La dictée de textes : il s'agit de la forme de dictée la plus répandue. C'est la dictée traditionnelle à laquelle on pense immédiatement en parlant de dictée.

Dans ce mémoire, nous n'allons travailler que sur la dictée traditionnelle.

### 1.1.2 Stratégies existantes

L'exercice traditionnel de dictée se déroule souvent en plusieurs passes effectuées sur le texte :

- Le tuteur lit l'énoncé complet pour donner le contexte sémantique du texte.
- La dictée procède ensuite par phrases et met en valeur le découpage morphosyntaxique[2] de l'énoncé (groupes syntagmatiques, ponctuations, etc.).
- D'autres passages peuvent être effectués sur un découpage de l'énoncé encore plus fin, notamment sur les mots.
- Finalement, cet exercice se conclut souvent par une lecture du texte complet permettant à l'apprenant de vérifier sa transcription.

Il est bien sûr recommandé d'écrire au tableau les mots pouvant créer des blocages.

Beaufort *(Beaufort, R. et al., 2011)* a particulièrement parlé de la rareté des essais d'automatisation de la dictée. Cependant, avec l'accroissement de la synthèse de la parole et grâce aux synthétiseurs de la parole, lire automatiquement un texte quelconque n'est plus un problème.

### 1.1.3 Les difficultés

Coniam *(Coniam, D., 1996 ; Coniam, D., 1998)* a notamment montré que la dictée est un très bon moyen d'estimer le niveau d'un étudiant. La dictée, souvent associée à la tradition scolaire comme exercice rébarbatif et inutile, est typiquement l'un des exercices semi-ouverts permettant d'évaluer les compétences linguistiques d'un élève. Cependant, elle est devenue une situation psychologique qui est souvent synonyme d'anxiété et d'angoisse parce que

[2] « L'étiquetage morpho-syntaxique consiste à associer une étiquette morpho-syntaxique à chaque mot, il repose sur la segmentation en mots et en phrases effectuée préalablement. »

l'élève a peur de faire des fautes. En effet, échouer dans sa dictée, fort courant, est bien déprimant. Cela dû à plusieurs raisons :

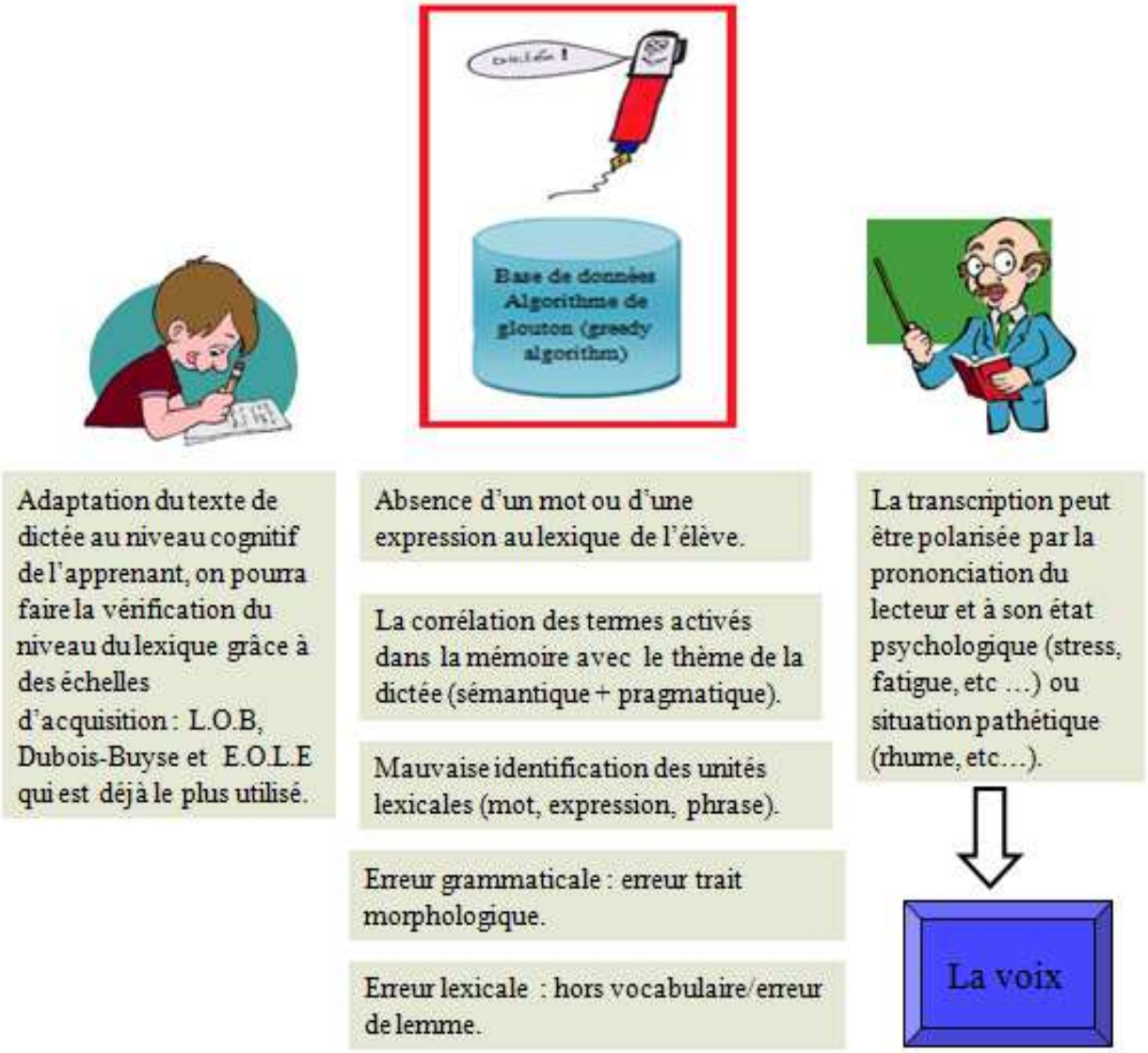



**Figure 2. Schéma représentant les difficultés rencontrées par les différents acteurs.**

Nos recherches nous ont amené à trouver plusieurs systèmes qui traitent la dictée automatique dans un usage didactique. Ces systèmes intègrent pour la plupart un module de détection d'erreur automatique dont le but est de diminuer le taux d'erreur de l'élève en lui donnant des commentaires et en lui rappelant des règles de grammaire appropriées.

**Remarque**

Les anglo-saxons ne font pas de la dictée car ils sont partants pour le fait qu'il faut produire de l'écrit, autrement dit, l'écrit pour eux ne doit pas être imposé.

## 1.2 Logiciels pour la dictée

### 1.2.1 Ordictée

Le logiciel de dictée automatique a été mis au point par Guyomard *(Guyomard, M. et al., 1997)* à L'Institut de recherche en informatique et systèmes aléatoires (IRISA), Il permet à un élève de réaliser sa dictée de manière autonome dans un environnement non stressant en utilisant la synthèse de parole.

Le système joue le rôle de l'instituteur, il dicte le texte, qui est lu par un synthétiseur de parole, à l'élève qui le tape sur un ordinateur tout en s'adaptant au rythme de la lecture et à la vitesse de frappe de ce dernier. Puis il corrige le texte tapé en présentant à l'élève les fautes commises.

Le logiciel est constitué de trois modules :

- le module « élève » qui effectue la dictée proprement dite
- le module « tuteur » qui permet à l'instituteur de créer ses propres dictées
- le module « concepteur » qui permet la gestion et le paramétrage de l'ensemble.

La fonction de suivi de frappe, c'est-à-dire l'adaptation du rythme de la lecture à la vitesse de frappe de l'élève, est fondée d'une part sur l'hypothèse *(Mercier, Guyomard et al., 2000)* que les fautes concernent en général la prononciation, et d'autre part sur la proximité phonétique des deux textes considérés (le texte tapé par l'élève et la référence du tuteur).

### 1.2.2 DICTOR

DICTOR a été développé par G. Conchi-Santiago à l'Institut de Recherche en Informatique de Toulouse (IRIT) *(Santiago-Oriola, C., 1998)*. Il s'agit d'un système d'aide à l'apprentissage de l'orthographe d'un système interactif de dictée. Ce système inclut un outil de vérification automatique qui est basé sur un algorithme stochastique d'alignement et des connaissances linguistiques.

La dictée est lue à l'étudiant par un système de synthèse de la parole et ce dernier corrige la saisie de l'étudiant. Le diagnostic et la présentation de ses résultats offrent des explications à l'étudiant à deux niveaux différents : le premier montre seulement le mot incorrect avec une description courte du genre d'erreur (orthographe, typographique) et le deuxième, demandé par l'étudiant si le malentendu reste, propose une explication plus complète sur l'orthographe, lexicologique ou règles grammaticales s'appliquant au terme dans ce contexte. D'après Santiago-Oriola, on considère que le système texte voix et le système d'enregistrement de

voix sont équivalents et que l'utilisation de l'une ou de l'autre sera seulement liée aux contraintes matérielles.

### 1.2.3 PlatON

Ce logiciel a été développé par Roekhaut *(Roekhaut, S. et al., 2013)* du CENTAL, le Centre de Traitement Automatique du langage de l'Université Catholique de Louvain en 2011. C'est une plate-forme internet révolutionnaire, dans ce contexte, explique Richard Beaufort "Nous avons modifié un système de synthèse de la parole pour qu'il puisse faire faire des dictées, en lisant les ponctuations et en respectant les arrêts ".

Selon un extrait de « La Meuse Luxembourg », sous le titre "PlatON va Vous Faire Des Dictées", du 8 septembre 2011 à propos de ce sujet : « Lorsque l'internaute a terminé de retranscrire le texte et de le relire, il l'envoie à l'ordinateur qui le corrige en l'espace d'une seconde. Son programme lui permet de comparer la copie à l'original. Il donne une brève explication des fautes commises, par exemple "il fallait utiliser le participe passé plutôt que l'infinitif. À terme, l'objectif est que la règle grammaticale complète soit fournie. Après plusieurs dictées, le logiciel analyse vos lacunes et propose des exercices adaptés. Ces derniers sont composés à partir d'extraits de textes d'auteurs. PlatON permet aussi à des enseignants de proposer leurs propres dictées qu'ils prononceront eux-mêmes ou confieront cette tâche au synthétiseur vocal. »

« La phase de la correction repose sur une analyse morphosyntaxique de l'original de la dictée. Cette analyse est produite par le système eLite *(Beaufort, R. & Ruelle, A., 2006)* qui réalise successivement un prétraitement, une analyse morphologique et une désambiguïsation contextuelle du texte. La désambiguïsation est obtenue par l'application d'un modèle de langue probabiliste » *(Beaufort R. et al., 2011)*.

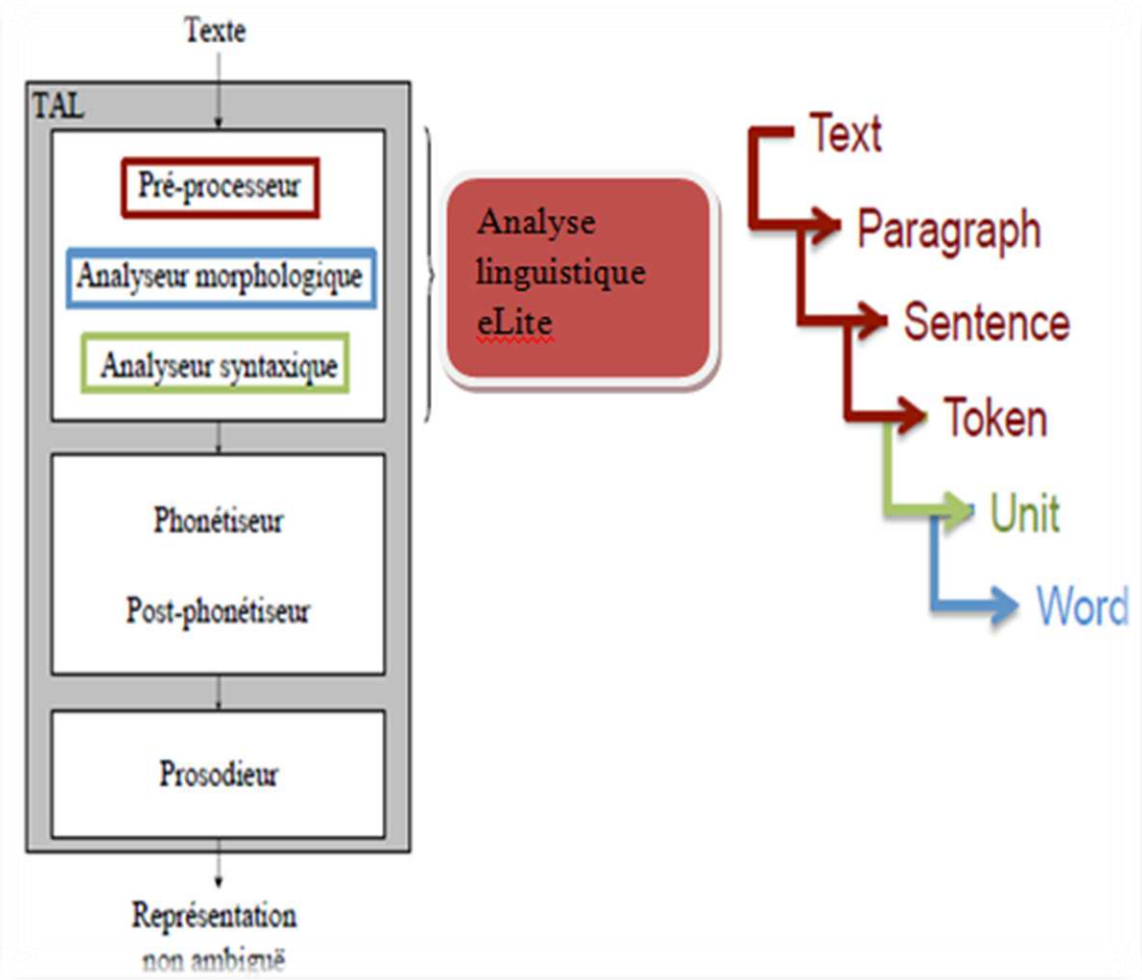



**Figure 3.** **Synthétiseur de la parole à partir du texte.**

Nous présenterons dans la section suivante le principe de fonctionnement du logiciel PlatOn. En effet, ce système est divisé en quatre phases : préparation, stratégies, évaluation et correction.

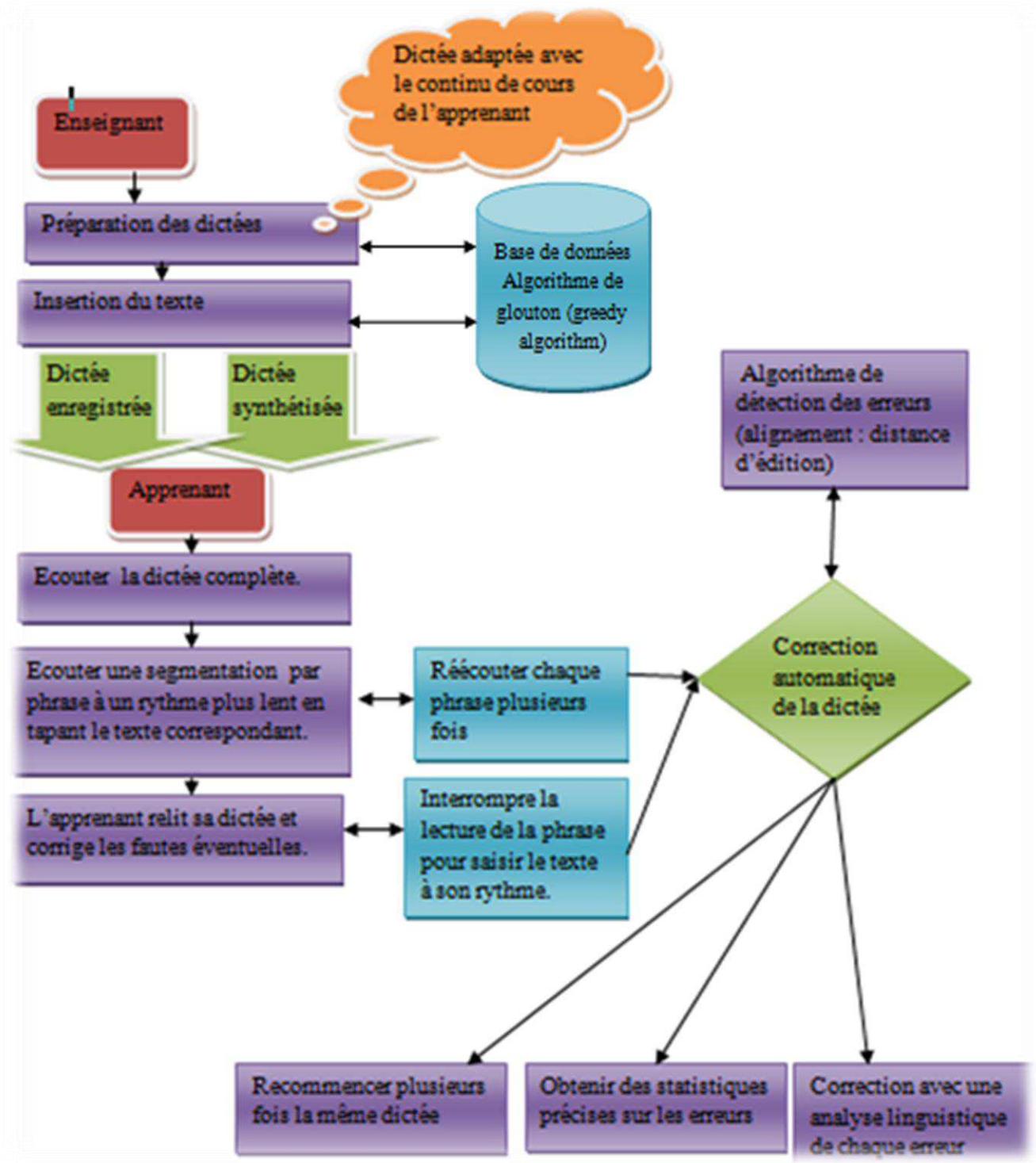



**Figure 4.** **Schéma synoptique du logiciel PlatON.**

### 1.2.4 La dictée interactive

Le logiciel développé à l'Université de Nice *(Ruggia, S., 2000)* propose deux modes d'utilisation :

- un parcours de simulation d'examen, où l'apprenant doit se soumettre à des contraintes de temps ;
- un parcours d'autoapprentissage, où l'apprenant gère lui-même l'écoute du texte (pauses et reprises).

A la fin de la dictée, l'apprenant a la possibilité de la corriger en réécoutant le texte autant de fois qu'il le souhaite. La correction automatique met en relief les fautes soulignant les mots en rouge. L'un des atouts du logiciel est qu'il permet la récupération pédagogique de la faute en donnant à voir et à entendre les mots porteurs de fautes.

### 1.2.5 Commentaires

En effet, les quatre logiciels présentés dans la section précédente se focalisent principalement sur la détection et la correction des erreurs.

La dictée automatique permet d'effectuer des exercices de manière autonome, en ayant la possibilité de jouer sur l'aide apportée par le tuteur lors de sa dictée.

Il faut pour ceci disposé d'un système de synthèse qui puisse adapter sa vitesse et son style d'énonciation à la partie de texte, à son contenu et à l'apprenant.

La conception d'un tel système soulève certains verrous scientifiques et technologiques: d'une part la conception de stratégies d'énonciation lors de la lecture (phrase complète, répétition d'un groupe de mots ou d'un mot) et la prise en compte par le système des difficultés supposées de l'apprenant.

### 1.2.6 Conclusions

La synthèse de la parole est une technologie en plein essor dont l'utilisation pour les logiciels assistés par ordinateur apparaît comme un domaine en plein extension et porteur d'espoirs.

Les systèmes de la dictée automatique actuellement disponibles se caractérisent par le fait que, d'une part, ils se concentrent sur les corrections orthographiques de façon précise et, d'autre part, ils permettent à l'élève de faire sa dictée en s'adaptant à son rythme.

La caractéristique de la dictée est que le tuteur énonce les phrases par unités plus petites (groupes, mots) en insistant éventuellement sur des difficultés possibles de correspondance phonème-graphème. Si l'on veut adapter la synthèse aux unités énoncées, il faut proposer un modèle de synthèse – notamment de génération de la prosodie – qui puisse prendre en compte cette synthèse à empan variable.

En guise de conclusion, on va analyser les caractéristiques des paramètres acoustiques d'un système de synthèse de la parole afin de mettre en place des stratégies d'énonciation du niveau textuel (découpages, répétitions, etc.) et du niveau sonore (phonétisation, prosodie, etc.). Des techniques d'apprentissage automatique (arbres de régression, réseaux de neurones[3], chaînes de Markov) seront mises en œuvre pour mettre en correspondance l'arbre syntaxique avec les trajectoires de paramètres prosodiques en fonction des niveaux de lecture du texte. On évaluera notamment la capacité du modèle SFC *(Bailly, G. & Holm, B., 2005)* à capturer les régularités de cette correspondance.

[3] Une technique de mise en correspondance des caractéristiques continues ou discrètes qui opère par combinaison de fonctions de régressions élémentaires.

*« Choisissez un travail que vous aimez et vous n'aurez pas à travailler un seul jour de votre vie. »*
*[Confucius]*

# 2 PROBLÉMATIQUE

## 2.1 Introduction

L'un des objectifs du projet ORTHOLEARN est de proposer des outils informatiques permettant de mettre en relief des unités orthographiques de tailles variables. Un premier logiciel dit de lecture synchrone assure une présentation synchrone des unités phonologiques correspondantes. La synchronie des informations orthographiques et phonologiques apparaît cruciale pour la mémorisation de la forme orthographique des mots lus. L'utilisation de « fenêtres » variables permet de guider et entraîner l'empan visuo-attentionnel (EVA).

La dictée permet de même de mettre en relief la correspondance entre l'oral et l'écrit en effectuant plusieurs passages sur le texte.

## 2.2 Répondre à un besoin

Lorsque l'on s'attaque à un tel projet, il est évident de s'interroger sur la nécessité d'un tel projet et sur la légitimité d'un travail de recherche pour cibler les objectifs à atteindre.

Dans notre cas, nous cherchons à développer un système de dictée automatique de textes à voix haute, prenant en charge la conduite de la dictée d'un apprenant à partir d'un texte source, et ceci avec le minimum de marquage spécifique. Ce projet ne traite ni des composants didactiques en amont (choix du texte et identification des difficultés attendues en fonction des compétences linguistiques de l'apprenant) ni en aval (identification des erreurs, notation, corrections, explications et enchaînement d'exercices complémentaires spécifiques).

La crainte d'échouer dans sa dictée, d'être pénalisé en ayant une note catastrophique et de perdre le rythme influent négativement sur la psychologie de l'élève. C'est cette crainte qui motive les élèves à chercher l'alternative appropriée vu qu'avec la dictée automatique l'élève lui-même va assumer la gestion de cette activité.

## 2.3 Analyser et modéliser les stratégies d'énonciation mises en œuvre en dictée orale

Il n'était pourtant pas évident que l'on puisse un jour faire une dictée automatique avec une voix synthétisée et se débarrasser de dictées préenregistrées avec une voix humaine.

L'évolution de la synthèse de la parole qui nous a amené à avoir une voix intelligible de haute qualité *(Santiago-Oriola, C., 1999)* a notamment montré l'efficacité de la voix synthétisée et sa fiabilité dans le contexte de la dictée. En effet, en comparant une dictée qui était lue par une voix humaine et une voix synthétisée, les résultats étaient presque semblables. D'après l'auteur, les élèves ont même, malgré la qualité inégale de la synthèse de l'époque, une tendance à préférer la voix synthétisée à la voix humaine.

En dépit de cette évolution importante, des problèmes majeurs se sont posés, et se posent encore, aux chercheurs vu que les scripteurs font toujours des erreurs catastrophiques pendant la dictée automatique. Pour pouvoir tenter de faire évoluer ces systèmes, il est donc nécessaire d'étudier de manière plus précise les caractéristiques des différents paramètres de la voix synthétisée.

## 2.4 Objectif de ce mémoire

Ce travail vise à fournir des solutions spécifiques pour traiter les difficultés orthographiques, lexicales et morphosyntaxiques de l'énoncé d'une dictée automatique dans une interaction Homme-Machine pour une catégorie particulière d'élèves qui sont en CM1-CM2 en générant la parole à un rythme adapté. Nous allons tenter de cerner des stratégies d'énonciation à partir de corpus de dictées orales qui peuvent améliorer la baisse du taux d'erreurs des élèves.

Ce travail se fera par la mise en place d'une série de tests qui a pour objectif de mettre en évidence le fonctionnement du système de synthèse vocale et les paramètres sur lesquels il agit pour obtenir des résultats satisfaisants.

L'état de l'art a démontré l'existence des stratégies pour la détection et la correction des erreurs, mais des stratégies satisfaisantes d'énonciation restent orphelines à ce jour, hormis les rendez-vous donnés par les signes de ponctuation, qui, énoncés par le tuteur, aident à la structuration de l'énoncé.

Notre premier objectif est d'analyser les caractéristiques des paramètres acoustiques d'un système de synthèse de la parole afin de mettre en place les stratégies convenables.

Notre second objectif est la formalisation de règles d'énonciation du niveau textuel (découpages, répétitions, etc.) et du niveau sonore (phonétisation, prosodie, etc.). Des

techniques d'apprentissage automatique (arbres de régression, réseaux de neurones, chaînes de Markov) seront mises en œuvre pour mettre en correspondance l'arbre syntaxique avec les trajectoires de paramètres prosodiques en fonction des niveaux de lecture du texte. On évaluera notamment la capacité du modèle SFC *(Bailly, G. & Holm, B., 2005)* à capturer les régularités de cette correspondance.

### 2.5 Attentes, hypothèses

Nous avons plusieurs attentes et hypothèses induites des travaux précédentes.

Le système proposé doit prendre en charge la conversion du texte en signal sonore, incluant les multiples répétitions de l'énonciation du contenu (découpages de l'énoncé à divers niveaux : texte, paragraphes, phrases, syntagmes, mots, etc.) en fonction des multiples passes nécessaires à l'apprenant pour comprendre, écrire et vérifier sa transcription littérale. On portera une attention particulière à la structuration prosodique du discours (cf. Figure 5).

On supposera que cette transcription se fera sur ordinateur : on proposera ainsi une stratégie permettant de prendre en compte les rendez-vous entre dictée et transcription.

Nos résultats seront mis en œuvre sur un synthétiseur du français existant, qui devrait in fine prendre en compte la coordination entre dictée et scripteur.

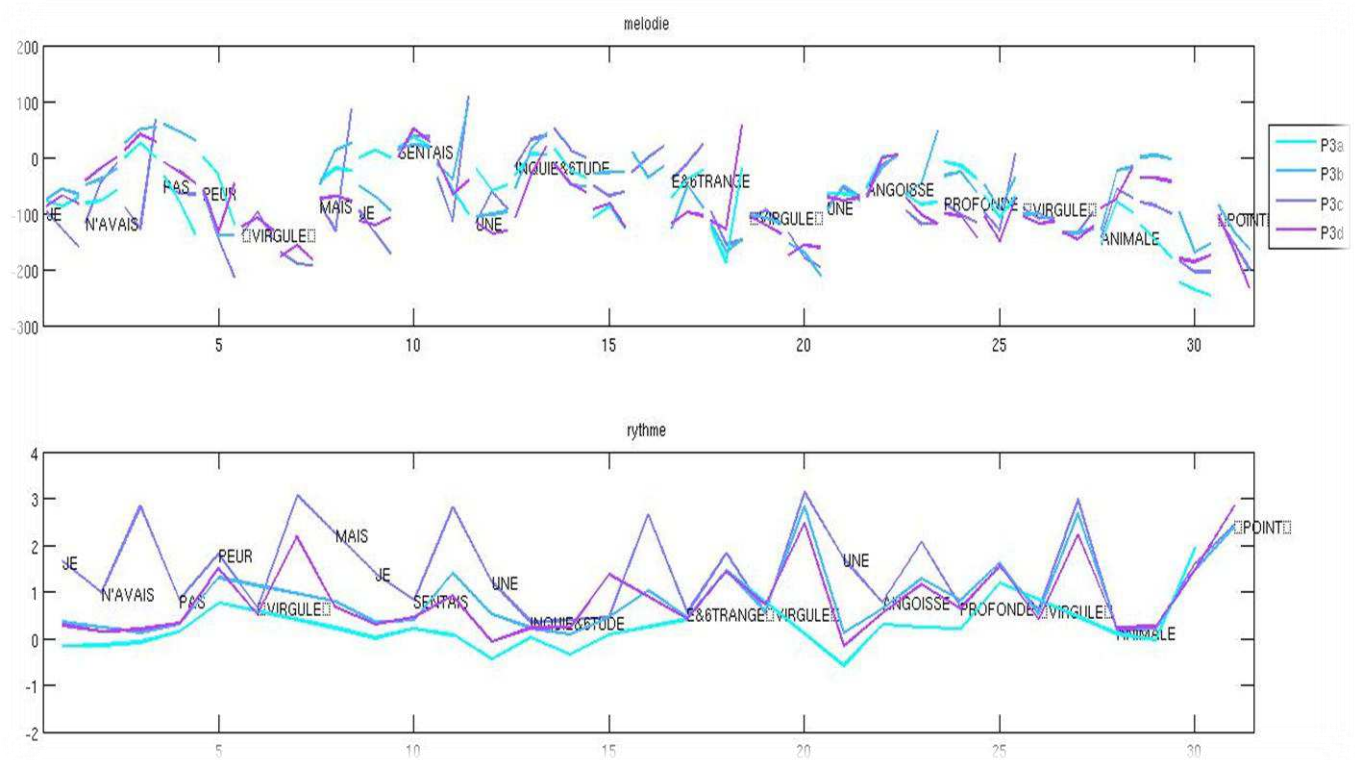



**Figure 5. Superposition des trajectoires prosodiques (haut : F0 ; bas : débit) d'une phrase dictée d'une seule traite à l'initiation de l'exercice pour familiariser l'apprenant au contexte sémantique ( a), par groupes (b), par mot (c) puis d'une seule traite pour vérification.**

En effet, selon ce schéma (figure 4), on voit que si la cohérence est globalement assurée, la coupure de l'énonciation en sous-unités induit des changements de stratégie clairement audibles. Ces figures illustrent le changement de prosodie dû aux diverses répétitions d'une même phrase par la tutrice. En haut le F0, en bas le rythme pour chaque syllabe. La version (a) est l'énonciation de la phrase complète : F0 modulé sans cassure, rythme assez plat... La version (b) est la segmentation en groupes prosodiques : on voit les chunks avec les maxima de ralentissement sur le rythme. La version (c) est une prononciation avec les mots isolés : on a des maxima de ralentissement à la fin de chaque mot. La version (d) est la version de vérification finale : plus lente. Dans (b), (c) et (d) les points et virgule sont énoncés. (a) et (d) sont bien prononcés d'une seule traite alors que (b) et (c) sont ici assemblées à partir de segments de paroles disjoints.

## 2.6 Conclusions

La dictée fait partie d'un ensemble d'exercices pédagogiques permettant l'acquisition de la correspondance entre la parole et sa transcription écrite. Lors de la dictée, le tuteur doit ajuster sa stratégie de lecture aux capacités linguistiques supposées de l'apprenant.

Une question fondamentale qu'il convient de se poser concerne la place du rythme dans ce processus : en effet, le tuteur doit proposer un débit de traitement de l'information compatible avec les processus cognitifs engagés par l'apprenant en fonction non seulement de son niveau de compétences mais aussi de son rythme actuel d'écriture qu'il soit cursif (maturité du contrôle moteur) ou par clavier (familiarité avec le clavier, type du clavier qwerty/azerty, fautes de frappe, etc.).

Notre travail d'analyse de corpus et de modélisation automatique n'a pas – à notre connaissance – d'équivalent dans la littérature, où les stratégies d'énonciation sont souvent réglées de manière ad hoc sans réelle assise sur l'expérimentation préalable. Nous espérons ainsi développer un système qui pourra s'enrichir automatiquement des performances réalisées par des tuteurs humains.

*« Ainsi, j'eusse blâmé, par exemple, le Phonographe de son impuissance à reproduire, en tant que bruits, le bruit… de la chute de l'empire romain… les bruits qui courent…les silences éloquents… et, en fait de voix, de ce qu'il ne peut clicher ni la voix de la conscience?… ni la voix -- du sang?… ni tous ces mots merveilleux qu'on prête aux grands hommes… ni le Chant du cygne… ni les sous-entendus… ni la voix lactée? Non! Ah! Je vais trop loin. »*
*[villiers, 1993]*

# 3 SYNTHÈSE AUDIO DU TEXTE AU SIGNAL ACOUSTIQUE

## 3.1 Introduction

La synthèse de la parole à partir d'un texte quelconque tapé au clavier consiste à générer un signal acoustique à partir d'une suite de caractères alphanumériques. C'est un problème mal posé au sens où, pour un locuteur donné, il y a une infinité de façons d'énoncer un texte. D'autre part, le débit d'information est drastiquement déséquilibré : quelques bits/s pour encoder un texte et plusieurs dizaines de Kbits/s pour encoder un signal de bonne qualité). La conversion texte-parole s'appuie donc sur un large ensemble de données et de connaissances sur la langue écrite et parlée. Un système de synthèse exploite ainsi un nombre important de ressources linguistiques, allant de dictionnaires morphologiques, de modèles syntaxiques à de larges corpus de parole paramétrés.

Par convention, nous noterons en lettres majuscules les caractères orthographiques donnés en exemple d'entrées textuelles.

## 3.2 Architecture d'un système de synthèse de parole

Un système de synthèse procède généralement en quatre grandes étapes :

1. Un traitement du texte permet de lever une structure grammaticale à partir d'une suite de caractères alphanumériques. C'est par essence le domaine du Traitement Automatique de la Langue Naturelle (TALN). On enchaîne souvent des modules de

prétraitement du texte (traitement des sigles, nombres, adresses mail, etc.), des systèmes d'analyse lexicale et morphosyntaxique combinés éventuellement avec des traitements de plus haut niveau – sémantiques ou pragmatiques – que l'on peut impliquer en fonction des domaines applicatifs

2. L'étape de transfert phonologique permet de passer de la structure de l'écrit à la structure de l'oral en complétant la structure précédente par des informations propres à l'oral, comme la phonétisation, l'organisation en syllabes et en groupes phonologiques divers (suivant la théorie intonative adoptée).
3. L'étape de génération effectue le passage de ces informations symboliques à une représentation paramétrique du signal. C'est l'étape charnière où l'on passe de représentations discrètes, généralement non spécifiques au locuteur, à une représentation continue de l'information.
4. L'étape de synthèse proprement dite, où la représentation paramétrique du signal est convertie en signal audible

On peut déjà noter qu'il est possible de s'affranchir de certaines de ces étapes. Certains systèmes s'affranchissent de l'étape de transfert phonologique en indexant directement des segments de parole par des caractéristiques textuelles ; de même l'étape de synthèse est inutile si la représentation paramétrique du signal adoptée est le signal lui-même. Mais la plupart des systèmes commerciaux adoptent ce modèle séquentiel où il est facile d'identifier et de mettre à jour les ressources.

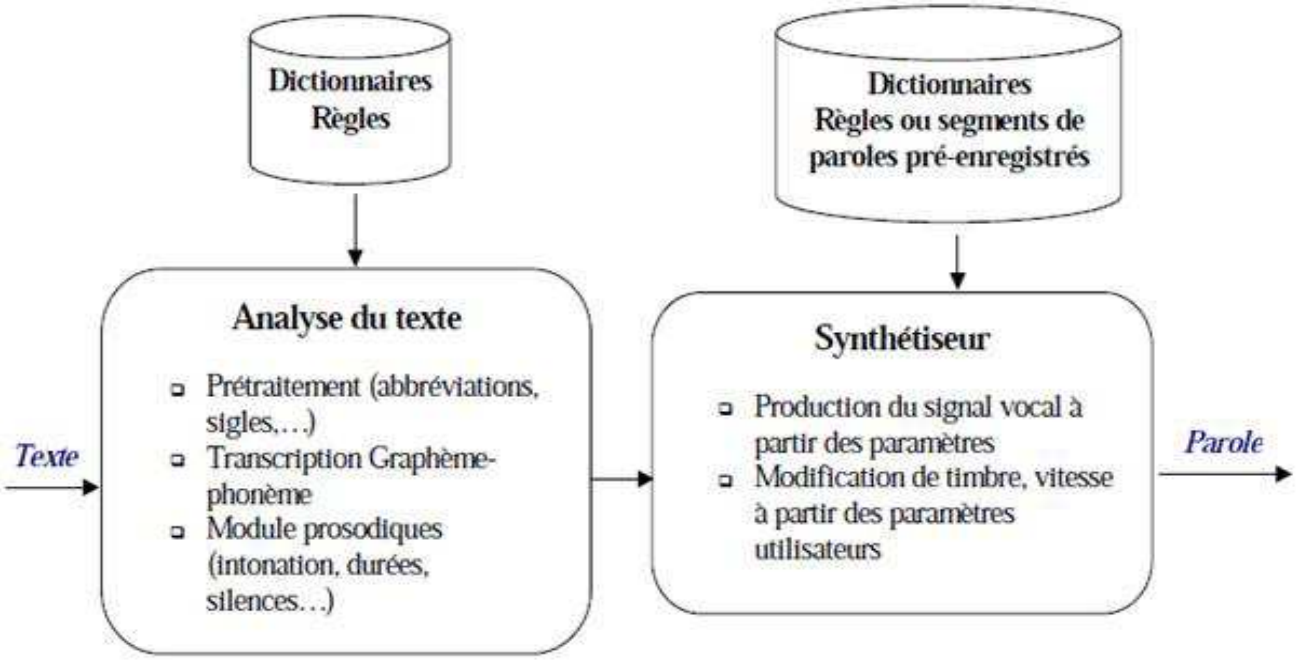



**Figure 6. Architecture d'un système de synthèse de parole.**

Nous allons à présent décrire chacun de ces modules, en particulier leur fonctionnement et leur structure.

### 3.2.1 Les traitements linguistiques

Ce module saisit, en entrée, une suite de caractères alphanumériques de la phrase cible[4] et produit en sortie la suite de mots la composant ainsi que la structure morphosyntaxique qui les organise. L'analyse linguistique est essentielle pour la traduction graphème-phonème[5] et le calcul de la prosodie.

*Pourquoi ce module est-il nécessaire ?*

La suite de caractères alphanumériques que l'on introduit en entrée du système de synthèse, même exempte de toute erreur de frappe et d'orthographe, comporte des ambiguïtés. Ces ambiguïtés peuvent être de nature très variée, notamment celles qui sont liées aux homographes. Il est donc indispensable d'examiner cette chaîne alphanumérique afin de déterminer la structure de la phrase et les éléments qui la forment *(Bailly, G. & Guibert, G., 2006)*.

Les ambiguïtés étant en nombre important, ce module va se diviser en un ensemble de sous-modules ayant chacun une fonction appropriée. Nous étudions plus en détails les traitements nécessaires à la désambiguïsation de l'entrée textuelle. Nous discernons le préprocesseur, l'analyse morphologique, l'analyse contextuelle, et l'analyse syntaxique *(D'Alessandro, C. & Richard, G., 2013)*.

#### 3.2.1.1 Préprocesseur

Le rôle de ce module est d'assurer dans la transcription en toutes lettres, abréviations, nombres, des chaînes non-orthographiques (plus précisément, qui ne sont pas uniquement constituées de caractères orthographiques) ou non composées uniquement de caractères orthographiques, en exploitant des lexiques spécifiques, des règles heuristiques ou de la grammaire régulière *(Bailly, G. & Guibert, G., 2006)*.

De même pour les sigles qu'ils soient écrits avec ou sans point après chaque lettre ne semble pas donner d'indication de prononciation ; prononçable (OVNI, CNET) versus non prononçable (CPAM, SNCF).

---

[4] Dans ce contexte, le mot cible désigne la phrase à synthétiser.
[5] Appellations équivalentes : phonétisation automatique ou bien transcription orthographique-phonétique.

#### 3.2.1.2 Analyse morphologique

Les unités lexicales du français sont composées d'affixes, qui se subdivisent en plusieurs types (préfixes, suffixes, infixes, interfixes et circonfixes) et d'une base lexicale, qui s'appelle aussi radical ou lexème. Les affixes peuvent s'ajouter les uns aux autres ; un mot comme « anticonstitutionnellement», par exemple, s'analyse grossièrement ainsi :

| Préfixes | | Radical | Suffixes | | |
|---|---|---|---|---|---|
| Anti- | Con- | -stitu | -tion | -nelle | -ment |

La prononciation tient compte généralement de ces frontières morphologiques. Ainsi, la graphie S se prononce généralement [z] entre deux voyelles, sauf dans le cas où elle commence un morphème, comme dans ASYMÉTRIE ou CONTRESENS. De même, le suffixe flexionnel ENT marquant le pluriel d'un verbe est généralement muet alors que la finale ENT d'un radical nominal est généralement prononcé [ã].

Le but de ce sous-module est à la fois d'identifier les mots (l'espace n'étant pas forcément séparateur de mot), de leur associer des fonctions grammaticales possibles et d'identifier les différents morphèmes qui les composent. On établit ainsi les morphèmes – préfixes, suffixes, désinences[6] et racines des lexèmes *(Boite, R. et al.)* – sachant qu'il existe deux catégories de mots : les mots grammaticaux et les mots lexicaux[7] ; cette analyse s'exécute exclusivement sur les mots lexicaux (en nombre a priori infini : le Petit Robert compte 56 000 formes canoniques, mais on estime à 12 000 le nombre de mots apparus en français ces vingt dernières années) alors que les mots grammaticaux (qui forment le squelette syntaxique de la phrase : déterminants, pronoms, prépositions, conjonctions), en nombre fini (de l'ordre de quelques milliers), seront traités comme des unités lexicales *(Bailly, G. & Guibert, G., 2006)*.

Comme mentionné plus haut, ce sous-module a aussi pour objectif la détermination des catégories grammaticales possibles de chaque mot, complétées éventuellement par certaines variables grammaticales (nombre, genre…) grâce à ses *désinences*. Les systèmes diffèrent notamment par le jeu des catégories grammaticales (appelées *Part-of-Speech* ou PoS en anglais) considérées. On distingue souvent les noms, adjectifs, verbes et adverbes parmi les

---

[6] « Appelée aussi un morphème flexionnel, élément variable à la finale d'un mot, qui, ajouté au radical sert à marquer chacune des formes verbales […] ou nominales […] » *[ATILF]*. Pour les adjectifs, les substantifs et les pronoms il y a des désinences de genre et de nombre. Pour les verbes il y a des désinences de personne et nombre et celles qui indiquent l'infinitif, le gérondif et le participe.

[7] Décomposition en morphèmes. Exemple : « soit » combine le morphème « être » et le morphème du conditionnel présent du 3 éme personne du singulier.

mots lexicaux et les déterminants, pronoms, prépositions et conjonctions parmi les mots grammaticaux. L'analyse syntaxique peut cependant bénéficier d'un jeu plus large et donc plus précis, distinguant par exemple les adjectifs antéposés (limités à une vingtaine en français) des adjectifs postposés, les auxiliaires des verbes conjugués, infinitifs, participes présents et passés, etc. La plupart des systèmes exploite plusieurs dizaines de telles catégories grammaticales *(Bailly, G. & Guibert, G., 2006)*.

En conclusion, cette analyse permet de réduire la taille des lexiques, d'obtenir des informations sur la catégorie syntaxique des mots et d'aider ainsi le transfert phonologique.

#### 3.2.1.3 Analyse Syntaxique

Communément, chaque mot orthographique, pris isolément, a souvent plusieurs catégories grammaticales possibles. Si l'inspection du contexte local peut permettre de résoudre une partie des ambiguïtés, ces dernières ne peuvent être généralement appréhendées qu'en considérant la phrase en entier, sinon le contexte général de l'énonciation (comme l'ambiguïté nominale FILS ou adverbiale PLUS). L'objectif de l'analyse syntaxique est à la fois d'assigner à chacune des unités lexicales constituant la phrase une étiquette grammaticale unique (« Tagging » en anglais) *(Constant, M., 2011)* mais aussi d'identifier leurs relations hiérarchiques.

Les synthétiseurs actuels exploitent rarement des analyseurs syntaxiques complets et se contentent souvent de l'assignation d'une catégorie morphosyntaxique unique à chaque mot. On parlera donc plus spécifiquement de « filtre syntaxique » au sens où ces analyseurs filtrent les hypothèses émises à l'étape précédente au regard d'un ensemble de contraintes de séquencement des catégories. On distingue deux types de filtres syntaxiques : (1) ceux qui exploitent des règles heuristiques et requièrent une grande expertise de la langue avec l'avantage de pouvoir considérer des règles de séquencement à horizon variable (2) les modèles statistiques (exploitant généralement la probabilité d'observer une suite de n classes lexicales, appelés *n-grams* en anglais) qui optimisent un critère de vraisemblance sur l'ensemble de la phrase et nécessitent de grosses bases de données *(Richard, G., 2008)*.

Il s'avère que la syntaxe a un rôle fondamental sur la désambigüisation des mots et sur leur organisation en groupes hiérarchiques. La sémantique et la pragmatique y représentent également ce rôle crucial. Néanmoins, il n'existe actuellement aucun analyseur sémantique ou pragmatique capable d'analyser un texte quelconque et de résoudre ainsi des ambiguïtés du type préposition de complément de nom versus verbe comme dans « il a peint la jeune femme en noir ». Notons cependant que les modèles statistiques offrent une solution sous-optimale en

privilégiant les structures les plus observées, donc vraisemblablement le complément de nom dans le cas évoqué *(Bailly, G. & Guibert, G., 2006)*.

### 3.2.1 Le transfert phonologique

Nous rassemblons ici sous le terme de transfert phonologique l'ensemble des opérations permettant de passer de la structure linguistique de la forme écrite du texte à la structure phonologique de son énonciation. Ces opérations sont essentiellement des manipulations d'objets symboliques, d'entités discrètes en nombre fini.

#### 3.2.1.1 La traduction graphème-phonème

Le français est une langue à orthographe profonde : sans considérer l'énonciation des sigles, noms propres et des noms d'emprunts récents, la relation entre la graphie d'un mot courant et sa prononciation peut être très irrégulière. Cette prononciation dépend à la fois de la succession des lettres assemblées (par exemple, le graphème CH peut se prononcer [k] ou [ʃ] selon qu'il est suivi ou non de R ou L) mais aussi de la classe lexicale (COUVENT se prononce [kuvã] ou [kuv] selon que le mot est un nom ou un verbe) et de la liaison possible entre les mots (par exemple, GROS se prononce [gRɔz] devant AVION mais [gRɔ] devant AVAIT).

A l'issue des traitements précédents, on dispose d'un étiquetage morphosyntaxique de la phrase à prononcer, complété parfois par des indicateurs relationnels inter-mots et une segmentation du mot en constituants morphémiques. Le module de traduction graphème-phonème peut donc s'appuyer sur ces informations pour effectuer la transcription mot à mot.

On distingue deux approches : une approche s'appuyant sur un ensemble de règles morphophonologiques et une approche opérant par mise en correspondance statistique de tronçons phonographiques. Dans ce dernier cas, des algorithmes de classification (arbres de décision, chaînes de Markov, etc.) sont alors entraînés sur des lexiques de mots dont les graphies et les transcriptions phonétiques correspondantes sont alignées. L'alignement comporte alors des lettres muettes et des lettres associées avec des couples de phonèmes (par exemple XAVIER aligné avec [g&z] [a] [v] [j] [e] [_]). Dans les deux cas, les transcriptions subissent un post-traitement phonologique, qui permet de gérer la réalisation des consonnes de liaison latentes et les autres phénomènes inter-mots d'harmonisation vocalique ou d'assimilation de voisement *(Bailly, G. & Guibert, G., 2006)*.

#### 3.2.1.2 La syllabation

La syllabe est une unité très importante pour structurer l'oral. Certains modèles d'apprentissage de la lecture et de l'écriture (cf. le modèle DIAMS) considère même qu'elle constitue un pont de passage fondamental entre écrit et oral. Le nombre de syllabes d'un mot, les frontières syllabiques ont une incidence importante sur la prononciation et l'organisation rythmique de l'énonciation. Ainsi, les consonnes de liaison entre mots, même interrompues par une pause, se syllabifient avec la rime initiale du deuxième mot alors même que la graphie source est portée par le premier.

La syllabe a aussi une grande importance dans la programmation motrice et donc dans le phénomène de coarticulation. La plupart des lexiques de sons intègrent ainsi le marquage des frontières syllabiques, permettant ainsi de distinguer la prononciation de la séquence [ãfRã] dans CHIFFRE ANCIEN de celle dans GACHIS FRANÇAIS.

La syllabe a aussi une importance centrale dans la structuration métrique, qui gouverne l'alternance entre syllabes faibles et fortes, notamment dans les langues à accents.

#### 3.2.1.3 Le module phonologique

Si la syllabe est l'unité rythmique de base, de nombreux modèles phonologiques considèrent des unités de plus haut niveau que le mot. La phonologie prosodique considère ainsi l'unité intonative, délimitée par des pauses respiratoires ainsi que certains niveaux intermédiaires appelés groupes prosodiques délimités par des allongements et tons spécifiques.

De manière symétrique au module d'analyse syntaxique, le module phonologique est donc chargé de l'identification des diverses unités du discours oral ainsi que de leur organisation structurelle et des diverses fonctions linguistiques que ces unités doivent véhiculer. Ainsi, si les ponctuations (virgule, point, point-virgule, point d'exclamation, points de suspension, etc.) assument un pauvre marquage de la modalité de la phrase, notamment lors de la transcription de dialogues, on dispose de beaucoup plus d'outils sonores pour signaler la fonction d'un énoncé dans un dialogue et la position du locuteur par rapport à son discours.

Dans la plupart des systèmes de synthèse, le module phonologique se charge de segmenter la phrase en groupes prosodiques et assigner à ceux-ci un ensemble de marques spécifiques, encodant des informations linguistiques ou paralinguistiques portées par ces parties d'énoncé. Si les groupes prosodiques sont fortement contraints par la syntaxe, ils n'en sont pas une image directe : des contraintes rythmiques (notamment liés à la gestion de la respiration) interfèrent avec la nature séquentielle – gauche-droite – de la production de parole. Si le texte

est dévoilé en entier et peut donc être exploré de manière assez libre, le décodage de la parole fait face à une présentation purement séquentielle de l'information à un rythme qui semble être assez constant quelle que soit la langue *( Pellegrino, F., 2009)*.

Il y a plusieurs techniques de segmentation en groupes prosodiques d'une phrase donnée. Tout comme les traitements linguistiques vus dans la section précédente, s'affrontent les approches « basées règles » (règles qui peuvent être de primitifs heuristiques, par exemple basées sur la ponctuation, ou bien déduites d'analyses morphosyntaxiques, c'est-à-dire « la congruence syntaxe-prosodie » *(Mertens, P., 2009)* et/ou l'ajout de règles d'ajustement[8]) aux approches « basées apprentissage », souvent utilisées pour apprendre à projeter la structure syntaxique sur l'axe syntagmatique par un jeu de marqueurs (cf. ToBI).

N'oublions pas que, grâce à l'intonation, nous disposons d'un outil formidable pour associer de multiples informations linguistique à de multiples parties d'énoncé : à l'image d'un surligneur à multiples couleurs, nous pouvons à loisir segmenter une partie de discours et lui associer un ensemble de signaux conventionnels permettant de lui assigner une fonction spécifique au sein d'un ensemble de conventions linguistiques. Parmi celles-ci, on peut ainsi mentionner la simple mise en relief d'une partie ou de l'ensemble de l'énoncé, mais aussi lui associer une valence telle que le doute, le mépris ou d'autres affects. Ceci est évidemment possible si le texte est augmenté de marques, comme cela peut être le cas lors de génération de textes dans un cadre applicatif spécifique ou d'analyse de textes soumise à une expertise humaine.

### 3.2.2 La génération

Le module de génération prend en charge la conversion de la structure phonologique de l'énoncé, organisant généralement des unités phonologiques (phonèmes, syllabes, mots, groupes prosodiques et unités intonatives…) au sein d'un arbre strictement hiérarchique en un ensemble de trajectoires de paramètres caractéristiques du signal à obtenir.

On distingue souvent deux voies : une voie dite « segmentale » prenant en charge la génération des caractéristiques spectrales des sons constitutifs du discours, l'autre dite « suprasegmentale » prenant en charge la génération des caractéristiques des unités plus larges. Cette distinction est toute théorique : la plupart des systèmes actuels utilisent l'ensemble des informations phonologiques disponibles pour calculer les caractéristiques spectrales des phones et même parfois (voir HTS) font l'économie du calcul séparé des

[8] Principe d'eurythmie (harmonie des éléments).

caractéristiques des unités plus larges. Ces derniers systèmes supposent cependant que l'on dispose de corpus de parole très volumineux permettant ainsi d'aborder sans crainte l'explosion combinatoire que suppose une telle stratégie.

#### 3.2.2.1 La génération prosodique

La génération prosodique a pour objectif d'instancier les caractéristiques prosodiques des divers groupes prosodiques en fonction des informations linguistiques qu'on leur demande de véhiculer. Les vecteurs essentiels de cet encodage sont les contours mélodiques et les durées des syllabes.

##### 3.2.2.1.1 Caractéristiques mélodiques

Pour engendrer le contour mélodique en fonction de la structure prosodique de la phrase, trois méthodes de synthèse sont souvent proposées : les modèles de commande mélodique, les modèles de contours pré-stockés et les modèles directs.

*Les modèles de commande.* Ces modèles supposent qu'on ne génère pas directement la courbe mélodique mais un ensemble de commandes qui excitent des réponses spatio-temporelles spécifiques. Trois modèles sont ainsi proposés dans la littérature : le modèle de Fujisaki (excitation de source vocale) *(Fujisaki, H. et al., 1984)*, les modèles de points cible (cf. Momel) et enfin le modèle de l'école hollandaise. Le modèle de Fujisaki suppose que le contour mélodique est le résultat de la superposition de réponses de filtres linéaires à différents types d'entrées (échelon et impulsions). Les caractéristiques de ces filtres sont supposées invariables dans le temps et sont déterminées par la physiologie de l'individu *(Bailly, G. & Guibert, G., 2006)*.

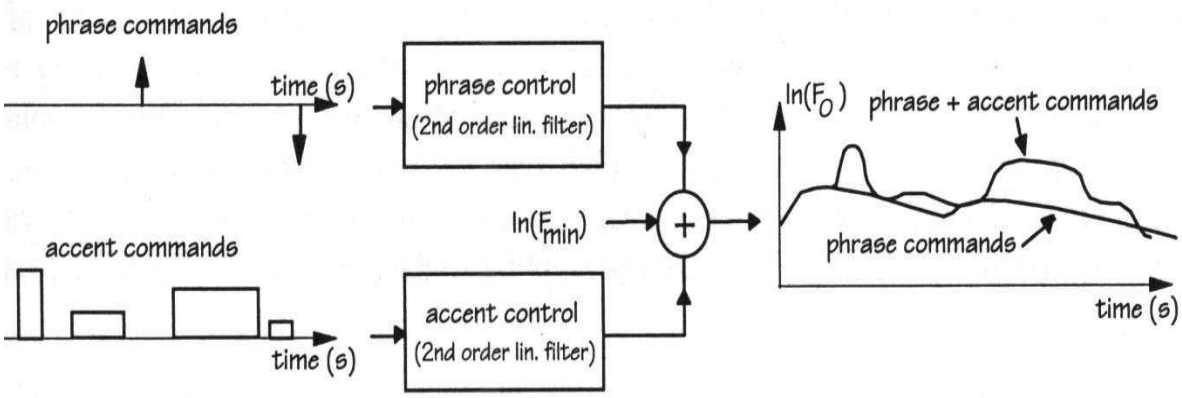



**Figure 7. Modèle de Fujisaki (d'excitation de source vocale).**

Dans les modèles par points cible, on suppose que l'information est portée par les extrema de la fréquence fondamentale et que les fonctions de transition ne sont pas contrôlées explicitement.

Le dernier modèle est celui de l'école hollandaise (développé à l'Institut voor Perceptieonderzoek dans les années 1960) qui suppose que l'information prosodique est portée par des contours élémentaires en nombre fini évaluant entre deux lignes de déclinaison.

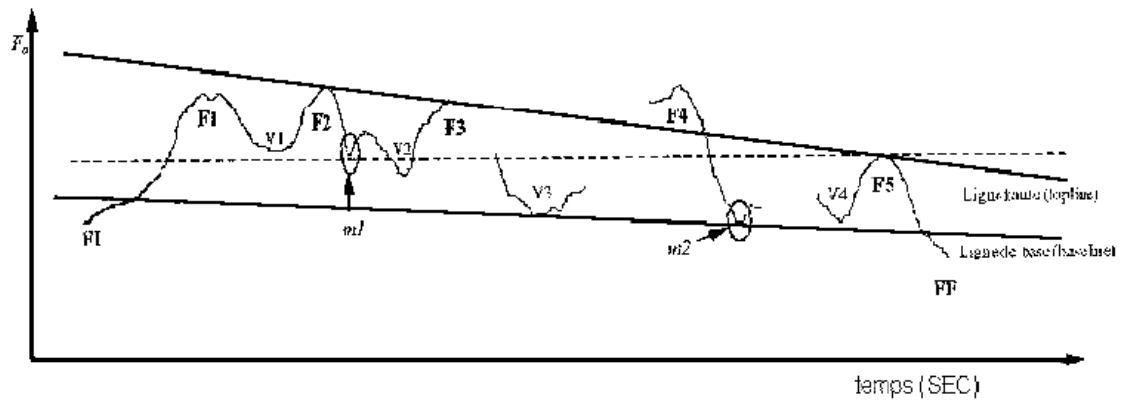



**Figure 8. Stylisation (école hollandaise) : ligne de déclination et stylisation en contours standards.**

*Les modèles de contours pré-stockés* indexent des dictionnaires de formes mélodiques par des indices phonologiques et phonotactiques. Ces modèles supposent l'existence de contours prototypiques dont la forme se déploie de manière dynamique sur une partie d'énoncé en fonction de la taille de ce dernier *(Bailly, G. & Guibert, G., 2006)*.

*Les modèles directs* procèdent quant à eux à une cartographie directe des liens entre caractéristiques phonologiques et phonotactiques et caractéristiques prosodiques. Chaque syllabe est alors caractérisée par un nombre de caractéristiques discrètes élémentaires en fonction de son appartenance à divers groupes prosodiques (natures, longueurs et positions des groupes…) et de caractéristiques mélodiques en nombre constant (par exemple, stylisation de chaque syllabe par un spline ou échantillonnage du contour en un nombre fixe de points…). Des techniques de mise en correspondance (arbres de régression, réseaux de neurones, etc.) sont alors mises en œuvre pour apprendre cette correspondance entre caractéristiques symboliques et paramétriques *(Bailly, G. & Guibert, G., 2006)*.

#### 3.2.2.1.2 Caractéristiques rythmiques

La majeure partie des modèles de rythme sont des modèles directs qui calculent la durée de chaque phonème de la phrase en fonction de caractéristiques phonologiques et phonotactiques. Si les premiers modèles de durée restreignaient l'interaction des caractéristiques phonologiques et phonotactiques à des additions, multiplications ou sommes de produits *(Bartkova, K. & Sorin, C., 1987 ; Klatt, D. & O'Shaughnessy, D., 1981)*, les modèles actuels procèdent par apprentissage direct.

A noter que certains modèles de médiation syllabique ont été proposés *(Campbell, W.N. & Izard, D.)* qui supposent que le rythme contrôle une horloge syllabique dont la fréquence instantanée est fonction des caractéristiques phonologiques et phonotactiques courantes. Cette fréquence instantanée peut être contrainte par un poids syllabique mais ceci est d'un deuxième ordre. Un microcontrôleur est alors sollicité pour répartir la durée de chaque période aux unités segmentales constitutives.

#### 3.2.2.1.3 Commentaires

Il y a bien sûr des systèmes de génération de la modification de l'énergie du signal et d'autres paramètres spectraux – notamment de pente spectrale – par rapport à la structure prosodique, mais ces modèles sont peu utilisés et contribuent plus modestement à la qualité de la synthèse.

Notons cependant que les systèmes préalablement mentionnés désunissent les paramètres prosodiques à engendrer, la cohésion étant essentiellement assurée par la distribution séquentielle de caractéristiques phonologiques et phonotactiques identiques à chaque module. Cependant, l'arrivée d'outils de mise en correspondance (en anglais « mapping») performants, une carte de passage entre paramètres d'entrée et de sortie, a permis la naissance de modèles prosodiques multiparamétriques combinant indices segmentaux et suprasegmentaux.

### 3.2.2.2 La génération segmentale

Ainsi, nous avons à notre disposition toutes les informations, acquises à partir des modules linguistiques et prosodiques, obligatoires pour générer le signal acoustique, soit l'identité des phonèmes, leur durée et la progression de la fréquence fondamentale F0. Il demeure désormais à appliquer ces informations de manière à produire un signal de parole.

Il y a diverses méthodes de synthèse, nous les dissocions en trois grandes catégories : les systèmes de synthèse par règles, les systèmes de synthèse par sélection/concaténation et les systèmes statistiques. Les premiers ont pour objectif de modéliser et d'analyser le phénomène

de production du signal acoustique, tandis que les autres exploitent de larges corpus de données étiquetées. Nous étudions plus en précision ces trois catégories de systèmes leur mode de fonctionnement, leurs avantages et leurs désavantages *(Bailly, G. & Guibert, G., 2006)*.

#### 3.2.2.2.1 Synthèse par règles

Cette catégorie de système fut la première à voir le jour. Réellement, il existe deux raisons à ceci : (1) les techniques de l'époque ne permettaient pas de stocker beaucoup de données, (2) les chercheurs de l'époque s'inscrivaient dans une tradition de phonétique descriptive de déchiffrage des mécanismes de planification de la production de la parole, notamment les phénomènes de coarticulation. L'obstacle principal du synthétiseur par règles réside dans la nécessité de règles suffisamment générales pour échapper à une description combinatoire des transitions et cibles des phonèmes en contexte. On mentionnera, par exemple, le modèle de coarticulation d'Öhman *(1966)* dans lequel l'articulation perturbatrice des consonnes se superpose à l'articulation continue des voyelles.

D'après G. Richard *(2008)*, ce système a quelques avantages tels que la faible quantité de données à stocker et la progression des connaissances sur la parole. Néanmoins, il a aussi quelques désavantages tels que l'établissement fastidieux de règles qui dépendent en grande partie de la langue et en moindre partie du locuteur.

#### 3.2.2.2.2 Synthèse par sélection et concaténation

L'idée fondamentale de ce type de système est de capturer le phénomène de coarticulation au sein d'unités acoustiques de tailles variables que l'on va concaténer. Ce type de système ne nécessite pas de connaissances explicites sur la production de parole, les obstacles majeurs étant au niveau du choix des unités à concaténer et au niveau du lissage acoustique post-concaténation.

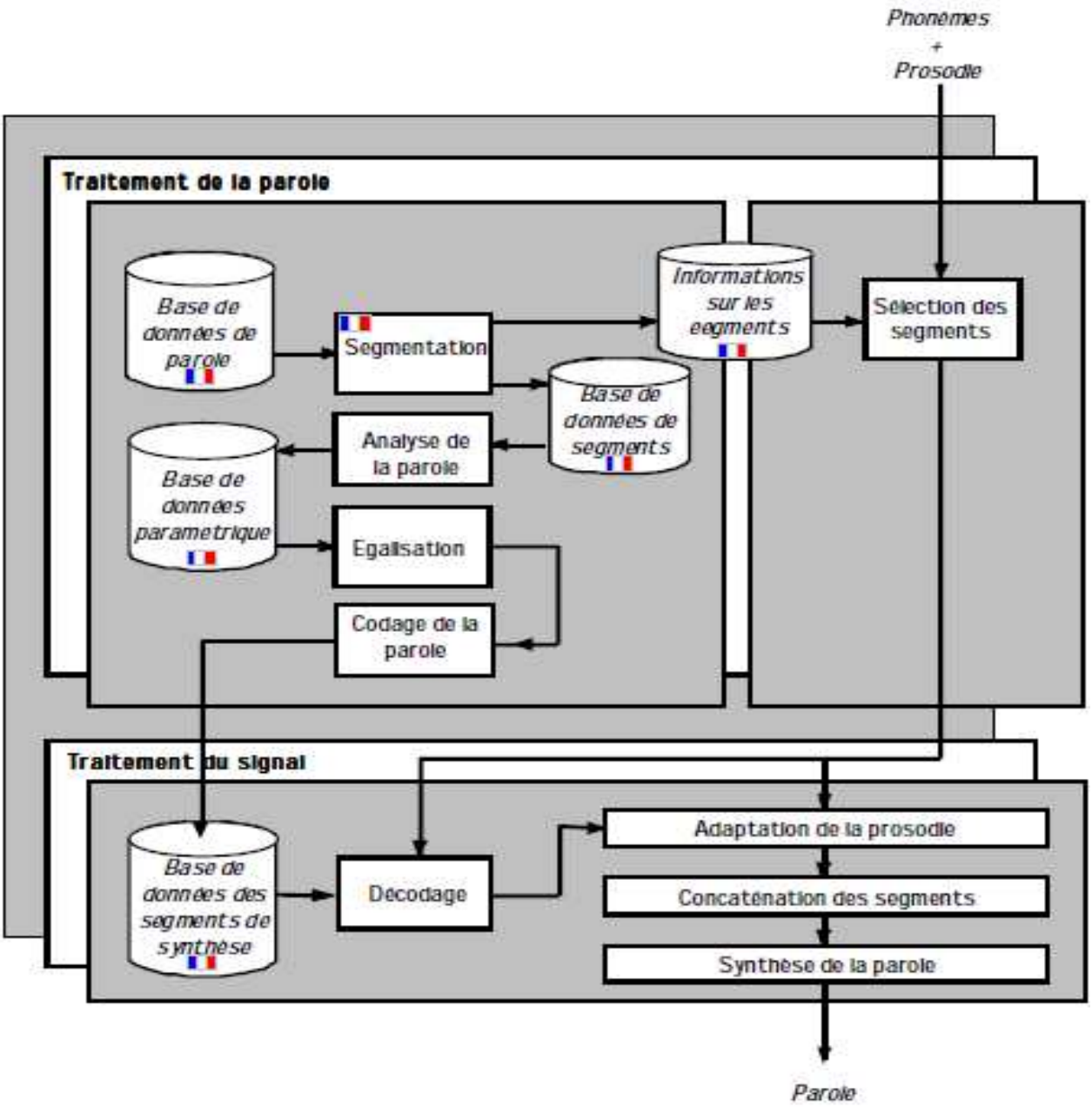



**Figure 9. Synthétiseur par concaténation : les opérations qui dépendent de la langue sont indiquées par un drapeau.** ***(Dutoit, T.)***

### *3.2.2.2.2.1 Choix des unités élémentaires*

La sélection des unités de base est primordiale : ces unités doivent encapsuler autant que possible le phénomène de coarticulation, être commodément concaténables (homogénéité de la voix) et autoriser une couverture large de la langue. Si la syllabe est un candidat évident et a été largement exploitée en Japonais, le nombre de syllabe en français est dissuasif. Le diphone[9] convient bien à ces objectifs : on encapsule la portion de signal acoustique comprise

[9] « Environ 1200 en français ; environ 5 Mo d'échantillons sur 16 bits à 16 khz.» *(Richard, G., 2008)*

entre les parties stables de deux phones successifs, on concatène des portions stables du signal et on a un inventaire réduit de segments (environ 1400 pour le français). Nonobstant, la qualité du signal de synthèse s'améliore avec la taille des unités, par conséquent, d'autres types d'unités comme les dissyllabes[10] (début et fin sur des voyelles), les polyphones[11] (triphones, quadriphones) ou plus récemment les sandwichs (début et fins sur les consonnes) ont été présentées. Certains systèmes appliquent ainsi des polysons, soit des mixtures d'unités de longueur variées : demi-phones[12], phones, diphones, dissyllabes, polyphones, mots, etc. Lors de la sélection, ce sont les unités les plus grandes possibles qui seront conservées *(Bailly, G. & Guibert, G., 2006)*.

#### *3.2.2.2.2.2 Sélection des unités de synthèse*

Lorsqu'elles sont multi-représentées, les unités sélectionnées doivent présenter le maximum de critères segmentaux et suprasegmentaux[13] (typiquement : pitch et durée) liées aux unités ou caractéristiques (si on suppose la prosodie préalablement calculée) prosodiques.

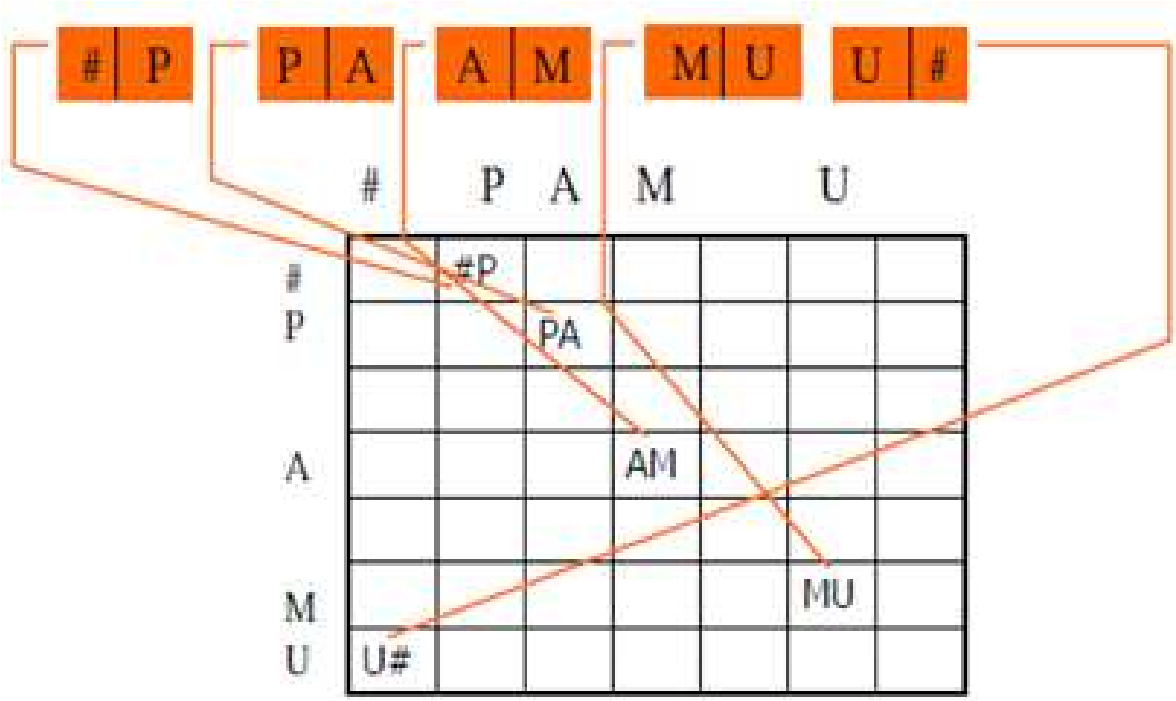



**Figure 10.Illustration d'une sélection statique (diphone), le symbole # est une indication de pause.**

---

[10] « Supérieur à 10 000 unités ; environ 50 Mo de données.» *(Richard, G., 2008)*

[11] « Ex en français : Triphones: 153 incluant [j], 54 incluant [w], 544 incluant [l] ou [r]. Quadriphones: 288 incluant [j] et soit [l] soit [r]. » *(Richard, G., 2008)*

[12] Est la plus petite unité.

[13] La plupart des linguistes de tradition américaine considèrent les notions de traits suprasegmentaux et de prosodie comme globalement équivalentes. *(Wilhelm, S.)*

Le coût de sélection est fréquemment composé d'un *coût cible* ($C_t$), pré-calculé au moment de la présélection qui désigne la distance entre un diphone cible et un diphone candidat, et un *coût de concaténation* ($C_c$), résolu lors du parcours du treillis de possibilités qui désigne la qualité de la concaténation *(Colotte, V. & Beaufort, R.)*. Un algorithme de programmation dynamique ou de simplex permet de sélectionner le chemin optimal (au sens du coût de sélection minimal).

$$C(t,u) = \sum_{i=1}^{n} \Omega_t C_t(t_i, u_i) + \sum_{i=2}^{n} \Omega_c C_c(u_i, u_{i-1})$$

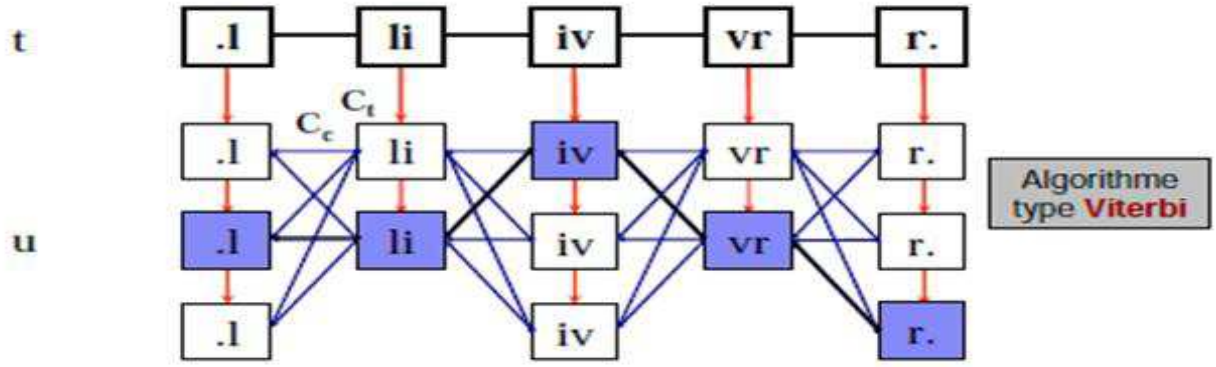



**Figure 11. Illustration de l'algorithme de sélection ; L'algorithme de Viterbi est en général utilisé pour optimiser cette recherche du meilleur chemin.**

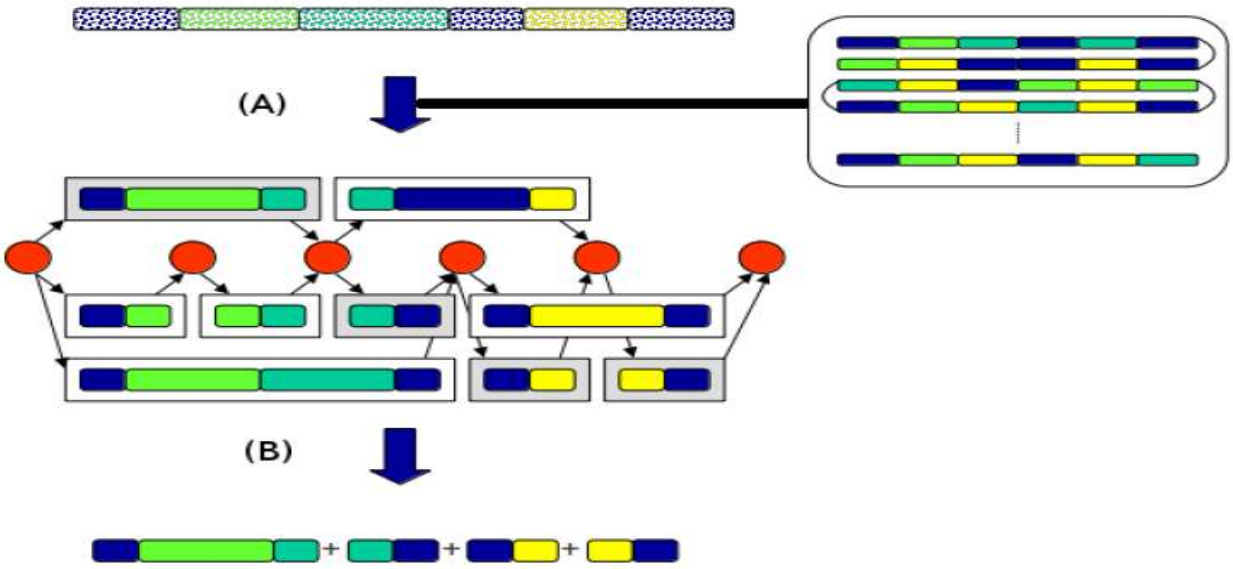



**Figure 12. « Principe de la synthèse par corpus. Dans un premier, en se basant sur une consigne définie et un corpus de parole annoté, un graphe d'unités candidates est constitué (A). En utilisant un algorithme de Viterbi, les unités optimales (au sens des coûts définis dans le système de synthèse) sont sélectionnées puis concaténées (B). »** ***(Le Maguer, S., 2013)***

La synthèse par corpus (SPC) généralise la méthode de synthèse par sélection et concaténation (cf. figure 15) à des unités de tailles et de natures variables. Ils ont considérablement perfectionné la qualité de la synthèse vocale. Leur succès a appuyé sur la concaténation d'unités acoustiques extraites de bases de données audio. Ces systèmes marchent parfaitement sur des corpus homogènes de parole lue ou de parole dite neutre. Pourtant, la qualité de ce type de synthèse se détériore au fur et à mesure que les corpus réduisent en taille ou accroissent en richesse et en expressivité prosodique. Dans ces contextes, un modèle prosodique devient obligatoire afin de mener la sélection des unités acoustiques ou préciser les modifications prosodiques. L'apprentissage prosodique reste pourtant compliqué, vu que les facteurs influençant la prosodie sont assez complexes ; on parle ici de l'attitude ou l'intention du locuteur, ou encore la place de l'accent dans le groupe accentuel. Pareil dans les descriptions linguistiques les plus avancées appliquées en "Concept-To-Speech", il n'y a pas de relation biunivoque[14] entre description linguistique et prosodie *(Boidin, C. et al., 2008)*.

En guise de conclusion, la SPC permet d'avoir une parole synthétique de très haute qualité. C'est pour cela qu'elle s'est peu à peu imposée comme un standard dans le monde de la synthèse vocale. Les performances d'un système de SPC dépendent notamment de deux facteurs principaux : d'un côté de la richesse du corpus de synthèse, et de l'autre l'algorithme de sélection appliqué. Pour une revue plus exhaustive, le lecteur pourra se référer à *(Le Maguer, S., 2013)*.

#### 3.2.2.2.3 Synthèse paramétrique par HMM

Au milieu des années 1990, l'évolution des machines de calcul a donné naissance à une synthèse paramétrique basée sur des modèles statistiques dérivés des techniques de reconnaissance de parole. De nos jours, le système de synthèse HTS (HMM-based Speech Synthesis System, qui a été réduit en HMM Triple S puis en HTS) (cf. figure 18), est le plus répandu. Ce système a le privilège d'engendrer une synthèse intelligible et fluide en exploitant moins de données qu'un système par sélection d'unités. Néanmoins, inversement à la synthèse par sélection, le signal de la synthèse engendré par HTS est de plus faible qualité.

L'idée fondamentale est ici de décrire les trajectoires de phones en contexte par des chaînes de Markov entrainées sur des très grands corpus. Le contexte est ici à prendre au sens large, il inclut un ensemble important de descripteurs segmentaux et suprasegmentaux.

---

[14] Qualifie une liaison mathématique où à chaque élément d'un ensemble correspond un élément et un seul de l'autre ensemble

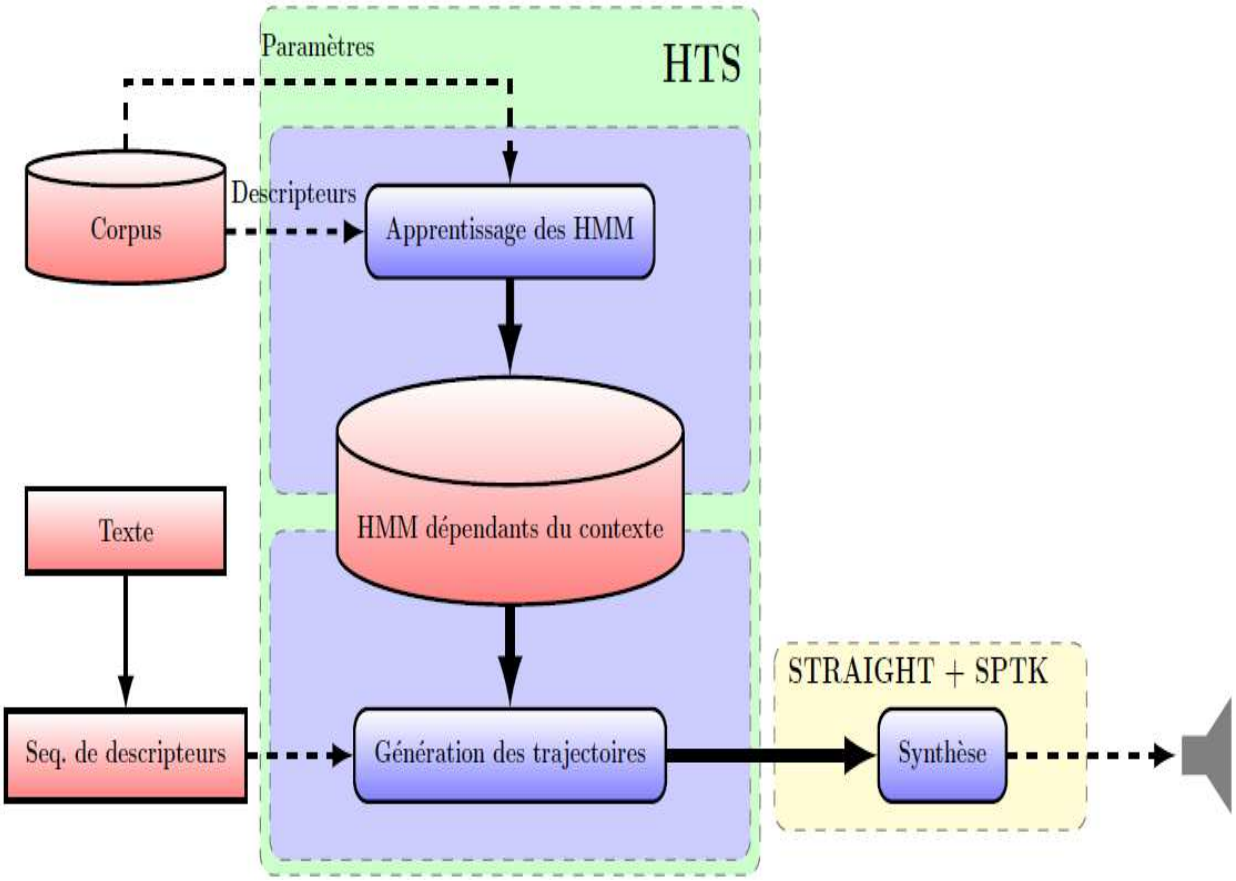



**Figure 13. Le système HTS : le corpus est composé du signal paramétrisé, en appliquant les outils STRAIGHT *(Kawahara 1999 ; Kawahara 2001)* et SPTK *(Fukada 1992)*, avec les descripteurs qui permettent d'identifier un segment. En s'appuyant sur ce corpus, les HMM sont appris. La séquence de paramètres qui correspond à la séquence de descripteurs définis par les outils d'analyse linguistique, est générée lors de la phase de synthèse. Les outils STRAIGHT et SPTK permettent d'avoir le signal acoustique à partir des paramètres générés par HTS. *(Le Maguer, S., 2013)***

Le Maguer *(Le Maguer, S., 2013)* a évalué, dans le cadre d'une synthèse en français, l'influence des descripteurs sur la modélisation effectuée par HTS. En effet, il existe deux limites importantes résultant de la combinatoire des descripteurs ***(cf. annexe A)***. Primo, il est impossible d'enregistrer un corpus permettant de couvrir l'ensemble des combinaisons possibles. Cela signifie que, lors de la phase de synthèse, des combinaisons que l'on a besoin de synthétiser peuvent ne pas avoir été vues lors de la phase d'apprentissage. Secundo, comme le système HTS appuie sur une modélisation statistique, le nombre d'occurrences associées à une combinaison de descripteurs peut aussi être excessivement faible afin obtenir un modèle convenable. Les informations contextuelles sur les quatre phonèmes adjacents suffisaient largement à couvrir la variabilité spectrale des sons alors que l'ajout d'informations syntaxiques était déterminant, pour la prosodie mais ne permettaient pas d'approcher asymptotiquement l'erreur minimale de prédiction.

En conclusion, HTS requiert des volumes importants de données et une large couverture des contextes segmentaux et suprasegmentaux. Le recours à une paramétrisation uniforme du signal engendre une perte de qualité du signal de synthèse. Néanmoins, cette dernière permet de manipuler le signal plus naturellement ce qui rend la synthèse paramétrique plus flexible que les systèmes par corpus. Actuellement, HTS est un des systèmes de synthèse référent.

### 3.2.3 Techniques de synthèse

Les représentations paramétriques du signal étant ainsi déterminées, il est parfois nécessaire de disposer d'une représentation permettant d'effectuer les changements au niveau temporel et au niveau spectral de manière à ce que la prosodie calculée soit appliquée. Les techniques PSOLA[15] (Pitch Synchronous Overlap-Add) sont les plus utilisées dans ce cadre. La version temporelle de cet algorithme s'appelle TD-PSOLA[16] (Time Domain Pitch Synchronous Overlap-Add). Le principe est d'effectuer un découpage du signal de parole en une série de signaux élémentaires se recouvrant dans le temps. En choisissant la fenêtre astucieusement, on peut garantir que le signal est inchangé si aucune modification n'y est apportée, ce qui est garanti par la fenêtre de Hanning ($\cos^2+\sin^2=1$). Dans le cas d'un signal voisé, on aligne le centre de chaque fenêtre sur l'excitation maximale de chaque période. Pour les portions non-voisées, on génère des pseudo-périodes et on les traite de manière spécifique dans le cas où on doit les dupliquer Ensuite, on additionne les signaux fondamentaux, en faisant varier les espacements entre signaux élémentaires et on les duplique/supprime pour s'harmoniser aux contraintes prosodiques. Cet algorithme est facile à réaliser, peu coûteux en calcul et donne de très solides résultats. En revanche, il peut y avoir des cas de discontinuités lors de la concaténation (discontinuité de pitch, de phase et d'enveloppe spectrale).

D'autres systèmes ont été ainsi introduits, tel LP-PSOLA[17] dans lequel un modèle paramétrique de type LPC[18] (Linear Predictive Coding) est couplé à la méthode TD-PSOLA, ou encore MBROLA (Multi Band Resynthesis Overlap and Add) dans lequel l'ensemble des unités de la base sont auparavant resynthétisées sur plusieurs bandes (il en va de même pour NU-MBROLA qui est une extension de MBROLA pour des unités non uniformes) *(PEETERS, G. & Le Beux, S., 2009).*

---

[15] Cette technique a une synchronie à la période fondamentale tant à l'analyse qu'à la synthèse. Ceci permet un contrôle à la fois du déroulement de l'axe temporel et de la hauteur du signal.
[16] La technique, dite d'addition recouvrement de fenêtres temporelles synchrones avec le pich, applique le principe de la réharmonisation spectrale directement sur le signal de parole.
[17] L'approximation de l'enveloppe spectrale est explicite et plus précise.
[18] Un des avantages de la synthèse LPC provient des excellentes propriétés de compression des paramètres mis en œuvre.

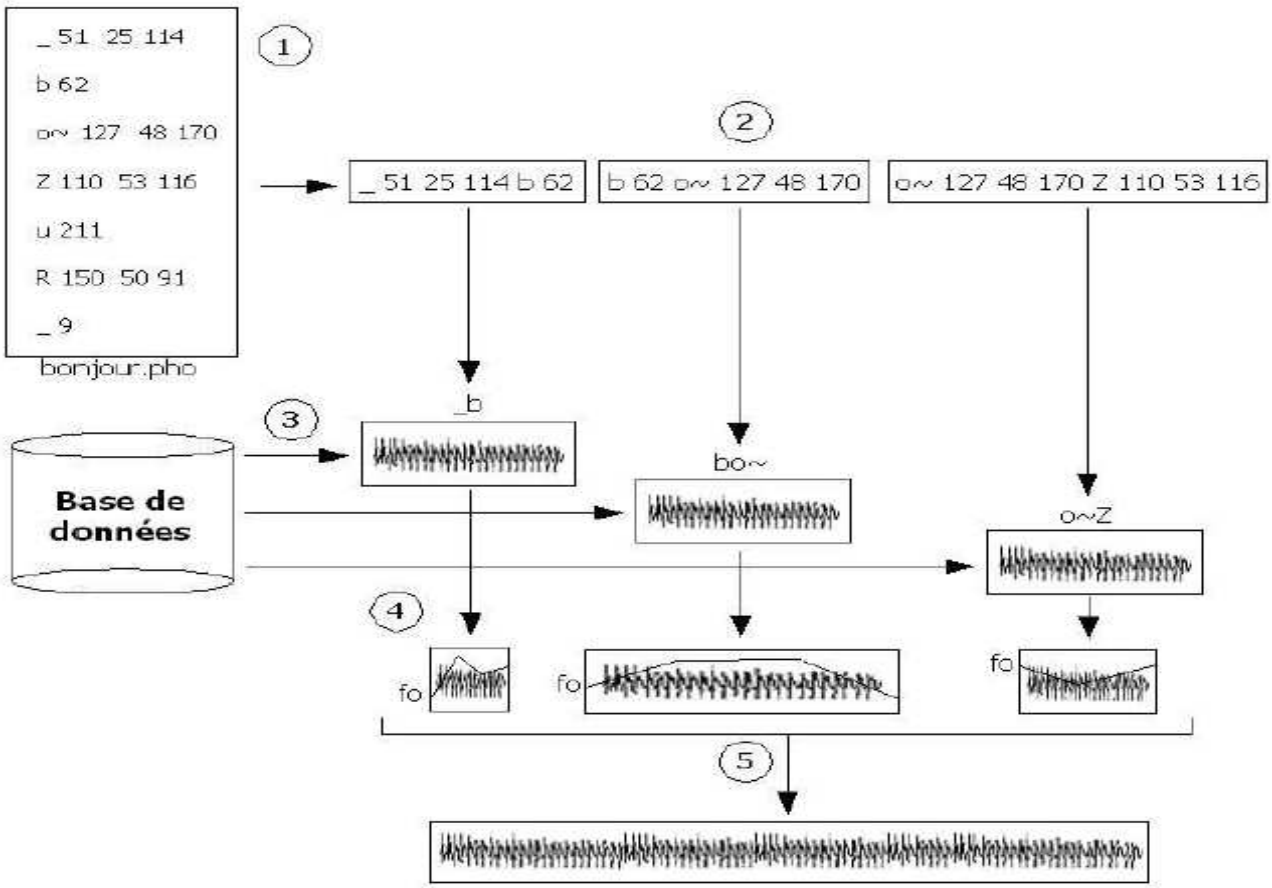



**Figure 14.** **Principe de fonctionnement de MBROLA.**

Les algorithmes ci-dessus sont fréquemment appliqués dans le cas où le nombre d'unités est peu important dans le dictionnaire et/ou au moment où l'unité de base est petite.

Dans le cas de gros dictionnaires où les unités sont multi-représentées, d'autres méthodes, orientée sur la sélection et les coûts combinés, sont appliquées. Effectivement, moins le signal sera modifié, plus il conservera un aspect naturel. Dans le cas de gros corpus, les unités qui sont présentes dans de nombreux contextes avec des caractéristiques prosodiques diversifiées, vont probablement découvrir une chaîne d'unités qui, sans traitement du signal a posteriori, sera proche de la prosodie définie par l'analyse de la phrase d'entrée. Les méthodes de sélection et le choix des coûts deviennent alors importants *(Bailly, G. & Guibert, G., 2006)*.

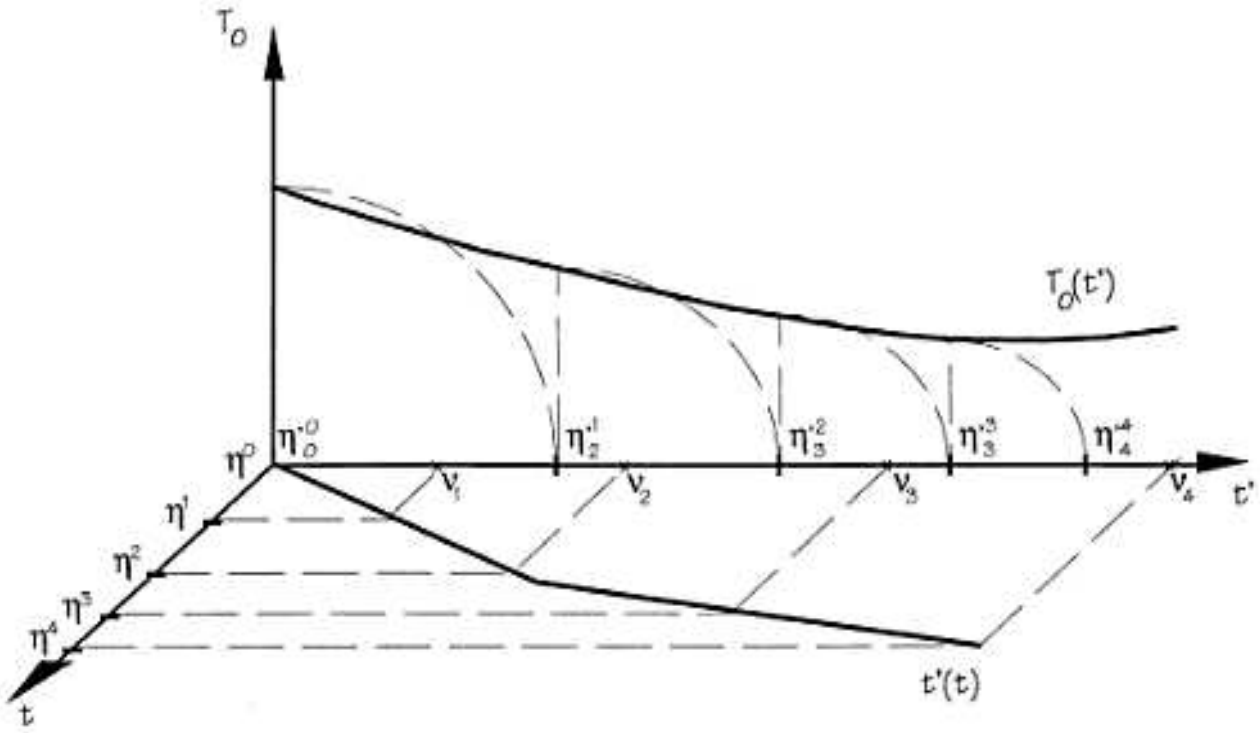



**Figure 15. Modification de la durée et du pitch avec TD-PSOLA. Partant de cinq fenêtres d'OLA sur l'axe temporel à l'analyse (t), on obtient quatre marqueurs de pitch sur l'axe du temps de synthèse (t'). Dans cet exemple, la deuxième fenêtre d'analyse a été éliminée ; la troisième a été dupliquée.**

En conclusion, des méthodes de lissage sont, conséquemment, utilisées et permettent la suppression de ces artefacts aux frontières. Elles se caractérisent par les paramètres spectraux, par la durée sur laquelle on les exécute et par le genre d'interpolation. Une approche probable est de faire un lissage au niveau de ces frontières. Dans le cas où la première voie ne prendrait pas en considération le contexte dans lequel sont situées les unités à concaténer, une deuxième voie est de réaliser le lissage en fonction de l'information contextuelle *(Bailly, G. & Guibert, G., 2006)*.

## 3.3 Conclusions

Dans ce chapitre, nous avons vu que la synthèse de la parole (audio) à partir du texte était un mécanisme compliqué faisant agir plusieurs modules, eux-mêmes formés de sous-modules ayant des tâches particulières. De ce fait, à cause de la complexité des signaux d'entrée (entrée textuelle) et de sortie (signal audio), il n'existe pas de solution unique à la problématique de la synthèse de parole à partir de textes.

Dans le cadre de ces travaux du master, nous avons utilisé un système de synthèse par concaténation de polysons multi-représentés qui utilise une modélisation distincte et explicite de la prosodie, dont nous allons donner les principes ci-après.

*« La voix ne peut porter la langue et les lèvres qui lui donnent des ailes. Seule, elle affronte l'air.»*
*[Khalil GIBRAN, le Prophète]*

# 4 MODÈLE PROSODIQUE UTILISÉ : SFC

## 4.1 Introduction

Deux approches complémentaires, mais pourtant opposées dans leurs aboutissements, ont été présentées dans la littérature pour spécifier la trajectoire du paramètre principal de la structuration intonative qui est la ligne mélodique.

## 4.2 Analyse ascendante (bottom up)

L'analyse ascendante (cf. figure 15) propose de caractériser la prosodie comme une succession d'événements saillants et de caractéristiques définies a priori. Des méthodes de stylisation automatique de la ligne mélodique ont été ainsi proposées qui considèrent la ligne mélodique comme une suite de tons syllabiques *(D'Alessandro, C. & Richard, G., 2013)*, d'accents : modèle TILT (Taylor, P., 2000), une superposition de formes prédéfinies (Fujisaki) ou une interpolation entre points-cibles (Hirst, D., 1991 ; Pierrehumbert, J. 1981).

Dans la prosodie autosegmentale, ces évènements sont généralement les feuilles d'arbres phonologiques plus ou plus profonds (voir discussion sur la récursivité éventuelle des structures phonologiques, selon Gussenhoven (1991)). La prosodie autosegmentale distingue notamment des unités intonatives constituées d'un ou plusieurs groupes prosodiques (ou accentuels selon les modèles). Enfin, un modèle phonologique *profond* charge d'attribuer à ces éléments structurants des fonctions linguistiques (dont notamment la correspondance de ces évènements avec les représentations morphosyntaxiques voire sémantiques ou pragmatiques) *(Bailly, G., 2007)*.

En guise de conclusion, cette approche vise à aller de la prosodie vers les fonctions par le biais des processus de stylisation progressifs. Ainsi, les guides d'étiquetage recommandent souvent de ne s'intéresser qu'aux signaux en évitant de se laisser influencer par un modèle théorique ou un quelconque biais perceptuel.

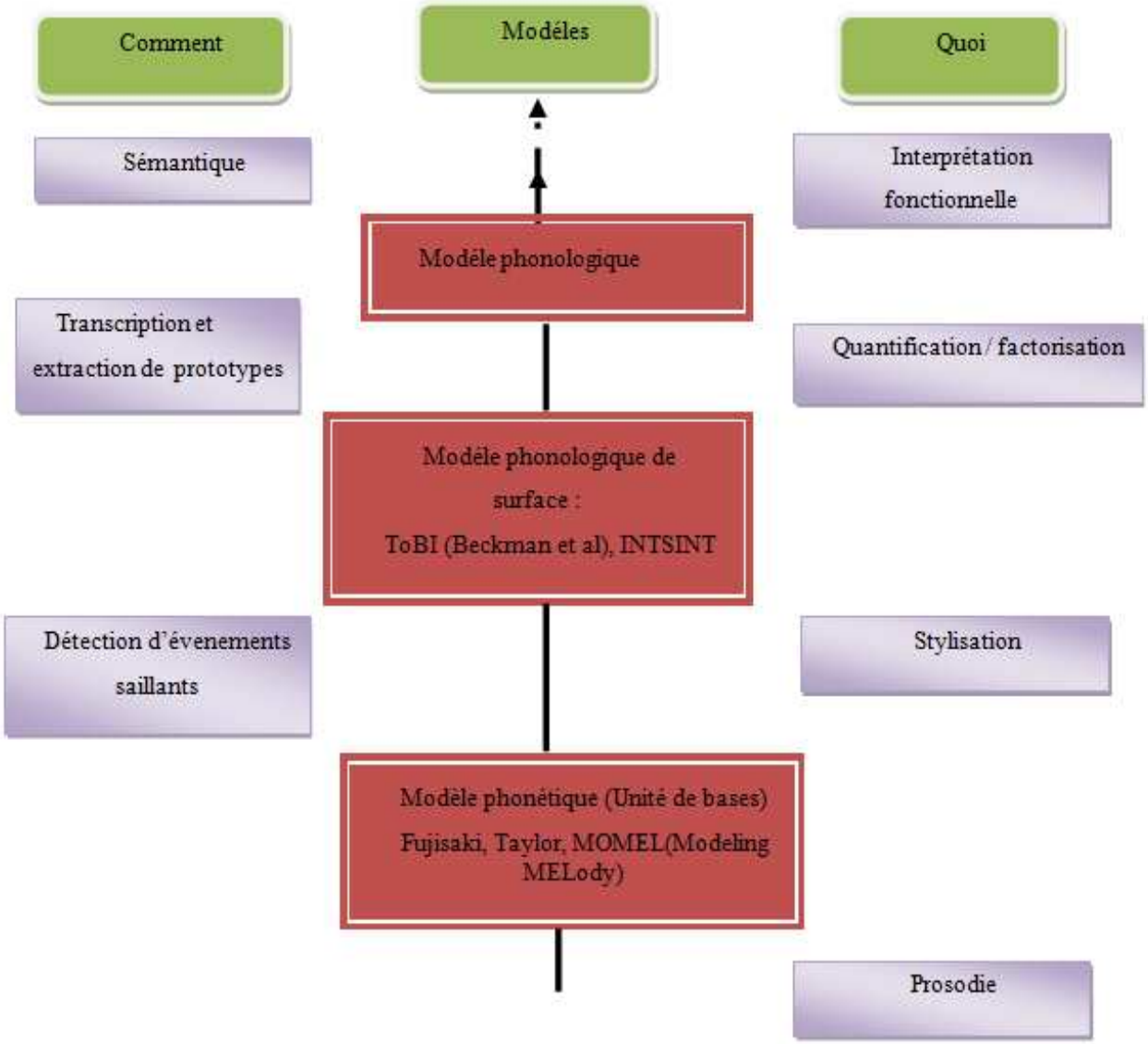



**Figure 16.** **Analyse ascendante.**

### 4.3 Analyse descendante (top down)

En miroir, l'analyse descendante (cf. figure 16) vise à induire des formes prosodiques en partant des fonctions linguistiques. Les formes ou évènements ne sont plus forcément saillants mais caractéristiques : ils portent de l'information, *i.e.* ont une représentativité statistique adéquate pour transmettre des contrastes linguistiques ou paralinguistiques.

A l'opposé de la méthode précédente qui part d'une stylisation semi-automatique de signaux, cette approche implique d'avoir à disposition des corpus importants. La couverture statistique, la représentativité des contrastes fonctionnels que l'on veut caractériser doivent être suffisants pour dériver les formes caractéristiques de manière contextuelle (position, empan, nombre de fonctions impliquées, etc.). Ainsi, la méthodologie d'extraction de

concours mélodiques suggérée par Aubergé *(1992)* exige la construction de phrases où la position et la longueur des unités linguistiques sont systématiquement variées *(Bailly, G., 2007)*.

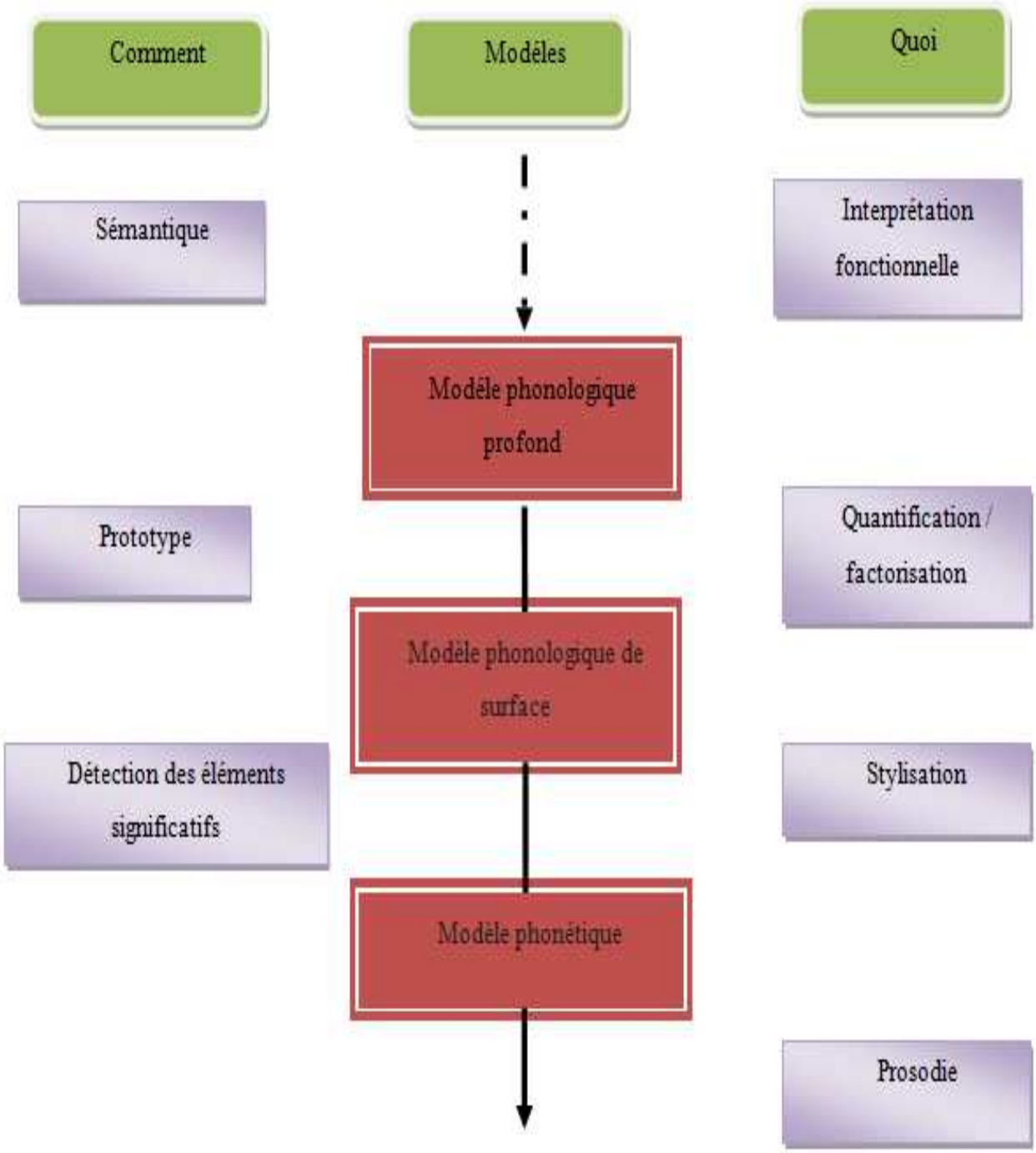



**Figure 17. Analyse descendante**

## 4.4 Motivation du SFC

Le SFC proposé par Bailly et Holm *(Bailly, G. & Holm, B., 2003)* n'a bien évidemment pas émergé du néant. Il doit beaucoup aux travaux de Morlec et Barbosa, et à travers eux aux propositions originales d'Aubergé.

### 4.4.1 Un lexique structuré de formes prosodiques globales

Aubergé *(1991 ; 1992)* fonde un lexique de contours prosodiques, en appliquant la méthode de paires minimales qui a été introduite par Delattre *(1969)*. Elle concrétise l'idée que l'intonation et la syntaxe coopèrent dans le but de transmettre le message sous la contrainte d'un alignement souple : unités intonatives et constituants syntaxiques ayant des points d'ancrage communs. C'est ainsi qu'elle suggère la notion de rendez-vous entre structure intonative et syntaxique.

Aubergé distingue quatre niveaux syntaxiques (phrase, proposition, groupe et sous-groupe) avec lesquels sont combinés des contours prosodiques indépendants. Ces derniers interfèrent par simple superposition dans le but de fabriquer le contour global de l'énoncé. Notons bien qu'en analyse, les contours sont obtenus d'une manière hiérarchique en soustrayant les niveaux déjà modélisés.

Au niveau de la phrase, des contours moyens (c'est ce que délivrent les générateurs de contours du SFC : un contour moyen des instances de la fonction socio-communicative sur un empan donné) sont rangés dans le lexique en fonction des attributs qui les distinguent. L'analyse des niveaux sous-jacents est faite de la même manière, après avoir retiré la contribution des niveaux supérieurs du contour observé. Ainsi, cette méthode permet de quantifier l'impact de certaines propriétés linguistiques sur la constitution prosodique des énoncés : soit, en d'autres termes : quand et comment syntaxe et prosodie respectent leurs rendez-vous ?

Soulignons que dans la première mise en œuvre de ce modèle, les contours des niveaux non-terminaux sont des lignes de déclinaison ordinaires et que seules les unités en bas de l'échelle hiérarchique accèdent à des contours "détaillés".

Cette proposition suppose donc un encodage très direct d'informations linguistiques par un contour prosodique correspondant : à chaque sous-groupe est associé un contour du lexique indexé par les attributs qui le marquent. Ces contours sont enchaînés, et suivent les lignes de déclinaisons des niveaux supérieurs. Les contours manipulés sont donc : associés au niveau

hiérarchique, coextensifs aux unités identifiées auparavant et caractérisés par des attributs assez concis *(Bailly, G. & Holm, B. 2003)*.

### 4.4.2 Contours dynamiques et coarticulés

Morlec *(1997)* fournit à ce modèle une mise en œuvre d'une plus ample souplesse, en supprimant la contrainte des lignes de déclinaisons associées aux niveaux supérieurs. Aux contours stockés dans un lexique, sont substitués des générateurs de contours (implémentés sous forme de réseaux de neurones récurrents) paramétrés de manière simple par la localisation absolue et relative des points d'ancrage du contour.

La génération dynamique des contours admet de se libérer de la contrainte de devoir les stocker pour toutes les longueurs possibles. En outre, elle permet l'extraction de contours de haut niveau qui ne sont plus contraints à être des lignes de déclinaison.

De même, les travaux de Morlec sur les attitudes prosodiques ont mis en évidence l'existence de clichés mélodiques *(Fónagy, I. et al., 1984)* représentés par un modèle d'expansion du mouvement. Grâce à une expérience de dévoilement progressif de phrases prononcées avec différentes modalités, (Aubergé, V. et al., 1997) ont ainsi montré que les auditeurs avaient accès de manière très précoce à la fonction attitudinale d'un contour de phrase. Ces expériences confirmées par les travaux de Van Heuven *(Van Heuven, V.J. et al., 1997)* vont dans le sens d'une comparaison dynamique entre contour réalisé et prototypes stockés avec une décision possible dès que le point d'unicité du contour est atteint.

Chez Morlec, les contours manipulés ne sont plus associés de manière univoque aux unités : chaque générateur de contours encode une fonction qui peut indistinctement être appliqué à toute unité de discours. A la spécification des unités chez Aubergé est substitué un marquage assez simple des unités, qui indique soit une fonction intrinsèque à l'unité (mise en emphase, valeur modale ou attitudinale, etc.) soit une relation qu'elle entretient avec ses voisines. Ce marquage relationnel est l'héritage de Bailly *(Bailly, G., 1983)* ***(cf. annexe B)*** qui, à son tour, s'est inspiré de Hays *(1964)*.

A chaque marque est associé un générateur de contours, qui est responsable de la génération d'un contour élémentaire sur l'unité marquée (et éventuellement ses voisines dans le cas d'un marqueur relationnel). Le contour total est obtenu – comme chez Aubergé – par superposition/addition de tous les contours élémentaires.

Notons bien que la souplesse obtenue par la génération dynamique des contours est réglée par une plus grande flexibilité de la relation entre les représentations linguistiques et prosodiques : indépendamment du niveau phrase, les contours constituent le résultat de

l'interaction de toutes les relations de dépendance entre unités, ceci de manière très récursive (voir les travaux de Holm sur l'énonciation de formules mathématiques *(Holm, B., Bailly, G. & Laborde, C., 1999)*. Le travail sur les empans et la forme des contours a été continué par Brichet et Aubergé *(2002)* pour l'analyse de l'accent d'emphase portant sur le mot ou la syllabe.

### 4.4.3 Contours multiparamétriques

Barbosa et Bailly *(Barbosa, P. & Bailly, G., 1994)* ont démontré l'existence de contours rythmiques. En appliquant un modèle d'élasticité phonétique à la syllabe et en considérant la pause comme une propriété émergente de coefficients d'élasticités limités *(Barbosa, P. & Bailly, G., 1997)*, ils ont proposé de considérer une caractérisation mélodique et rythmique de la prosodie d'un énoncé comme un ensemble de valeurs associées à chaque syllabe (cf. figure ci-dessous).

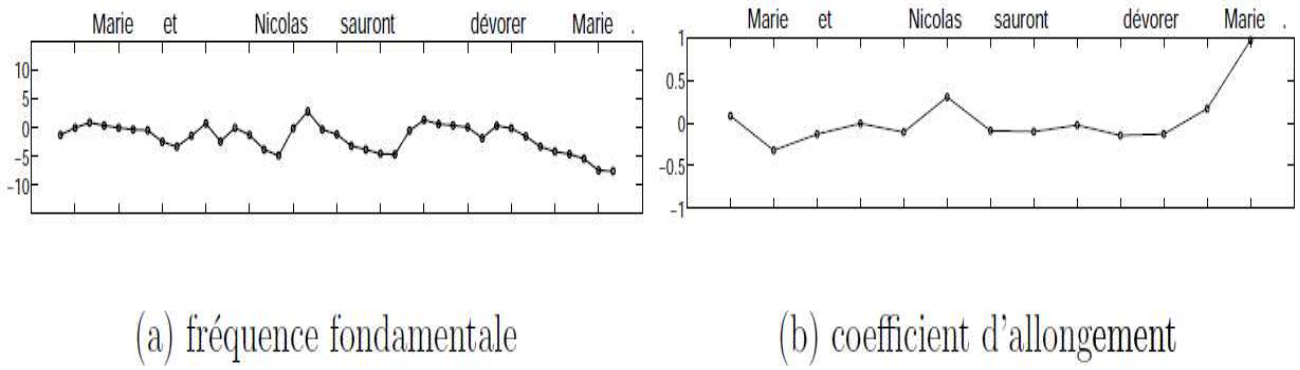



**Figure 18. Contour multiparamétrique. L'ensemble des deux courbes représente le contour multiparamétrique de la phrase. Les traits sur l'abscisse indiquent les GIPCs19. Les mots de la transcription orthographique de la phrase sont alignés avec la fin de leur dernier GIPC *(Bailly, G. & Holm, B., 2003)*.**

### 4.4.4 Modèle SFC

D'après Holm (*2003)*, la prosodie d'un énoncé est le reflet plus ou moins immédiat de fonctions communicatives que le locuteur utilise dans le but de transmettre son message. Ces fonctions sont en nombre assez limité, mais une même fonction est éventuellement utilisée à plusieurs reprises au sein d'un même énoncé : on parle ainsi d'instances de la fonction. Une instance de fonction est caractéristique de la fonction à laquelle elle est associée et est paramétrée par sa portée, *i.e.* la partie de l'énoncé à laquelle elle se réfère ou les parties qu'elle met en relation.

[19] Group Inter Perception Center

En effet, la manifestation prosodique d'une instance de fonction est un contour multiparamétrique coextensif à sa portée. Ces contours associés à l'action d'une fonction sont nommés *contours fonctionnels* et la suite de syllabes est nommée *l'empan du contour*. Ce dernier correspond à la portée d'une instance de fonction.

Les contours fonctionnels associés à une même fonction constituent une famille de contours qui est morphologiquement harmonieuse. Le SFC associe à chaque fonction un générateur de contour qui a la charge de générer l'ensemble des contours associés à cette fonction, en prenant en considération les portées des instances impliquées.

Les instances de fonction se combinent par superposition additive des contours fonctionnels pour engendrer le contour total. Ainsi, le SFC met en évidence la liaison immédiate entre intentions communicatives et implémentation prosodique (cf. ci-dessous) *(Bailly, G. & Holm, B., 2003)*.

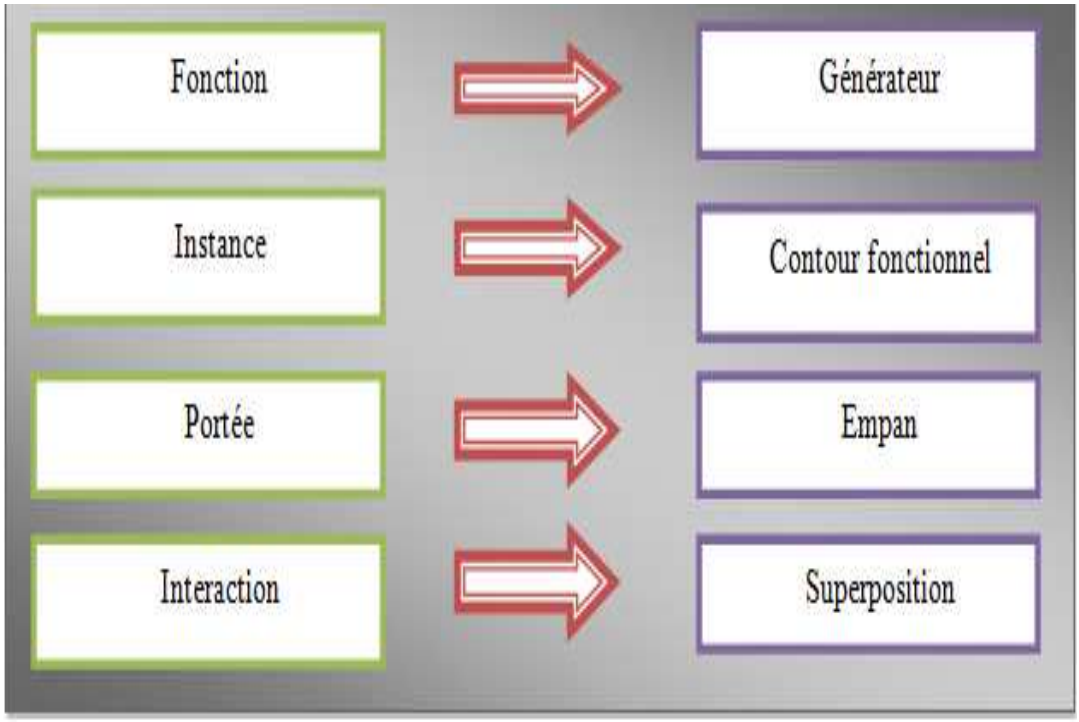



**Figure 19. Hypothèses du SFC**

#### 4.4.4.1 Fonctions et leurs portées

Holm *(2003)* a étudié quatre fonctions principales :

**Fonctions attitudinales :** La portée de ces fonctions est la phrase complète. Les attitudes étudiées sont : déclaration (DC), question (QS), doute-incrédulité (DI), évidence (EV), exclamation (EX) et ironie de soupçon (SC) *(Bailly, G. & Morlec, Y., 1997)*.

**Fonctions hiérarchisantes des constituants syntaxiques : *(cf. annexe B).*** Il distingue quatre fonctions exprimant la relation de dépendance entre deux parties d'un énoncé : la dépendance gauche (DG), la dépendance droite (DD), indépendance (ID), interdépendance (IT).

**Fonctions de segmentation :** La morphologie caractéristique des contours prosodiques aux alentours des articles et pronoms ce qui explique l'existence d'une fonction de segmentation, appelée X, dont la portée serait uniquement l'article (ou pronom) et le mot suivant : (le) X (livre). Elle servirait à mettre en relief ce dernier par rapport au premier.

**Fonctions d'emphase et de focalisation (EM) :** la portée d'une telle fonction serait simplement la syllabe, le mot ou le groupe de mots mis en valeur *(Brichet, C. & Aubergé, V., 2002)*. Comme pour la fonction de segmentation, la portée pourrait inclure les parties d'énoncé par rapport auxquelles la mise en relief s'effectue *(Bailly, G. & Holm, B., 2003)*.

#### 4.4.4.2 Générateurs et familles de contours

Le SFC associe à toute instance d'une fonction communicative un contour multiparamétrique dont l'empan correspond à la portée de la fonction. Tous ces contours fonctionnels qui appartiennent à une même fonction constituent une famille homogène. A l'intérieur de ces familles, la morphologie des contours est exclusivement déterminée par des contraintes phonotactiques (nombre de syllabes de l'empan, indexation relative et absolue de chaque syllabe par rapport à la fin et au début de l'empan). Les dispositifs capables de générer ces familles sont nommés générateurs. Il y en a un par fonction communicative et ils doivent exécuter la mise en correspondance (« mapping » en anglais) entre configurations phonotactiques et contours prosodiques. Ces générateurs sont implémentés par des réseaux de neurones *(Bailly, G. & Holm, B., 2003)*.

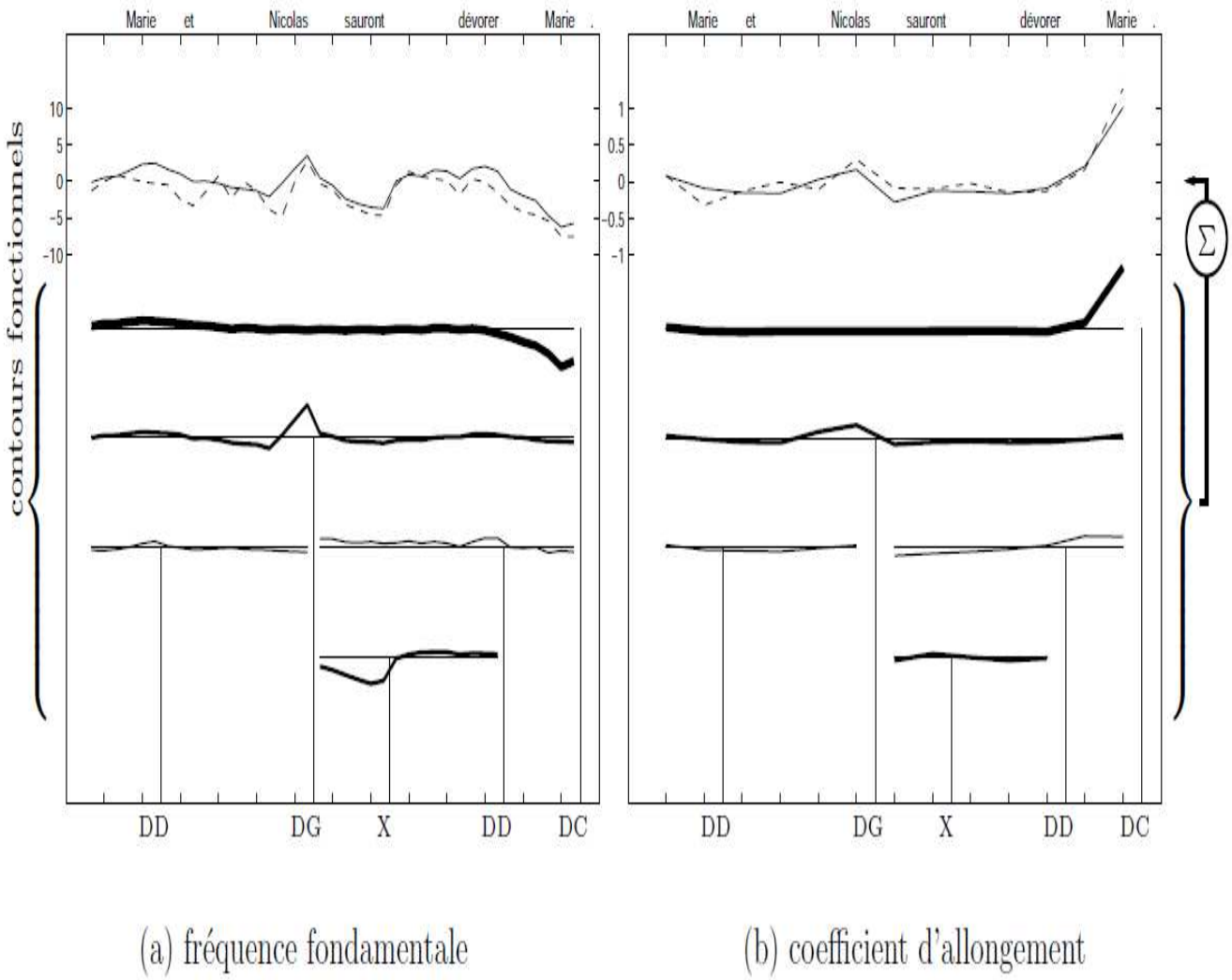


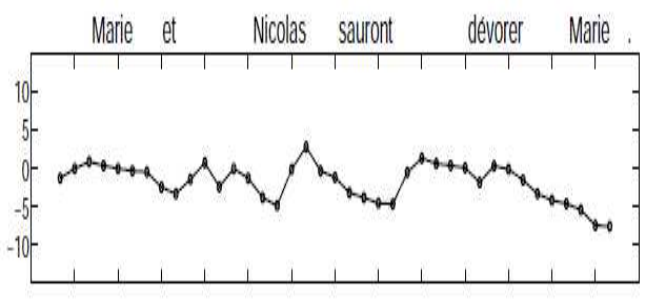

**Figure 20. « Exemple de superposition. Les résultats des prédictions du SFC seront, par la suite, souvent représentés sous cette forme. On y trouve tout en haut la transcription orthographique (les mots étant alignés avec la fin de leur dernier GIPC). La courbe en tiret représente le contour global observé ; celle en trait plein le contour prédit par superposition des contours fonctionnels qui, à leur tour, se trouve dans la partie inférieure des figures. Les étiquettes des fonctions associées aux contours fonctionnels sont notées tout en bas à l'endroit du point d'ancrage[20] central du contour. »** ***(Bailly, G. & Holm, B., 2003)***

#### 4.4.4.3 Superposition

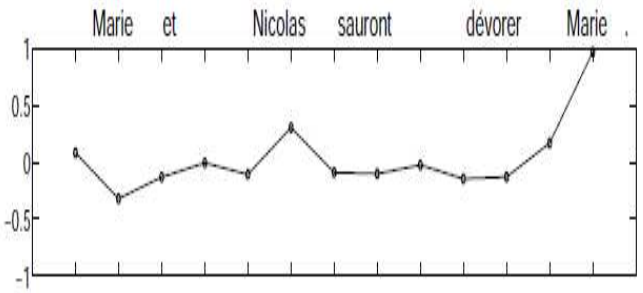
Le contour prosodique total prédit est obtenu par la superposition additive des contours fonctionnels. Il s'agit donc d'une simple interaction des manifestations prosodiques des différentes fonctions de haut niveau. Ces fonctions peuvent être utilisées à des divers niveaux hiérarchiques et plusieurs fois sur un même niveau. Les portées des instances de fonctions et avec eux les empans des contours fonctionnels peuvent se chevaucher et s'imbriquer.

[20] Ils sont le début et la fin de l'empan et, le cas échéant, la frontière entre groupes.

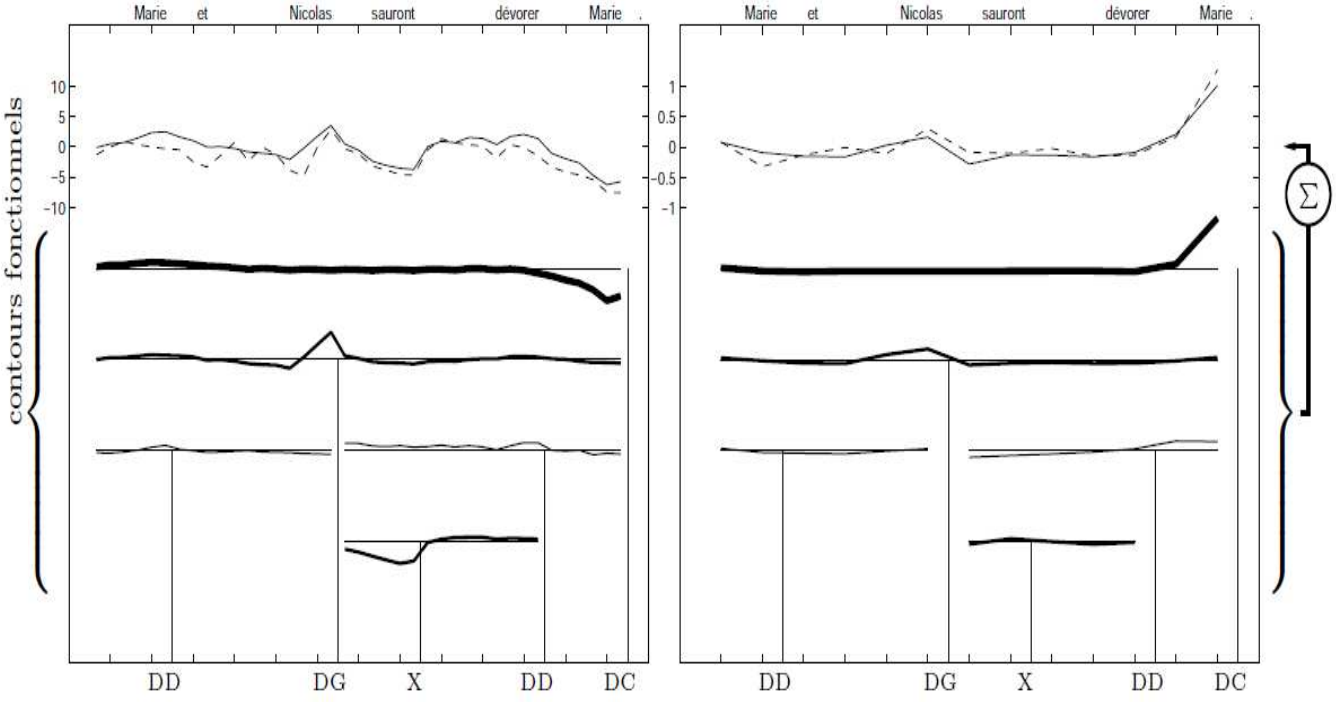
La figure ci-dessus présente un exemple de génération du contour global par superposition de contours fonctionnels. Le contour prosodique observé est figuré en haut des graphiques en tirets. Sur les quatre niveaux de la partie inférieure sont figurés les contours fonctionnels avec leur point d'ancrage et le nom de la fonction associée. Ils sont engendrés par quatre générateurs différents (un par fonction). Les deux contours DD sont issus du même générateur, ils font partie de la même famille. En superposant additivement les contours fonctionnels on acquière le contour global prédit (traits pleins, en haut des figures).




(a) fréquence fondamentale




(b) coefficient d'allongement




(a) fréquence fondamentale (b) coefficient d'allongement

En guise de conclusion, il existe plusieurs types de générateurs de contours : ceux qui sont responsables de l'encodage de la modalité (dans notre exemple, la modalité est affirmative DC), ceux qui sont responsables de la hiérarchisation des constituants syntaxiques (par exemple, DG marque la relation groupe nominal sujet/groupe verbal).

#### 4.4.4.4 L'inversion du modèle : l'apprentissage

Traber [1992] a indiqué que la prédiction directe de contours complexes à partir d'indices phonologiques et phonotactiques simples est possible avec un réseau de neurones. Mais l'inversion d'un modèle superpositionnel, *i.e.* l'apprentissage de contours fonctionnels à partir de l'observation d'un corpus de contours supposés être le résultat de multiples superpositions, est un problème mal-posé : en effet, la décomposition n'est pas unique (10=5+5=5+3+2=3+1+1+3+2…). Il faut donc ajouter des contraintes supplémentaires. Par exemple, une possibilité est d'appliquer une forme spéciale – aux contours élémentaires – comme aux modèles du type Fujisaki. Une autre possibilité est de définir une procédure d'inversion. Morlec *(1997)* adopte cette voie en suggérant un apprentissage hiérarchique partant des contours porteurs et additionnant alternativement les contours portés.

En guise de conclusion, le SFC s'affranchit de la notion de hiérarchie des contours : les contours qui ont le même type peuvent être porteurs ou portés. Ainsi, chaque contour est associé à une fonction communicative et les contours associés à une même fonction sont homogènes. En appliquant la méthode de génération, chaque fonction se voit affecter un générateur responsable de la création de tous les contours associés. L'entrée des générateurs est exclusivement de nature phonotactique et représente l'empan de manière assez simple.

#### 4.4.4.5 Analyse par la synthèse

La figure suivante schématise le cycle d'analyse par la synthèse. En effet, pour paramétrer les générateurs de contours, il est nécessaire de boucler trois processus :

1. **Synthèse :** On prend l'ensemble des "fonctions métalinguistiques", on sélectionne les générateurs de contours correspondants et on superpose et additionne les contours générés. A l'initialisation, cette opération génère des contours nuls.
2. **Analyse :** On calcule l'erreur entre contours prédits et générés, puis on répartit cette erreur entre les contours élémentaires qui contribuent à chaque syllabe. Ces erreurs élémentaires s'ajoutent à ce que les contours savent déjà prédire et deviennent les nouveaux contours élémentaires d'apprentissage des générateurs (on appelle ces ensembles des familles de contours : ils sont supposés avoir été engendrés par le même générateur fonctionnel).

3. **Apprentissage :** L'ensemble des générateurs sont alors entraînés à générer au mieux les familles de contours.

On boucle ces trois phases jusqu'à ce que les générateurs ne changent plus. On peut alors utiliser le processus de synthèse pour générer les contours de nouvelles phrases, qui est le même que celui utilisé pour paramétrer les générateurs de contours : c'est pourquoi le processus d'apprentissage est appelé analyse par la synthèse. Les fonctions métalinguistiques (les marqueurs et leurs empans) sont supposés connus et on apprend la manière dont les générateurs de contours synthétisent les contours élémentaires qui se superposeront et s'additionneront pour prédire au mieux les contours globaux observés. On procède donc d'une analyse descendante, du contenu vers la substance (des marqueurs vers la prosodie) et pas l'inverse *(Bailly, G. & Holm, B., 2003)*.

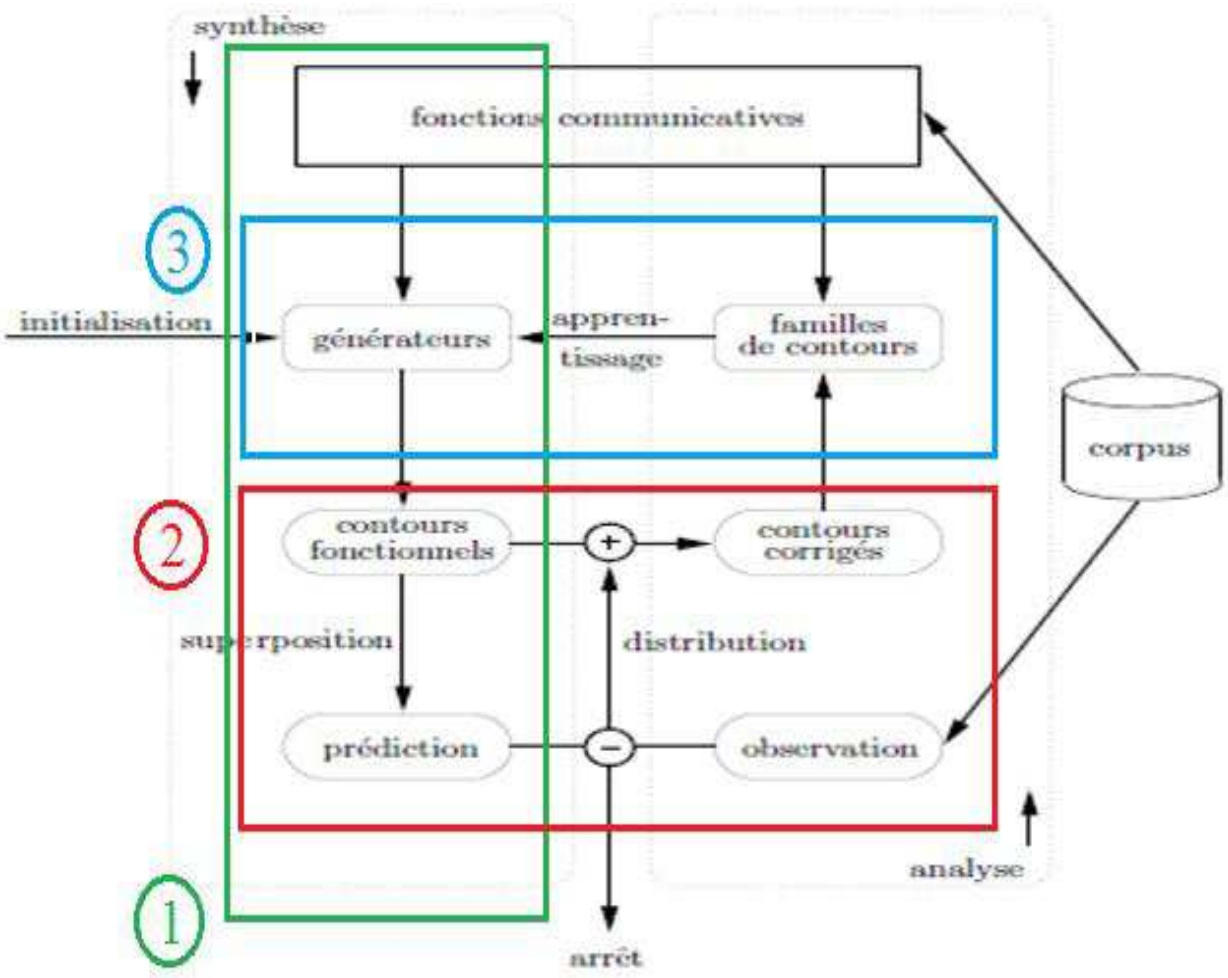



**Figure 21. Cycle d'analyse par la synthèse.**

## 4.5 Commentaires

Les modèles SFC et HTS procèdent tous deux de l'analyse descendante. En effet, HTS est basée sur une grande base de données, voire plusieurs heures d'enregistrements. Il effectue une régression directe entre un jeu de descripteurs qui permet de caractériser symboliquement la fonction de chaque phonème et des caractéristiques phonétiques et prosodiques du signal de parole. Ainsi, le HTS applique un apprentissage massif de mise en correspondance entre consignes et substance, du haut vers le bas. Cependant, ce système reste une boite noire (en anglais « blackbox») vu que l'on n'a pas d'étapes intermédiaires et nécessite une couverture combinatoire de tous les descripteurs.

En mettant en scène peu de fonctions à de multiples niveaux et en imposant un paramétrage réduit, le SFC considère que les relations entre les unités sont encodées par des contours globaux multiparamétrés dont la forme ne dépend que de la fonction et de la taille des unités connectées et ceci indépendamment de la position de ces unités dans le discours. Conséquemment, à l'opposé du HTS, le système SFC reste une boite blanche (en anglais « white box»), ou une boite transparente. SFC est plus adapté à des corpus de taille réduite, ce qui est notre cas (cf. chapitre suivant).

## 4.6 Conclusions

Dans ce chapitre, nous avons présenté brièvement l'analyse ascendante et l'analyse descendante. Ces deux approches visent à mettre en correspondance fonctions linguistiques et substance prosodique. Jusqu'à présent, l'analyse ascendante a fait l'objet d'une attention particulière en phonologie prosodique. De manière assez remarquable, les méthodes actuelles utilisées en synthèse de la parole opèrent par analyse descendante, de l'étiquetage massif de données à la paramétrisation du signal, souvent en s'affranchissant de tout modèle prosodique intermédiaire.

Nous avons introduit le modèle prosodique SFC qui propose un encodage direct de fonctions linguistiques ou paralinguistiques par des contours multiparamétriques chevauchants. L'objectif de ce travail est de confronter le SFC à la synthèse de dictées, où plusieurs répétitions de la même phrase sont effectuées avec des styles d'énonciation différentes.

# PARTIE 2 : ÉTUDE EMPIRIQUE

*« La théorie, c'est quand on sait tout et que rien ne fonctionne. La pratique, c'est quand tout fonctionne et que personne ne sait pourquoi. Ici, nous avons réuni théorie et pratique : Rien ne fonctionne . . . et personne ne sait pourquoi ! »*
*[Albert Einstein]*

# 5 CORPUS & MÉTHODOLOGIE

## 5.1 Introduction

La base essentielle de tout système de synthèse est le corpus. Ce chapitre a pour but de présenter la méthodologie utilisée pour recueillir des données, analyser et caractériser le corpus.

## 5.2 Corpus

Dans le but d'étudier les stratégies d'énonciation et l'intonation dans le cadre de la lecture de dictée, un ensemble de quinze dictées lues, à voix haute ont été réalisées et enregistrées par une unique lectrice experte qui s'appelle Aude. Ces dictées de niveau CM1-CM2 sont issues du site français ladictéee.fr (voir www.ladictee.fr ) qui offre un ensemble diversifié de dictées et d'exercices grammaticaux pour les élèves. Ainsi, nos dictées sont simples et de durée comprise entre trois et cinq minutes. Notre première tâche était de faire une transcription écrite de la parole ***(cf. annexe C)***.

| Dictées | Titre | Durée | Nombre de mots |
|---|---|---|---|
| AUDE0000 | Avant l'orage. | 6 minutes et 47 secondes | 84 |
| AUDE0001 | Dans la jungle. | 7 minutes et 17 secondes | 119 |
| AUDE0002 | Drame en montagne. | 4 minutes et 56 secondes | 60 |
| AUDE0003 | Incendie dans les Landes. | 4 minutes et 9 secondes | 53 |
| AUDE0004 | La biche. | 6 minutes et 47 secondes | 85 |
| AUDE0005 | La chasse du renard. | 4 minutes et 35 secondes | 56 |
| AUDE0006 | La petite fille et les oiseaux. | 5 minutes et 32 secondes | 89 |
| AUDE0007 | La rivière. | 5 minutes et 45 secondes | 83 |
| AUDE0008 | Le loup. | 8 minutes et 17 secondes | 111 |
| AUDE0009 | Le repas d'un paysan. | 6 minutes et 18 secondes | 95 |
| AUDE0010 | Le vieux noyer et les oiseaux. | 4 minutes et 13 secondes | 65 |
| AUDE0011 | Pauvre minet. | 3 minutes et 30 secondes | 48 |
| AUDE0012 | Un coiffeur novice. | 7 minutes et 23 secondes | 105 |
| AUDE0013 | La petite école de campagne. | 7 minutes et 18 secondes | 98 |
| AUDE0014 | La leçon d'histoire. | 6 minutes et 10 secondes | 88 |
| Total | | 88 minutes et 57 secondes | 1239 |

**Tableau 1. Description du corpus par durée et nombre de mots.**

Bien que le nombre de dictées étudiées soit limité, il est nécessaire de signaler que les variations portent plus sur la manière de répéter les groupes et d'en introduire de nouveaux, que sur la formation des groupes à proprement parler. Ce point est fondamental, dans la mesure où c'est notamment pour cette étude, la formation des groupes qui nous importe.

### 5.3 Alignement et étiquetages des signaux

L'étude des stratégies d'encodage d'informations linguistiques par la prosodie nécessite d'avoir accès à des corpus alignés phonétiquement et étiquetés avec les informations phonotactiques, syllabiques, lexicales, syntaxiques, etc. Des outils d'annotation tels que Praat[21] *(Boersma P. et al., 2009)* permettent de définir et d'enrichir différentes couches

[21] Praat est utilisé dans le but d'analyser la phonétique et la parole et manipuler des signaux de discours acoustiques. Cet outil créé par Paul Boersma et David Weenik de l'Université d'Amsterdam est utile aux phonéticiens pour la conduite d'analyse spectrale et d'exploration de la tonalité, les formants et l'intensité de la parole.

d'annotation. Ils offrent aussi la possibilité d'aligner manuellement ces différentes annotations avec le son. En revanche, l'alignement manuel montre deux problèmes cruciaux. Primo, un tel traitement nécessite un temps important : de 130 à 800 fois la durée du son *(Schiel, F. et al., 2003 ; Kawai, H. et al., 2004)*. Secundo, ce travail requiert des phonéticiens experts et une excellente entente entre les différents annotateurs. Dans ce but, des outils d'alignement automatique sont fréquemment utilisés pour produire un alignement à la fois reproductible et cohérent. Les méthodes procédant par alignement du signal avec des modèles de Markov cachés (HMM) de phones en contexte produisent les meilleurs résultats d'alignement phonétique *(Adell, J. et al., 2005; Van Niekerk, D. et al., 2009)*.

Nos dictées ont été phonétisées et alignées automatiquement par un outil existant du GIPSA-Lab. Outre les positions des phonèmes constitutifs de chaque mot et ces pauses éventuelles, cet outil délivre :

1. Le graphème correspondant à chaque phonème : Le phonétiseur procède par arbre de décision entrainé sur un corpus de 150000 mots (dont le Robert) alignés *(Black, A., Lenzo, K. & Pagel, V., 1998)*.
2. La segmentation en syllabes : règles de segmentation dans COMPOST (un serveur de synthèse multilingue) *(Bailly, G. & Alissali, M., 1992 ; Alissali, M. & Bailly, G., 1991)*.
3. La classe lexicale de chaque mot (parmi 23) : dictionnaire morphologique de 150000 entrées. Puis, filtrage par tri-grams *(Bailly, G. & Alissali, M., 1992 ; Alissali, M. & Bailly, G., 1991)*.
4. La modalité de la phrase : trivial. Distingue questions pleines des questions introduites (forme "est-ce que", pronom ou adjectif interrogatif).

Il peut aussi délivrer un marquage des dépendances syntaxiques, mais cette fonction automatique n'a pas été utilisée car nous avons préféré utiliser des outils de TAL plus robustes.

Ces alignements et informations ont ensuite été édités manuellement sous Praat.

Dans nos dictées, la lectrice effectue quatre passes sur le texte ***(cf. annexe D)*** :

a. **une lecture de compréhension**, où le texte est lu une première fois de manière intégrale suivie de deux répétitions de dictée, où les marques de ponctuation sont prononcées aux positions auxquelles elles apparaissent dans le texte afin que le scripteur puisse les écrire. Ainsi, nous utilisons le terme « tronçon de phrase » pour faire référence aux versions (b) et (c).

b. **une dictée de groupe de mots.**

c. **une dictée de mots,** où pratiquement tous les mots non grammaticaux sont prononcés de manière isolée.

d. **une version dite de vérification**, permet au scripteur de vérifier l'orthographe de chaque phrase.

| Version | ***Énonciation*** |
|---|---|
| | *Je m'appelle Aude… Je commence la dictée :* |
| a | *Avant l'orage.*<br>*Des nuages violets passaient sur nos têtes, et la lumière bleuâtre baissait de minute en minute, comme celle d'une lampe qui meurt. Je n'avais pas peur, mais je sentais une inquiétude étrange, une angoisse profonde, animale. Les parfums de la colline étaient devenus des odeurs et montaient du sol, presque visibles.*<br>*Plusieurs lapins passèrent, aussi pressés que devant les chiens, puis des perdrix surgirent sans bruit du vallon, et se posèrent à trente pas sur notre gauche ...*<br>*Texte de Marcel Pagnol.* |
| b | *Avant l'orage PT* |
| c | *Avant l'orage PT* |
| b | *Des nuages violets passaient sur nos têtes VRG* |
| c | *Des nuages violets passaient sur nos têtes VRG* |
| b | *Et la lumière bleuâtre baissait de minute en minute VRG* |
| c | *Et la lumière bleuâtre baissait de minute en minute VRG* |
| b | *Comme celle d'une lampe qui meurt PT* |
| c | *Comme celle d'une lampe qui meurt PT* |
| | *Je relis la phrase en entier.* |
| d | *Des nuages violets passaient sur nos tête VRG et la lumière bleuâtre baissait de minute en minute VRG comme celle d'une lampe qui meurt PT* |
| … | … |

**Tableau 2. Les quatre passes sur le texte.**

### 5.3.1 Remarque sur la ponctuation

Remarquons bien que, dans la version (a), les marques de ponctuation ne sont pas prononcées verbalement, mais la lectrice y associe souvent des pauses. En revanche, dans chacune des versions (b), (c) et (d), les marques de ponctuation sont prononcées aux positions auxquelles elles apparaissent dans le texte pour que le scripteur puisse les reproduire. Notons aussi que certains mots non usuels, notamment les noms propres, sont souvent prononcés puis épelés dans les versions (c) et (d).

## 5.4 Analyse des corpus

Nous rappelons que notre hypothèse forte est que l'énonciation à voix haute consiste à faire écouter des fragments syntaxiques prélevés sur un arbre syntaxique complet. Ainsi, nous faisons l'hypothèse que des diverses versions d'énonciation des groupes de mot ne modifient que la forme des contours mélodiques et rythmiques et non pas leur fonction. Nous avons pour ceci analysé de manière fine le corpus de 15 dictées orales. Nous avons effectué ainsi :

1. Manuellement, le modèle de marquage de Bailly ***(cf. annexe B),*** dans chaque version (a). En effet, un ensemble restreint de marques prosodiques peuvent lier les constituants (mots, groupes, phrases, clauses) : interdépendance **IT** reliant deux unités adjacentes dirigé par le même gouverneur et l'indépendance, dépendance gauche **DG** reliant la tête d'un sous-arbre (le " gouverneur " ou " mère ") avec son unité dépendante immédiatement précédent (" sœur "), dépendance droite **DD** reliant le gouverneur avec son unité à charge suivant immédiatement et indépendance **ID** lorsque aucune des relations simples précédentes peuvent être identifiés.
2. Le marquage manuel a ensuite été transféré automatiquement aux versions (b), (c) et (d) par alignement.

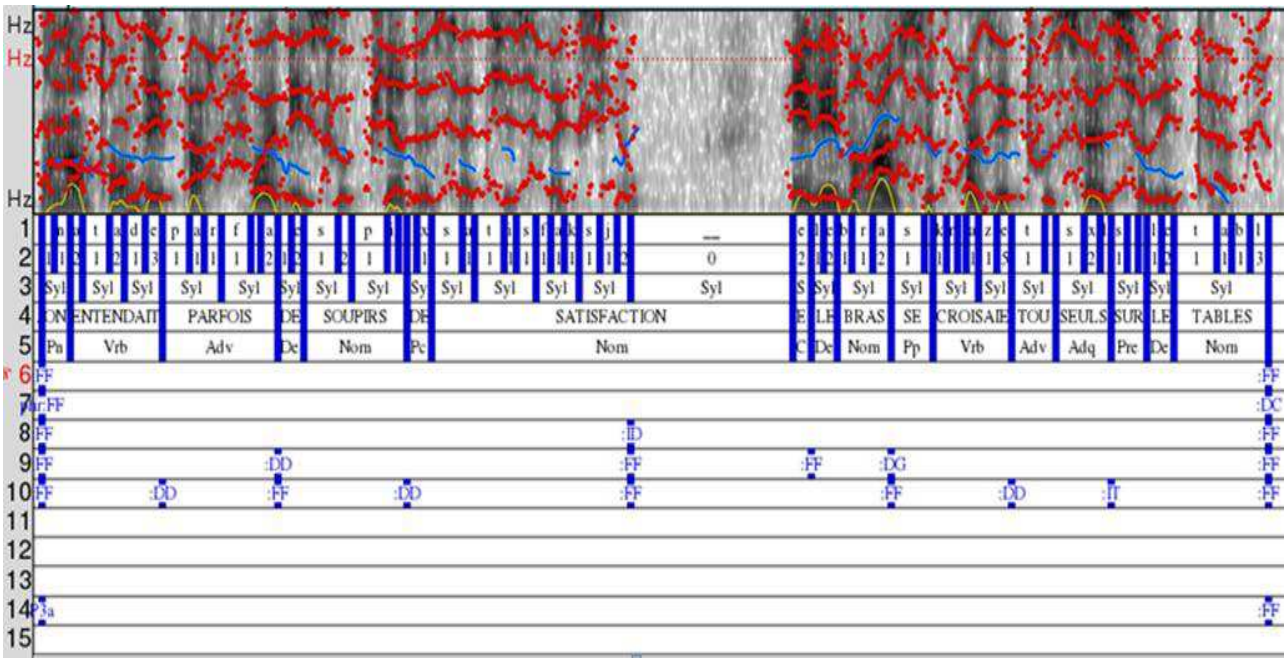



**Figure 22.** **Marquage d'une version (a) d'un échantillon de parole « On entendait parfois des soupirs de satisfaction et les bras se croisaient tout seuls sur les tables. » à l'aide de Praat. Le [:FF] balise délimitant le début et la fin des unités. D'autres tags sont utilisés pour marquer les fonctions métalinguistiques : [:DC] pour la phrase déclarative, [:IT] entre clauses, [:DG] et [:DD] entre les groupes nominaux et verbaux.**

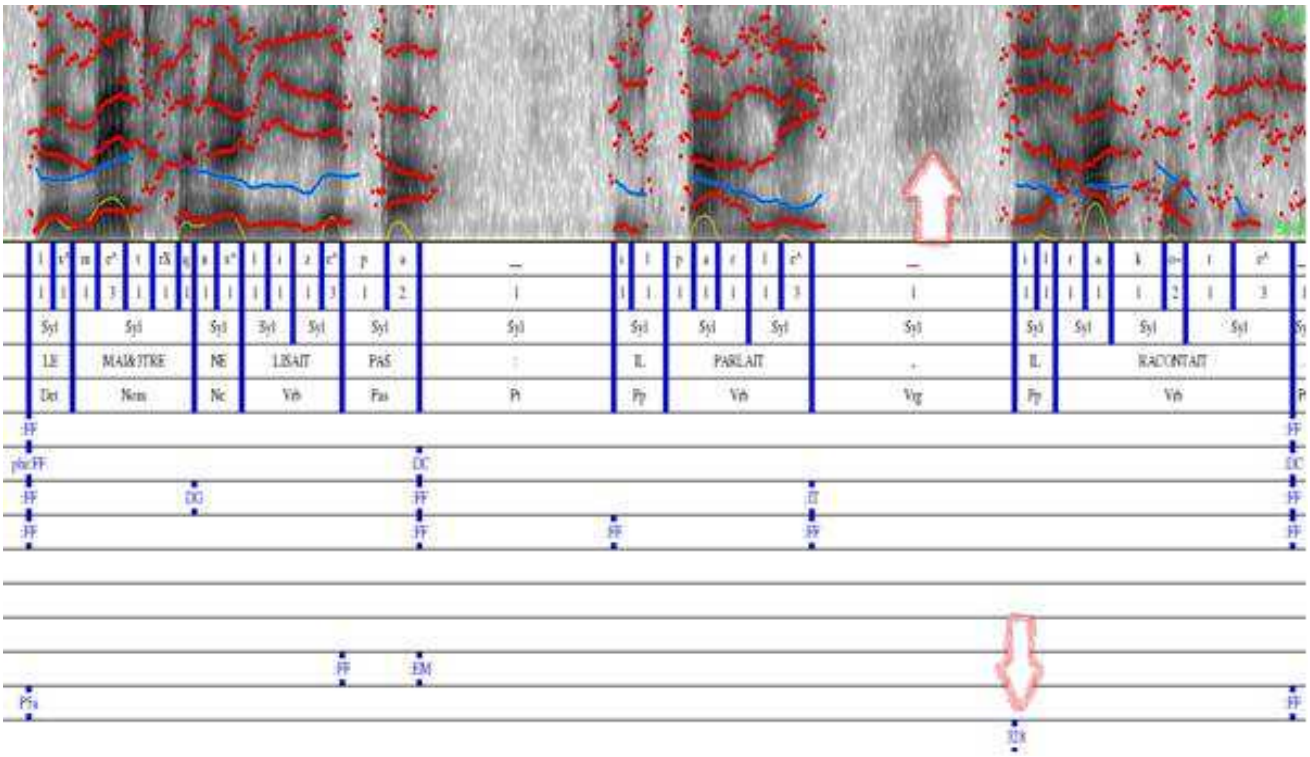



**Figure 23.** **Attention portée aux contours intonatifs, aux signes de ponctuation et aux pauses. Un échantillon de parole « Le maître ne lisait pas : il parlait, il racontait.».**

Pour chaque texte, dans chaque version (a), nous nous focalisons sur deux types d'informations. Primo, les découpages prosodiques, *i.e.* la manière dont les textes ont été segmentés en groupes en appliquant le modèle de marquage. Secundo, la forme du contour intonatif observé à la fin des différents groupes prosodiques. Une attention spéciale a été portée aux contours intonatifs, aux signes de ponctuation et aux pauses qui marquent en partie les découpages. Ainsi, si l'on a, dans la version (a), un groupe de mots marqué par **:ID**, ce dernier deviendra alors un marqueur de découpages dans les deux versions (b) et (c).

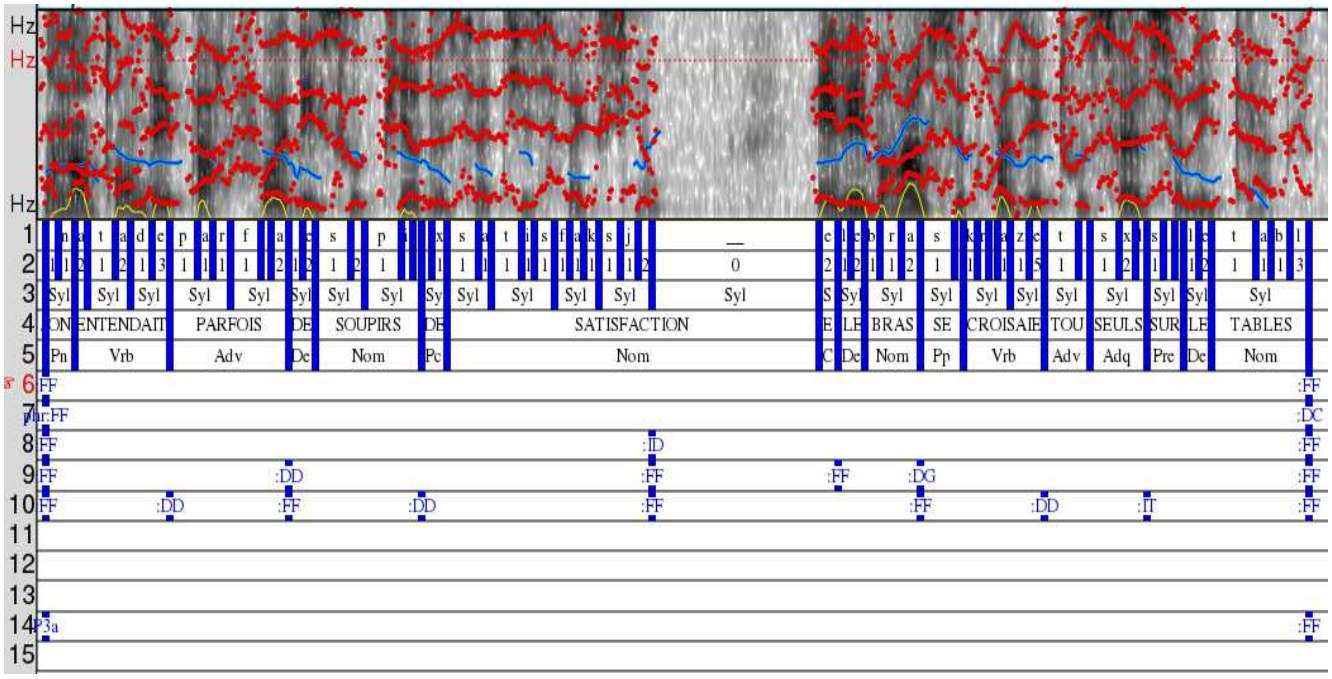



**Figure 24. Utilisation du marqueur ID entre deux propositions connectées par une coordination. Un échantillon de parole « On entendait parfois des soupirs de satisfaction et les bras se croisaient tout seuls sur les tables.»**

Enfin, la forme du contour intonatif observé à la fin des différents groupes prosodiques non terminaux dans chacune des versions (b) et (c) signale qu'il s'agit d'un tronçon de phrase non terminal. Nous avons ainsi utilisé un nouveau marqueur prosodique **:CT**, appelé marqueur de continuation**.** Ce dernier n'était pas utilisé dans les versions (a) et (d) car il est, bien évidemment, spécifique aux stratégies de découpage dans les deux versions (b) et (c) *i.e.* ce que l'on a déjà appelé les « tronçons de phrase ».

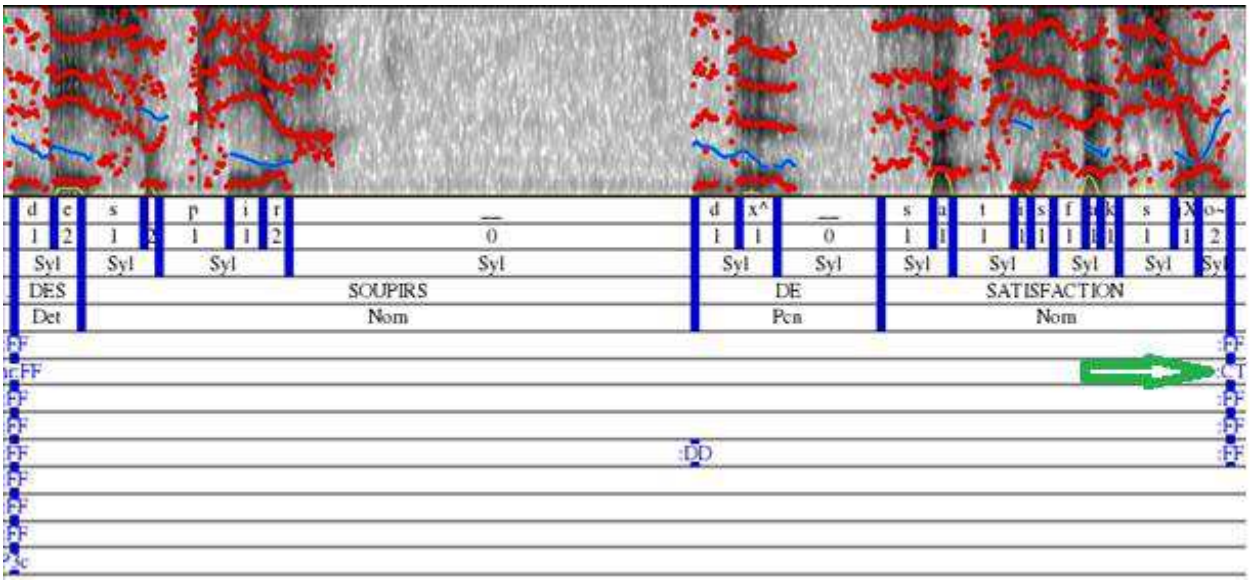



**Figure 25. Le marqueur ID dans la version (a) devient marqueur CT dans la version (c). Un échantillon de parole « Des soupirs de satisfaction.»**

En outre, dans les données observées, nous avons remarqué que ce marqueur continuatif était dupliqué « comme dans un miroir » dans l'énonciation du signe de ponctuation (cf. ci-dessous). Ceci était aussi le cas de l'énonciation des ponctuations finales qui dupliquaient la forme des marqueurs finaux des phrases qu'ils concluaient. Chaque énonciation de signe de ponctuation duplique donc le marqueur précédent.

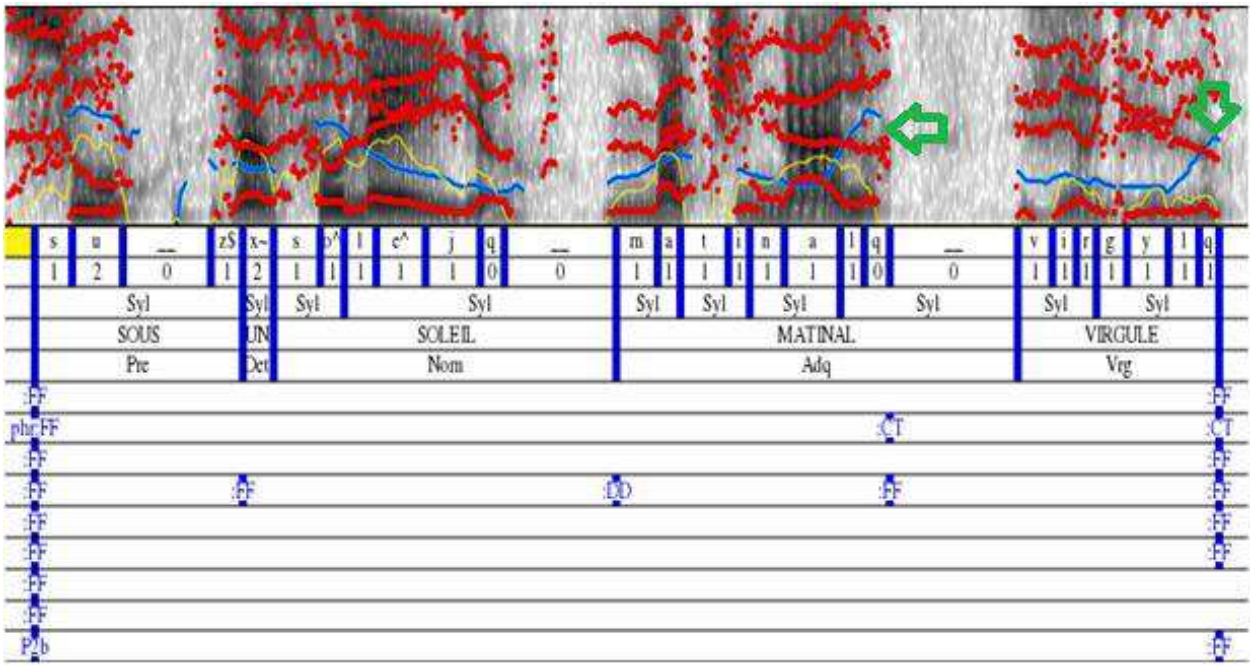



**Figure 26. Forme du contour intonatif à la fin des différents groupes prosodiques dans chacune des versions (b) et (c) : le cas d'un marqueur de continuation.**

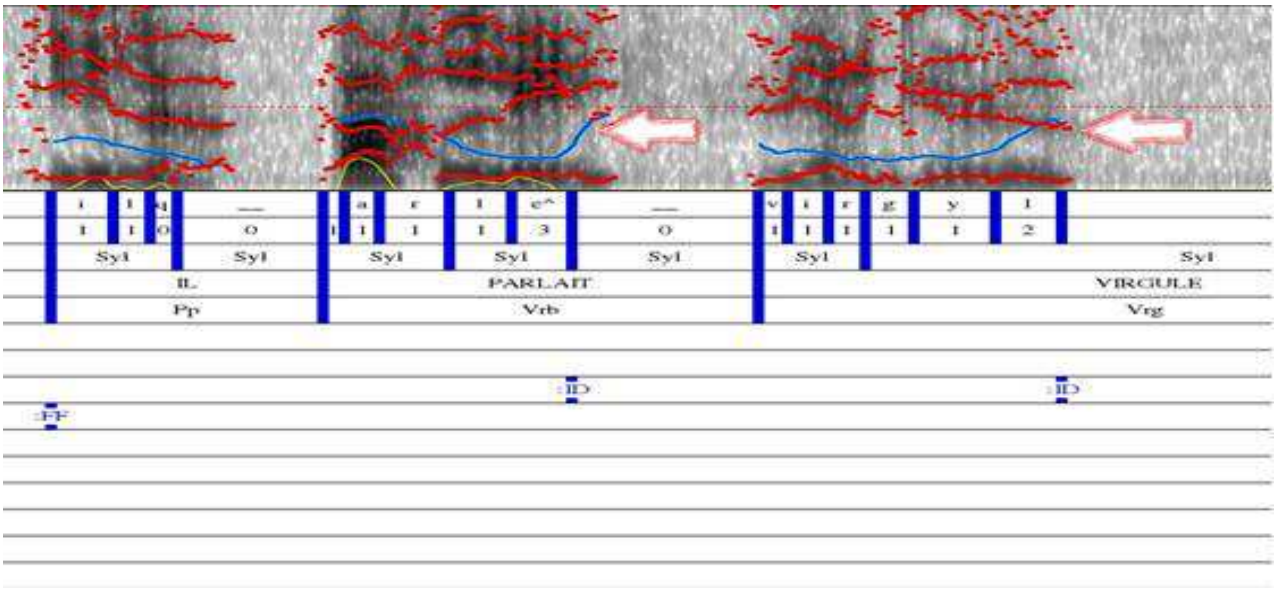



**Figure 27. Forme du contour intonatif à la fin des différents groupes prosodiques dans chacune des versions (b) et (c) : le cas d'un marqueur d'interdépendance.**

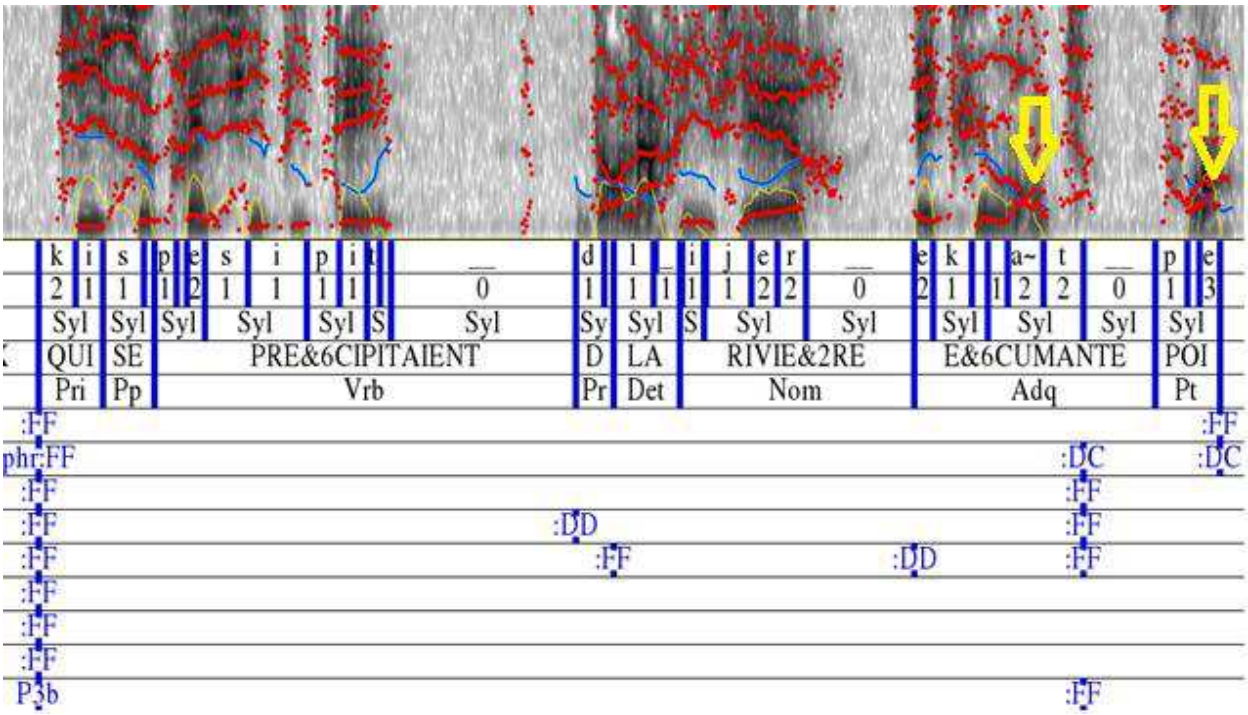



**Figure 28. Forme du contour intonatif à la fin des différents groupes prosodiques dans chacune des versions (b) et (c) : le cas d'une phrase déclarative.**

Notons bien que dans le cas de phrases longues sans signe de ponctuation, nous considérons un nombre maximal de syllabes dans un tronçon de phrase (soit 12 syllabes dans notre cas). Les phrases ont ainsi découpées de manière itérative en considérant une hiérarchie entre marqueurs : **:ID**, **:IT**, **:DG** puis **:DD**.

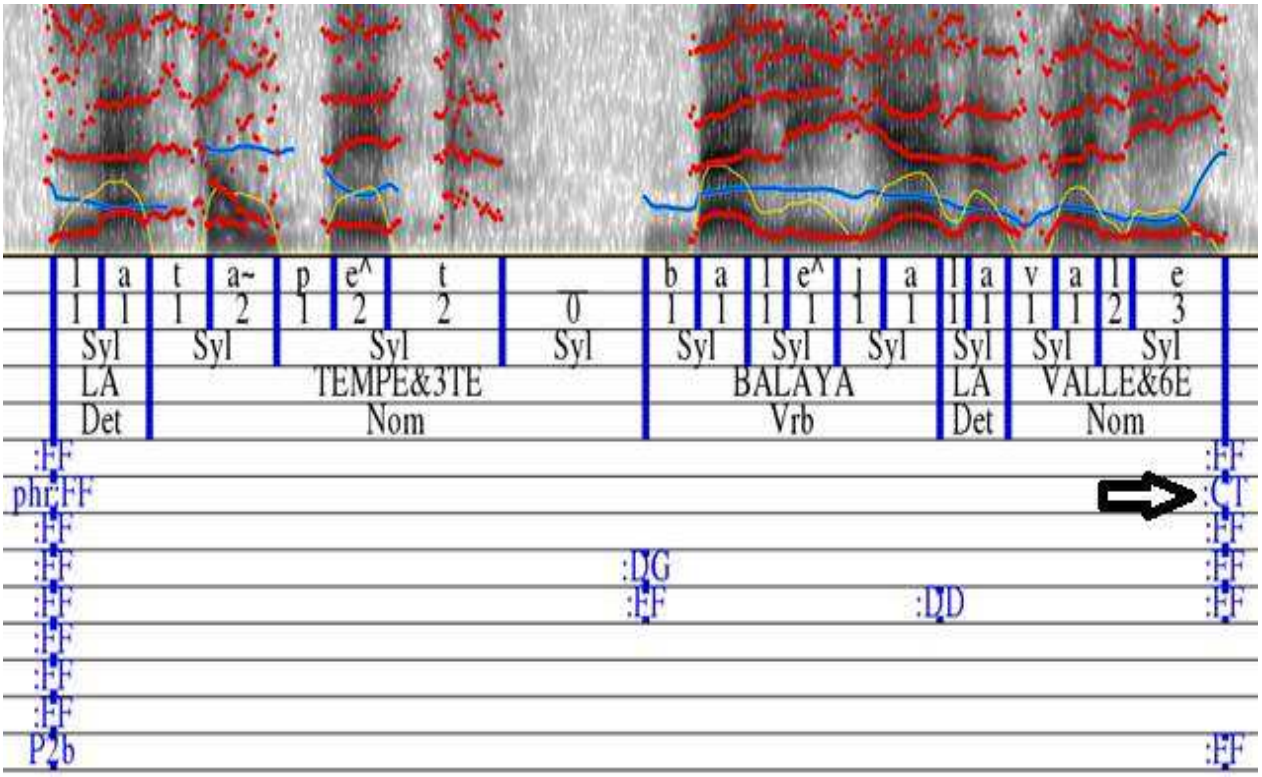



**Figure 29.** **Découpage d'un bout de phrase sans signe de ponctuation avec le marqueur :CT.**

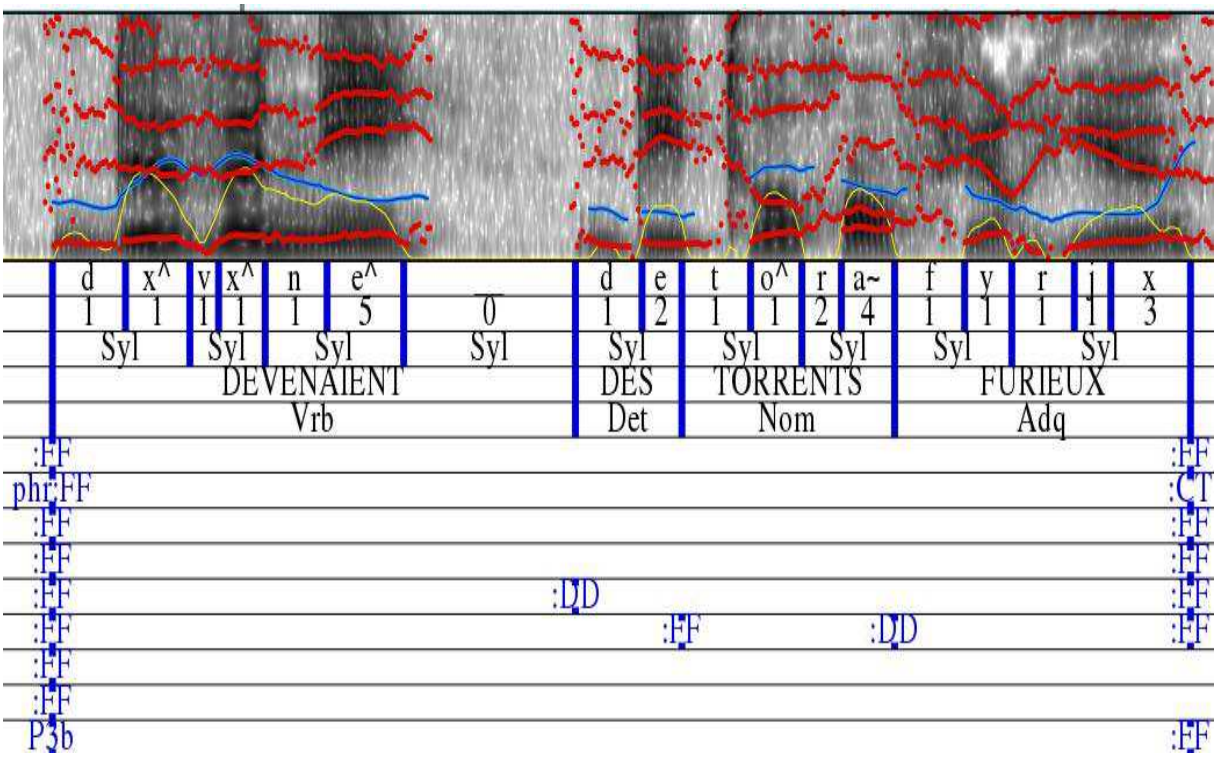



**Figure 30.** **Découpage d'un bout de phrase sans signe de ponctuation avec le marqueur :CT.**

## 5.5 Analyse morphosyntaxique

Afin de pouvoir appliquer le système de marquage, il est indispensable d'avoir une analyse morphosyntaxique de nos dictées. Au début, nous avons essayé d'utiliser Cordial de Synapse. Développé par la société Synapse Développement, il s'agit d'un « logiciel d'analyse syntaxique et sémantique », jugé comme l'un des étiqueteurs morphosyntaxiques les plus fiables. Il comporte neuf catégories (nom, verbe, adjectif, pronom, déterminant, adverbe, préposition, conjonction et ponctuation). Celles-ci sont divisées en sous-catégories (nom : commun / propre, pronom : personnel / démonstratif / indéfini / relatif / possessif, etc.) et précisent certains traits morphologiques (mode, temps, personne, genre, nombre) et phonologiques (débute par un h aspiré), ce qui donne un jeu de 128 étiquettes. Cordial est orienté vers la langue écrite : son corpus de référence, notamment, est un corpus de presse et de littérature. Son analyse se base sur la « phrase », notion déjà problématique à l'écrit, et totalement inutilisable pour l'oral *(Blanche-Benveniste, C., 1998)*. Néanmoins, la catégorie verbale étant souvent repérable immédiatement, sans ambiguïté, de par sa forme, nous avons estimé que l'étiquetage devait convenir pour la requête spécifique de dépendances à distance en infinitive. » *(Berard, L., 2012)*.

L'analyse délivrée par Cordial n'était pas convenable à nos attentes. En effet, trois problèmes majeurs sont apparus en lien avec le format du fichier de sortie, la négation et l'épellation.

### 5.5.1 Format du fichier de sortie

On s'est trouvé devant un problème technique qui se manifeste dans le format de sortie de Cordial : il s'agit d'un fichier d'extension .COR qui perd son architecture d'arbre lors de son ouverture avec n'importe quel autre éditeur du traitement de texte autre que Cordial.

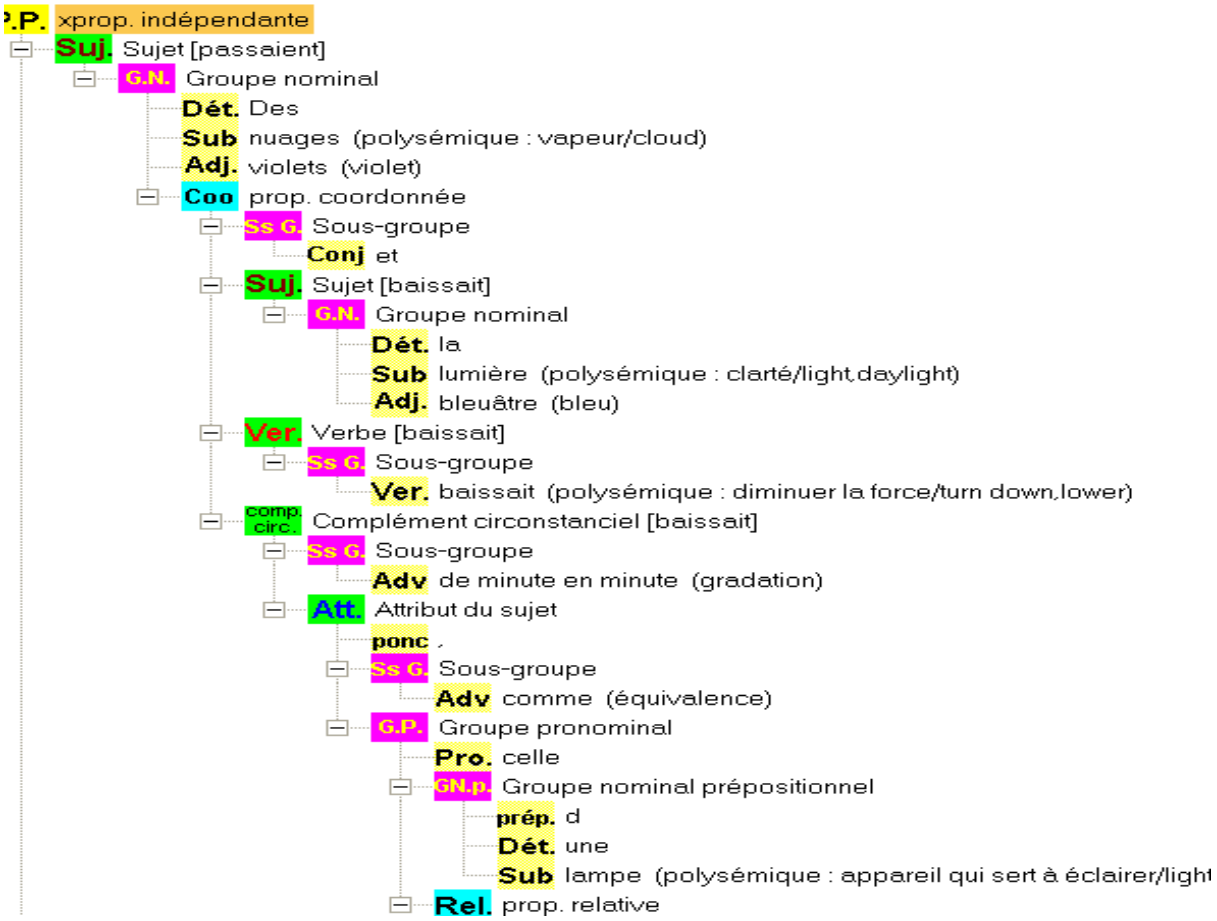



**Figure 31. Sortie de Cordial avec l'extension .COR**

### 5.5.2 La négation

La sortie de l'analyse morphosyntaxique traitant la négation ne répondait pas à nos besoins. En prenant par exemple le bout de phrase « je n'avais pas peur », nous constatons, lors de l'analyse morphosyntaxique, que Cordial considère que « n'avais pas » est un seul syntagme. Par contre, selon notre système de marquage, la négation s'applique plutôt au complément avec un marqueur **:DD** après chaque verbe dans la phrase, comme cela est illustré ici :

```
Gv (
        Pp(JE)
        Ne(N\\\')
        Vrb(AVAIS)
        <Gn;#DD>(
                Pas(PAS)
                Nom(PEUR)
                Vrg(\,)
        )
)
```

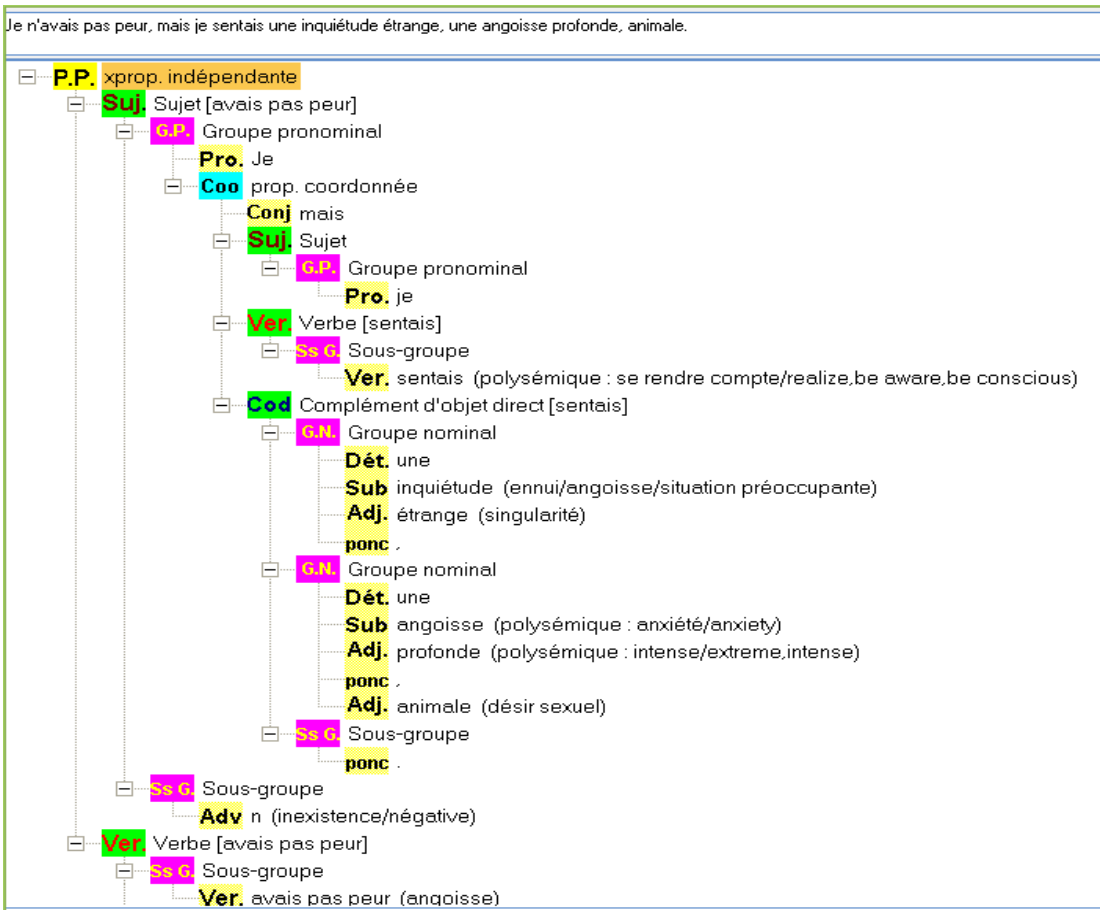



**Figure 32. Sortie de Cordial : la négation.**

### 5.5.3 Épellation

La sortie de l'analyse morphosyntaxique traitant l'épellation ne répondait pas à nos besoins.

Dans notre cas, l'épellation concerne les entités nommées, notamment les noms propres. En effet, notre lectrice épelle les noms des villes caractère par caractère, par exemple pour le mot « Bordeaux », elle prononce « B majuscule » sachant que le mot « majuscule » est prononcé suivi de « o » « r » « d » « e » « a » « u » « x ».

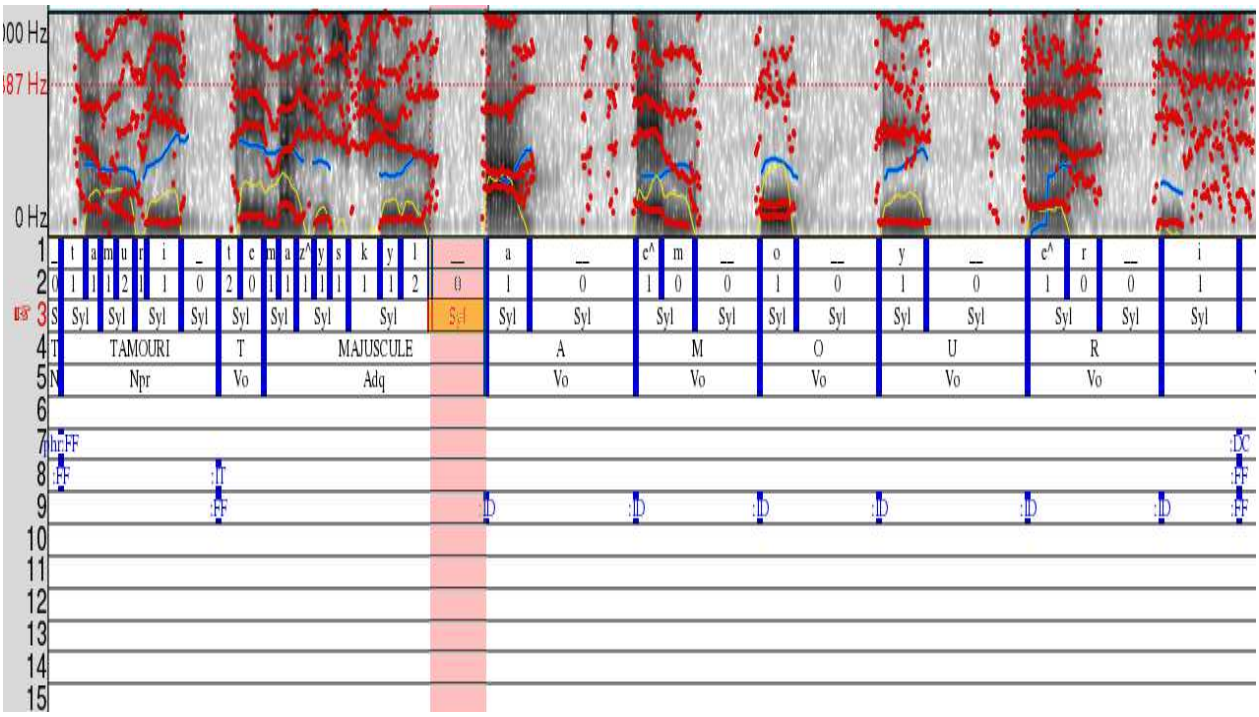



**Figure 33.** **Marquage prosodique de l'épellation.**

## 5.6 Prétraitement

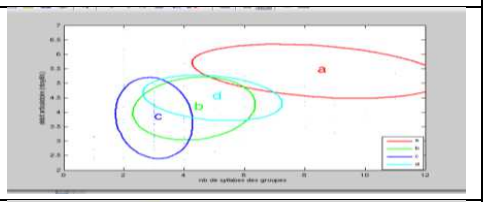
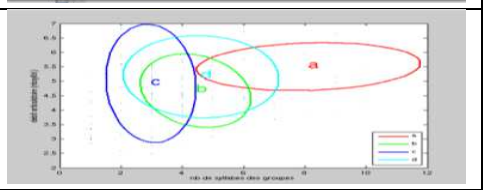
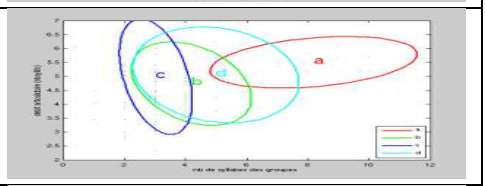
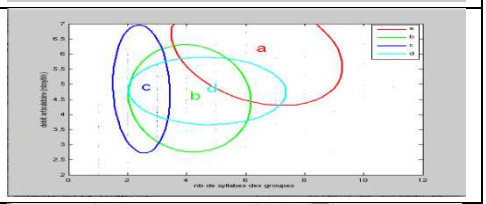
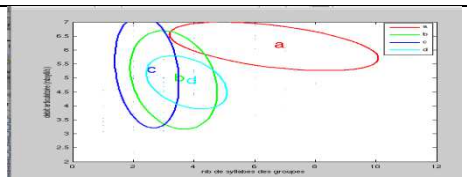
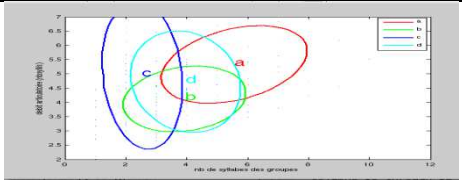
Face aux problèmes majeurs cités en amont, nous avons décidé de faire notre propre prétraitement. A cet effet, nous avons choisi le langage de programmation PERL qui est connu notamment pour sa puissance à gérer les expressions régulières. En partant d'une analyse morphologique en utilisant TreeTagger[22], nous avons créé nos propres règles de découpages ***(cf. annexe E)*** en s'appuyant sur : les catégories grammaticales de chaque mot, les signes de ponctuations dans le texte et enfin le modèle de Bailly permettant de gérer les différentes relations entre les mots dans la phrase. Notre programme perl appelé « dictée.pl » respecte bien les stratégies de prononciation de notre lectrice experte. Les résultats obtenus sont satisfaisants ***(cf. annexe F)***. Néanmoins, des corrections manuelles sont indispensables, notamment liées à la qualification et aux prépositions de complément de verbe/nom : ainsi

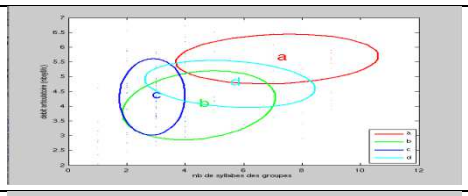
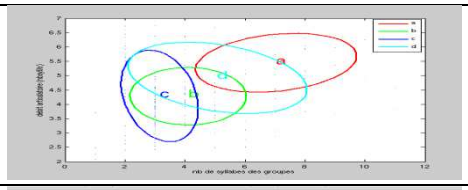
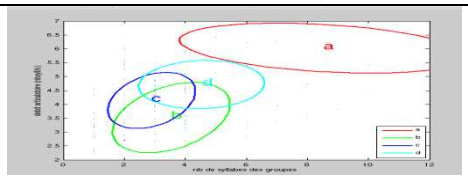
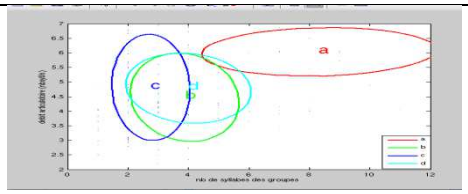
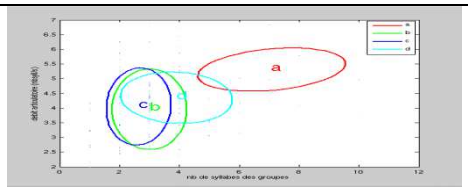
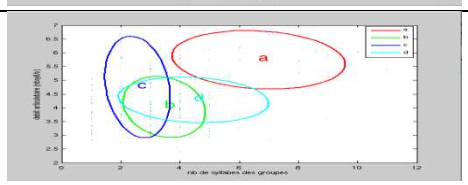
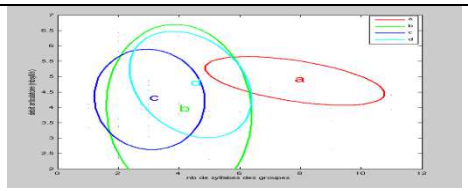
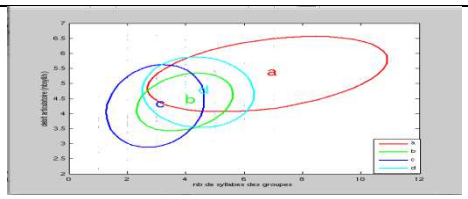
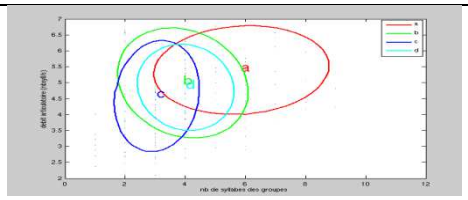
[22] TreeTagger est un outil qui permet d'annoter un texte avec des informations sur les parties du discours (genre de mots: noms, verbes, infinitifs et particules) et des informations de lemmatisation. Il a été développé par Helmut Schmid dans le cadre du projet « TC » dans le ICLUS (Institute for Computational Linguistics of the University of Stuttgart). TreeTagger permet l'étiquetage de l'Allemand, l'Anglais, le Français, l'Italien, l'Allemand, l'Espagnol, le Bulgare, Le Russe, le Grec, le Portuguais, le Chinois et les textes français anciens. Il est adaptable à d'autres langages si des lexiques et des corpus étiquetés manuellement sont disponibles. TreeTagger peut également être utilisé comme un « chunker » pour l'Anglais, l'Allemand et le Français (étiquetage des parties du discours, délimitation des groupes syntaxiques, étiquetage des groupes).» *(taln09.blogspot.fr/2009/02/etiquetage-morpho-syntaxique-et.html)*

dans « la pomme de terre sèche », l'adjectif « sèche » peut aussi bien qualifier « la pomme de terre » ou bien « la terre ». Ces corrections ont été faites sur la version (a). Enfin, nous appliquons l'apprentissage automatique sur chacune des versions (b), (c) et (d) pour apprendre à notre modèle SFC à avoir une bonne prosodie.

## 5.7 Corpus et caractérisations

En utilisant le langage de programmation Matlab, nous avons développé un script permettant la superposition des ellipses de dispersion des caractéristiques de réalisations des groupes de souffle pour les 4 versions (cf. tableau ci-dessous) (**a** : compréhension ; **b** : groupes syntaxiques ; **c** : mots de contenus ; **d** : relecture finale). En abscisse, la durée de phonation du groupe et en ordonnée, la durée de pause subséquente.

| Dictées | Titre | Superposition des ellipses |
|---|---|---|
| AUDE0000 | Avant l'orage. |  |
| AUDE0001 | Dans la jungle. |  |
| AUDE0002 | Drame en montagne. |  |
| AUDE0003 | Incendie dans les Landes. |  |
| AUDE0004 | La biche. |  |
| AUDE0005 | La chasse du renard. |  |

| | | |
|---|---|---|
| AUDE0006 | La petite fille et les oiseaux. |  |
| AUDE0007 | La rivière. |  |
| AUDE0008 | Le loup. |  |
| AUDE0009 | Le repas d'un paysan. |  |
| AUDE0010 | Le vieux noyer et les oiseaux. |  |
| AUDE0011 | Pauvre minet. |  |
| AUDE0012 | Un coiffeur novice. |  |
| AUDE0013 | La petite école de campagne. |  |
| AUDE0014 | La leçon d'histoire. |  |

En effet, on distingue trois comportements cruciaux : (**a**) est caractérisé par des groupes assez longs entrecoupés de pauses silencieuses relativement courtes ; (**c**) est au contraire constitué de groupes brefs suivis de longues pauses, permettant ainsi la transcription orthographique des mots ; (**b**) et (**d**) sont intermédiaires. La superposition moyenne des ellipses de dispersion des caractéristiques de réalisations des groupes de souffle pour les 4 versions (cf. ci-dessous).

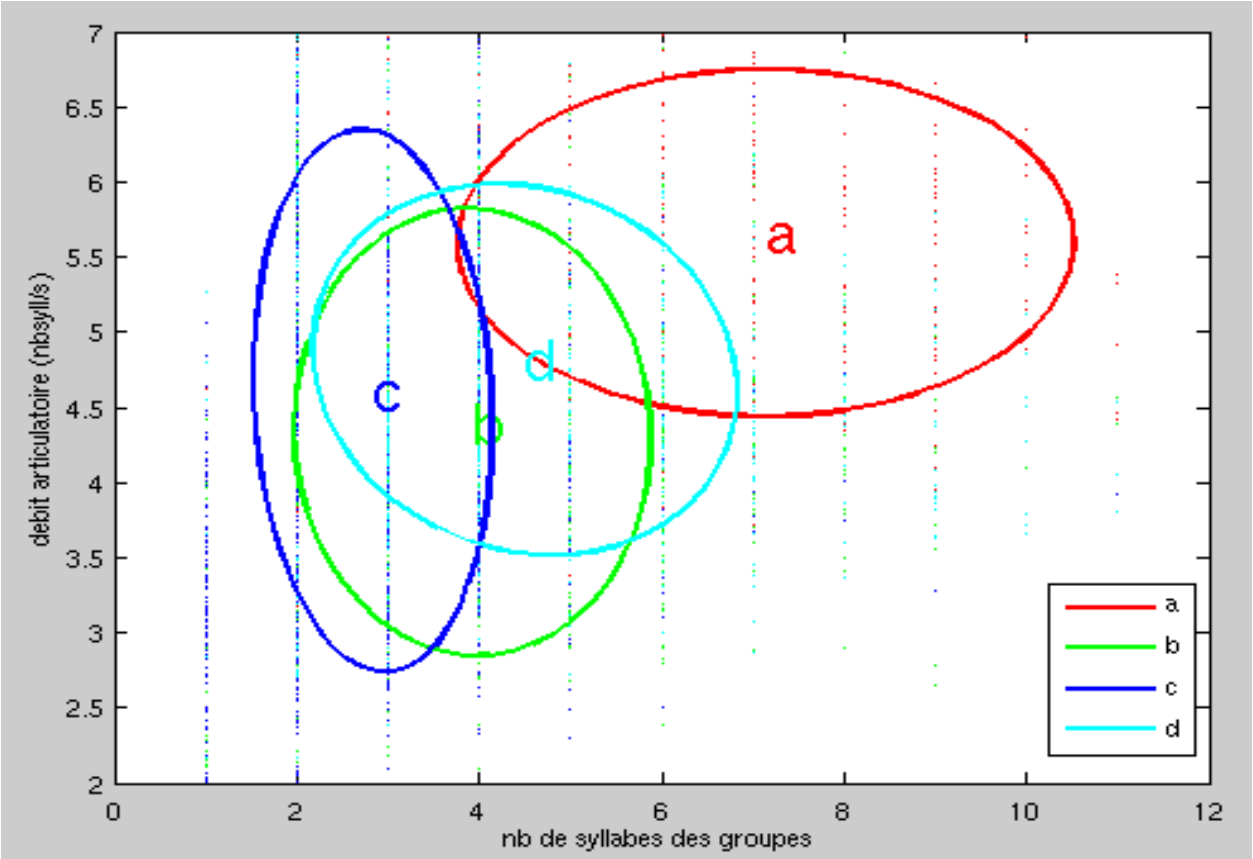



**Figure 34. Superposition des ellipses de dispersion des caractéristiques de réalisations des groupes de souffle pour les 4 versions ((a) : compréhension ; (b) : groupes syntaxiques ; (c) : mots de contenus ; (d) : relecture finale). En abscisse, la durée de phonation du groupe et en ordonnée, la durée de pause subséquente.**

## 5.8 Conclusions

Nous avons vu dans ce chapitre l'existence de verrous technologiques qui nous ont empêchés d'exploiter quelques outils existants (Cordial de Synapse). La solution que nous avons proposée consiste à créer nos propres règles de découpage en se basant sur l'analyse morphosyntaxique d'un nouvel outil (TreeTagger). Dans le chapitre suivant, nous faisons l'évaluation de notre système.

*« La perfection ce n'est pas de faire quelque chose de grand et de beau ; mais de faire ce que l'on fait avec grandeur et beauté ! »*
*[Swâmi Prâjnanpad]*

# 6 ÉVALUATIONS & RÉSULTATS

## 6.1 Introduction

Dans ce chapitre, nous expliquons et justifions, tout d'abord, pourquoi une telle évaluation est nécessaire. Pour cela, nous devrons répondre aux questions « Pourquoi évalue-t-on ? », « Qu'évalue-t-on ? » et « Comment évalue-t-on ? ». Ensuite, nous présentons les résultats des évaluations objectives et subjectives de notre système.

## 6.2 Évaluation

### 6.2.1 Pourquoi évalue-t-on ?

Un système est développé pour réaliser un but précis en répondant à un tel besoin. Ainsi, il faut tester, une fois la réalisation du système terminée, qu'il répond bien aux perspectives de la tâche qui lui appartient. Il est donc nécessaire de définir, dès le début, avec précision, les contraintes attachées à l'application terminale. C'est pourquoi, à partir du cahier des charges, nous déterminerons les solutions qui sont capables de résoudre les verrous attachés à la tâche qui lui incombe. Donc, nous sommes aptes de déterminer et de suggérer une série de tests permettant une évaluation efficace du système.

Notre hypothèse est d'étudier les modifications des stratégies d'énonciation, notamment les changements de contours prosodiques, de ces divers découpages et relectures du texte. Nous avons pour ceci analysé de manière fine le corpus de quinze dictées orales. Un ensemble restreint de marques prosodiques a été utilisé : notre hypothèse forte est que l'énonciation à voix haute consiste à faire écouter des fragments syntaxiques prélevés sur un arbre syntaxique complet. Nous faisons l'hypothèse que des diverses versions d'énonciation des groupes de mot ne modifient que la forme des contours mélodiques (c'est-à-dire une courbe mélodique la plus naturelle et cohérente possible) et rythmiques (de manière plus concrète, la durée de chaque phonème et des pauses) et non pas leur fonction.

Dans notre système, la lectrice effectue ainsi quatre passes sur le texte : une lecture (a) dite de compréhension, où le texte est lu une première fois de manière intégrale ; deux répétitions de groupes de mots, une version (b) lente suivie d'une version (c) où pratiquement tous les mots non grammaticaux sont prononcés de manière isolée ; puis une version (d) dite de vérification, permettant au scripteur de vérifier l'orthographe de chaque phrase.

Certes, notre système doit être capable de remplacer un lecteur humain c'est-à-dire avoir le maximum de ses « atouts», sans réunir d'autres désavantages. Les avantages d'une voix humaine sont qu'elle est intelligible et qu'elle s'adapte au rythme des scripteurs. Les points importants sur lesquels notre système de synthèse doit être performant sont donc l'intelligibilité et le rythme.

### 6.2.2 Qu'évalue-t-on ?

Selon la section préalable, il faut surtout définir le cadre adéquat des buts à réaliser afin d'évaluer sûrement un système. Cependant, même si la définition de la tâche finale est déterminée et complète, la sélection des critères demeure une difficulté complexe. Certes, il n'est pas évident de définir quel est l'ensemble des critères indispensables à l'évaluation du système.

Nous décomposons les critères d'évaluation en deux types : les critères objectifs et les critères subjectifs. Les premiers ne sont évalués que sur le système lui-même, pendant que les seconds demandent une expertise humaine. Les critères objectifs, même s'ils produisent une évaluation judicieuse du système, restent insuffisants pour énoncer un avis clair sur la qualité d'un système. Par conséquent, il est donc primordial d'exécuter les deux types d'évaluations afin d'avoir une idée stricte des points positifs et négatifs du système.

Dans le but d'évaluer notre système, nous  nous focalisons notamment sur les critères suivants :

- Niveau de la dictée
- L'intelligibilité de la dictée
- L'agrément de la voix
- La manière de lire la dictée
- La vitesse générale de lecture
- Le rendu des ponctuations
- La première lecture du texte
- La deuxième lecture en groupes
- La troisième lecture en mots
- La relecture finale pour vérifier votre transcription
- Ordre de préférence

### 6.2.3 Comment évalue-t-on ?

Dans le cadre de ces travaux, nous évaluons trois styles des dictées françaises :

1. DICTEE_AUDEo_* : nommée « originale » est l'original resynthétisé par la voix d'Agnès Afriat.
2. DICTEE_AUDEc_* : nommée «contrôle » est une condition où on a pris l'ensemble du marquage et fait un seul modèle moyen sans tenir compte des versions c'est-à-dire sans distinguer les quatre styles.
3. DICTEE_AUDEs_* : nommé «multi-style » est le résultat de notre travail où on a entrainé quatre versions différentes des mêmes marques.

Notons que les stimuli de tests sont générés par la **méthode de validation croisée** (Leave One Out). Pour chaque dictée, la dictée originale est exclue du corpus d'apprentissage et le modèle SFC est appris sur toutes les autres : il est alors utilisé pour générer la synthèse de la dictée. On peut ainsi relever que le modèle SFC de la version « contrôle » bénéficie d'environ quatre fois plus de contours pour chaque marque que la version « multi-style », qui distingue chaque marque instanciée dans chaque version.

### 6.2.4 Évaluation objective

Lors de la phase d'évaluation objective, tous les sujets ont réussi à faire leur dictée et à répondre aux questionnaires. Le tableau ci-dessous résume nos résultats obtenus par ***Sclite (en ligne :*** www1.icsi.berkeley.edu/Speech/docs/sctk-1.2/sclite.htm). En effet, l'évaluation objective nous a amené à quantifier différents types d'erreur ***(Cf. annexe J)*** : erreur grammaticale, erreur lexicale et erreur de reconnaissance. Nonobstant, une analyse quantitative seule n'est pas suffisante à l'établissement d'un diagnostic des points, forts et faibles, de notre système. Par conséquence, il est inévitable de passer à l'évaluation qualitative en termes d'intelligibilité, compréhension et charge cognitive.

SYSTEM SUMMARY PERCENTAGES by SPEAKER

| /tmp/tst | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| SPKR | # Snt | # Wrd | Corr | Sub | Del | Ins | Err | S.Err |
| org | 14 | 1214 | 96.5 | 2.7 | 0.7 | 0.0 | 3.5 | 84.6 |
| ctrl | 14 | 1318 | 94.6 | 4.9 | 0.5 | 0.3 | 5.7 | 78.6 |
| multi | 14 | 1194 | 96.3 | 3.0 | 0.7 | 0.1 | 3.8 | 84.6 |
| Sum/Avg | 40 | 3726 | 95.8 | 3.6 | 0.6 | 0.1 | 4.3 | 82.5 |
| Mean | 13.3 | 1242.0 | 95.8 | 3.6 | 0.6 | 0.1 | 4.3 | 82.6 |
| S.D. | 0.6 | 66.6 | 1.1 | 1.2 | 0.1 | 0.2 | 1.2 | 3.5 |
| Median | 13.0 | 1214.0 | 96.3 | 3.0 | 0.7 | 0.1 | 3.8 | 84.6 |

**Tableau 3. Résultats des scores de dictées : pourcentage de mots correctement orthographiés *versus* les fautes commises.**

### 6.2.5 Évaluation subjective

L'évaluation objective doit être complétée par une analyse subjective fondée sur la quantification de la capacité de notre système à fournir une bonne prosodie. Par conséquence, lors de la phase d'évaluation subjective, nous quantifierons dans un premier temps l'intelligibilité de la dictée, l'agrément de la voix, la manière de lire la dictée, la vitesse générale de lecture, le rendu des ponctuations, la première lecture du texte, la deuxième lecture en groupes, la troisième lecture en mots, la relecture finale pour vérifier votre transcription. Puis, nous déterminons la qualité de compréhension. Plutôt qu'un ordre de préférence global que l'on donnerait aux trois versions du système, il s'avère plus cohérent de représenter les résultats dans trois histogrammes différents (cf. ci-dessous).

## 6.3 Résultats

Afin d'évaluer la qualité de réalisation de nos productions, nous avons effectué notre test d'évaluation dans la chambre sourde du laboratoire. Nous avons eu quatorze sujets (cf. figure 34) dont quatre femmes et dix hommes. Les sujets ont au minimum le niveau du master. D'abord, nous leur avons leur demandé de répondre à un questionnaire général. Puis, leur tâche réside dans la transcription de trois dictées ayant trois styles différents présentant dans un ordre de présentation aléatoire, en répondant après chaque dictée à un questionnaire. La tâche finale est de donner un ordre de préférence de un à trois, qui nous permettra d'avoir un classement général de nos trois versions (le chiffre trois pour le meilleur et le chiffre un pour le moins bon) ***(cf. annexe G).***

| Sujet | Sexe | Age | Études | Domaine |
|---|---|---|---|---|
| 1 | F | 21 | Master | Ingénieur |
| 2 | F | 22 | Master | Ingénieur |
| 3 | F | 24 | Doctorat | Ingénieur |
| 4 | F | 43 | Master | SHS |
| 5 | M | 21 | Master | Informatique |
| 6 | M | 21 | Master | Ingénieur |
| 7 | M | 22 | Master | Ingénieur |
| 8 | M | 22 | Master | Ingénieur |
| 9 | M | 22 | Master | Ingénieur |
| 10 | M | 22 | Master | Ingénieur |
| 11 | M | 23 | Master | Ingénieur |
| 12 | M | 24 | Doctorat | Ingénieur |
| 13 | M | 41 | Doctorat | Ingénieur |
| 14 | M | 45 | Doctorat | Ingénieur |

**Figure 35. Tableau présentant la description des sujets.**

Les résultats par scripteur permettent d'avoir une étude détaillée mais complexe à analyser. La Figure 36 présente les résultats d'ordre de préférence par versions. Comme attendu, la meilleure satisfaction est pour la version « originale ». Les versions « contrôle » et « multi-style » ne sont pas distinguées. Concernant les versions de lecture, la Figure 38 montre que le problème principal de la version « multi-style » réside dans son rendu de la version (c), où chaque mot est énoncé de manière séparée.

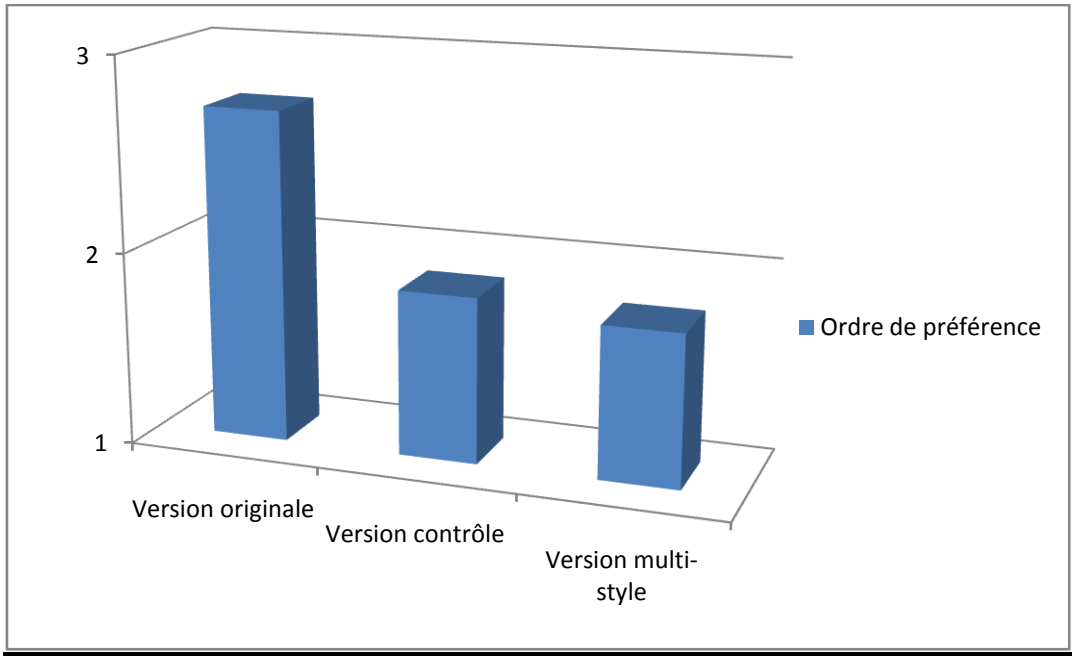



**Figure 36. Histogramme présentant les résultats d'ordre de préférence des trois styles.**

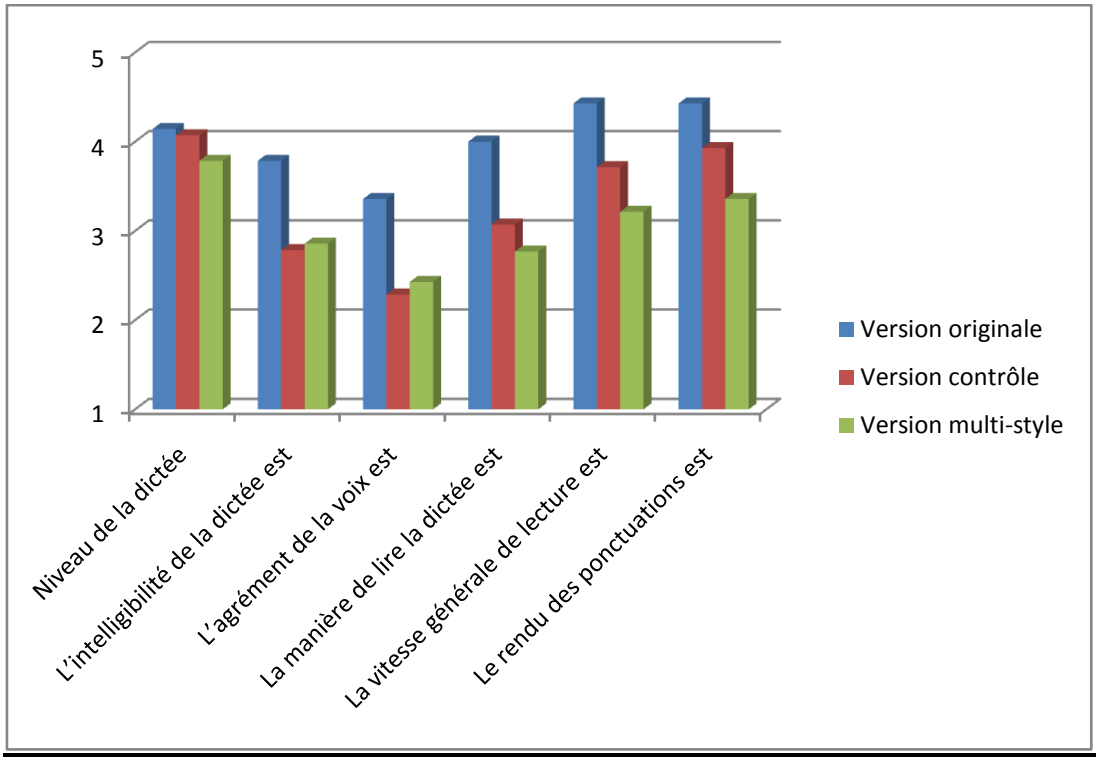



**Figure 37. Histogramme présentant les résultats d'évaluation du niveau, intelligibilité, agrément, manière de lecture, vitesse générale et rendu des ponctuations des trois styles.**

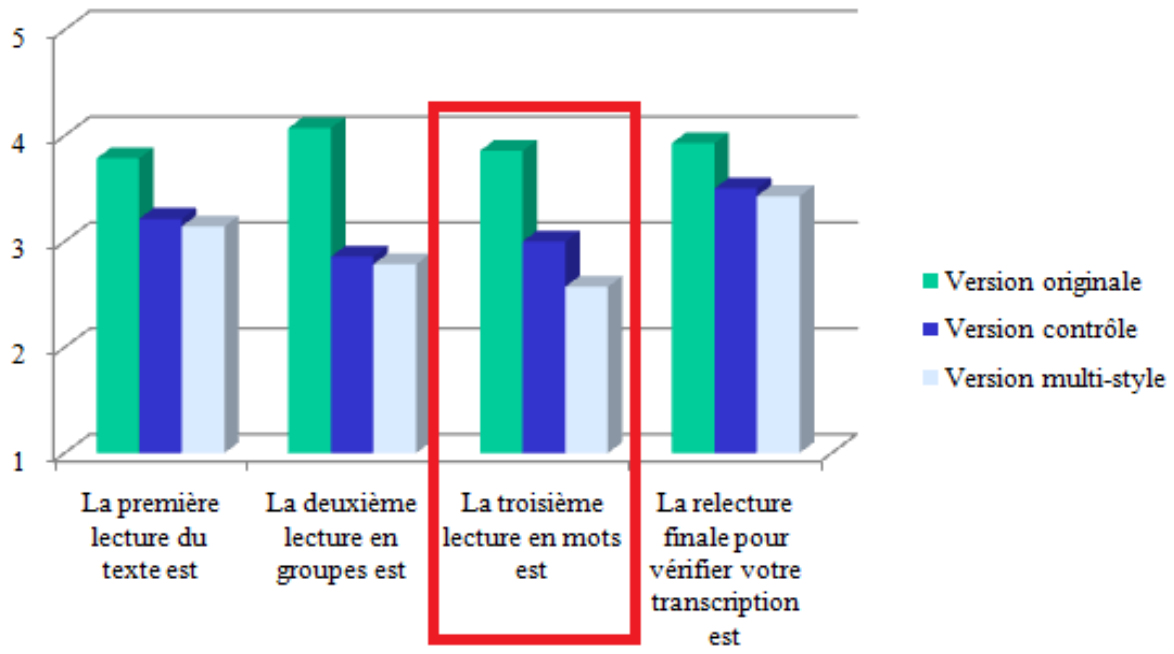



**Figure 38. Histogramme présentant les résultats d'évaluation entre les versions (a), (b), (c) et (d) des trois styles.**

## 6.4 Problèmes rencontrés

Au terme de cette étude, nous avons montré que le modèle prosodique SFC avait effectivement des répercutions sur la reconnaissance de certains mots qui étaient moins bien reconnus et par conséquent considérés par les sujets comme inconnus ou difficiles à identifier. Il faudrait donc étudier sur plus de sujets quels sont, statistiquement les mots les plus problématiques. Cependant, il serait intéressant de vérifier cette problématique en exploitant plus de corpus d'apprentissage automatique.

En conséquence, nous avons testé la capacité de ce modèle fonctionnel modulaire à capturer *les contours des virgules*. Ainsi, nous avons remarqué qu'il y avait un problème pour ces contours ce qui nuit au décodage des mots. Le contour montant est essentiellement porté par le « l » et le « schwa » ***(cf. annexe I)*** donc ignoré car le « schwa » n'est pas compté comme noyau syllabique. Tout cela permettrait de générer une – règle – (cf. ci-dessous) qui viendrait de compléter le modèle prosodique SFC.

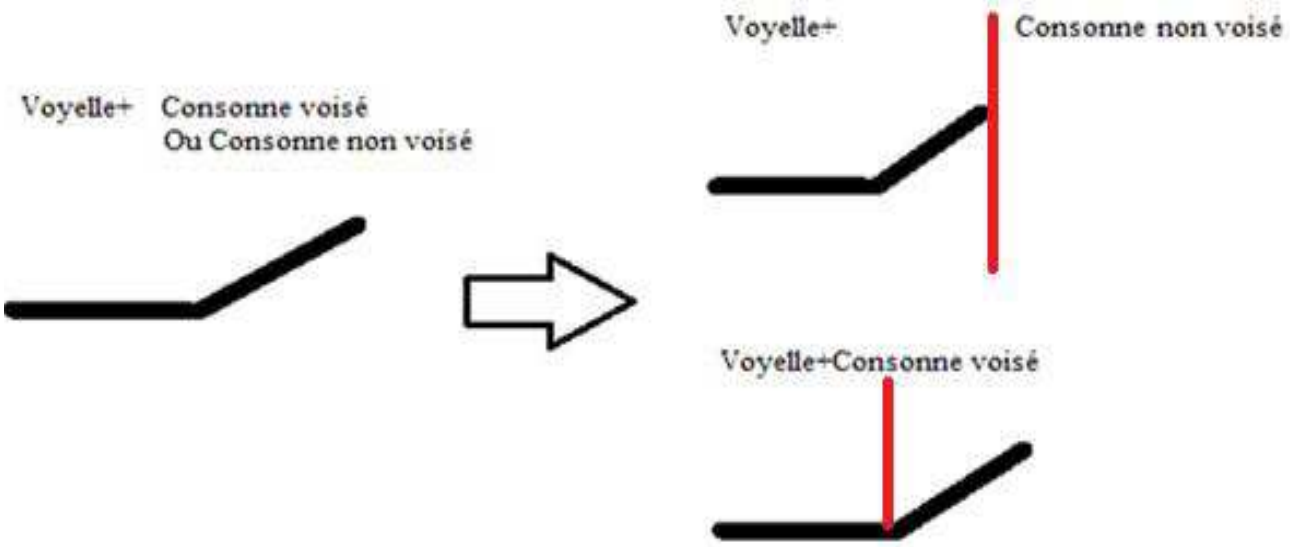



**Figure 39. Schéma présentant les deux cas de la rime**

Ainsi, cette règle nous permettrait d'anticiper les erreurs et y pallier. Il faudrait donc soit :

1. Styliser le F0 des syllabes non pas sur le noyau vocalique mais sur la rime (c'est faisable mais il faut reprogrammer de scripts en utilisant Matlab).
2. Changer manuellement, sous Praat, les finals schwas «q» par une voyelle pleine «x^», si cela est possible.

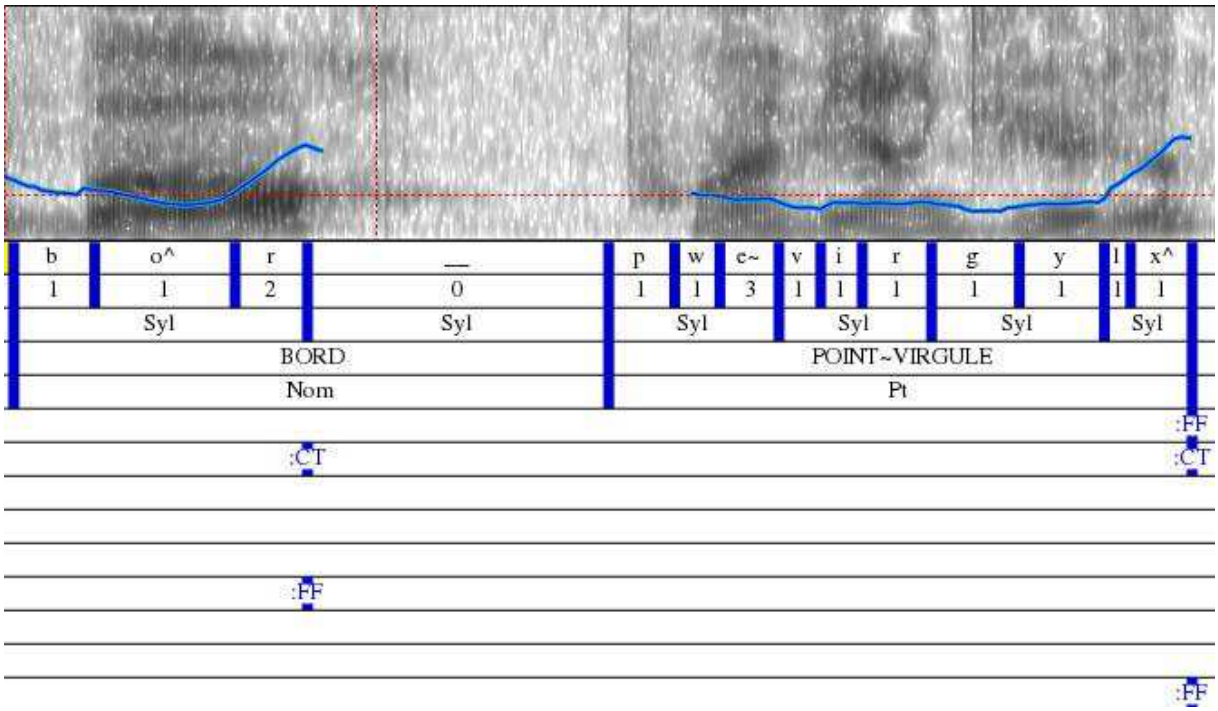



**Figure 40. Rime voisée : la montée finale est réalisée sur le [r] de « bord » alors que le contour est plat sur la voyelle.**

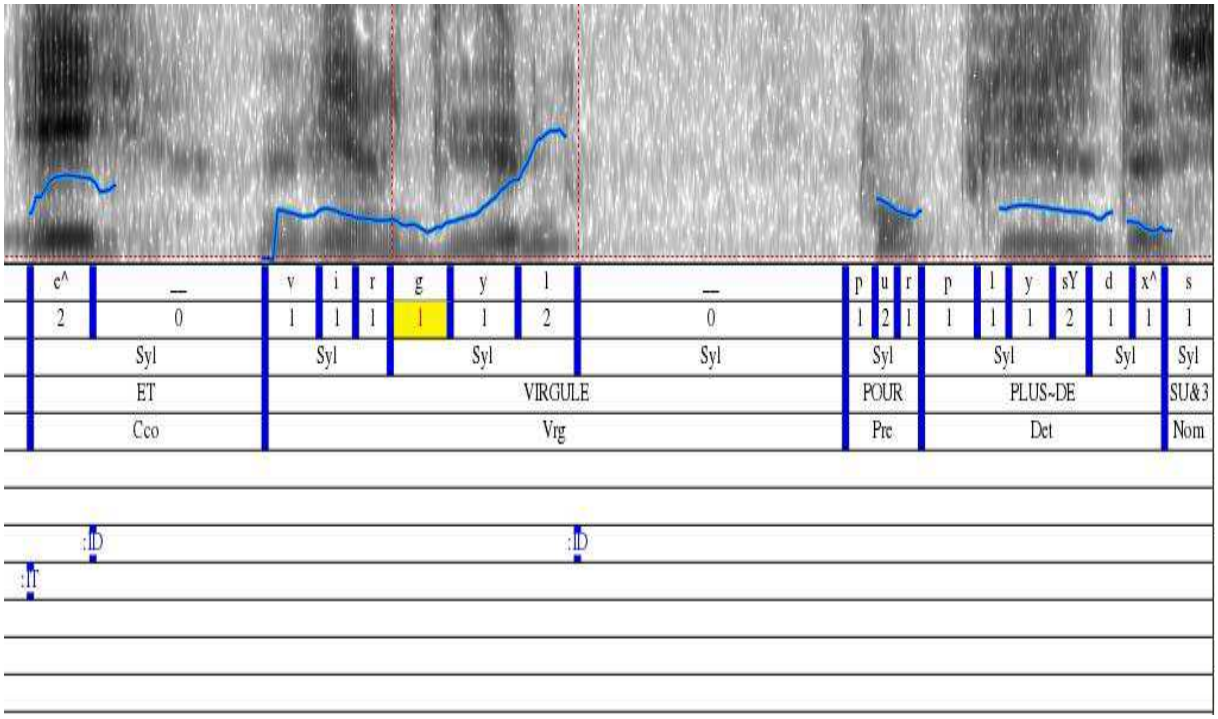



**Figure 41.** **Rime voisée : la montée finale est réalisée sur le [l] de « virgule ».**

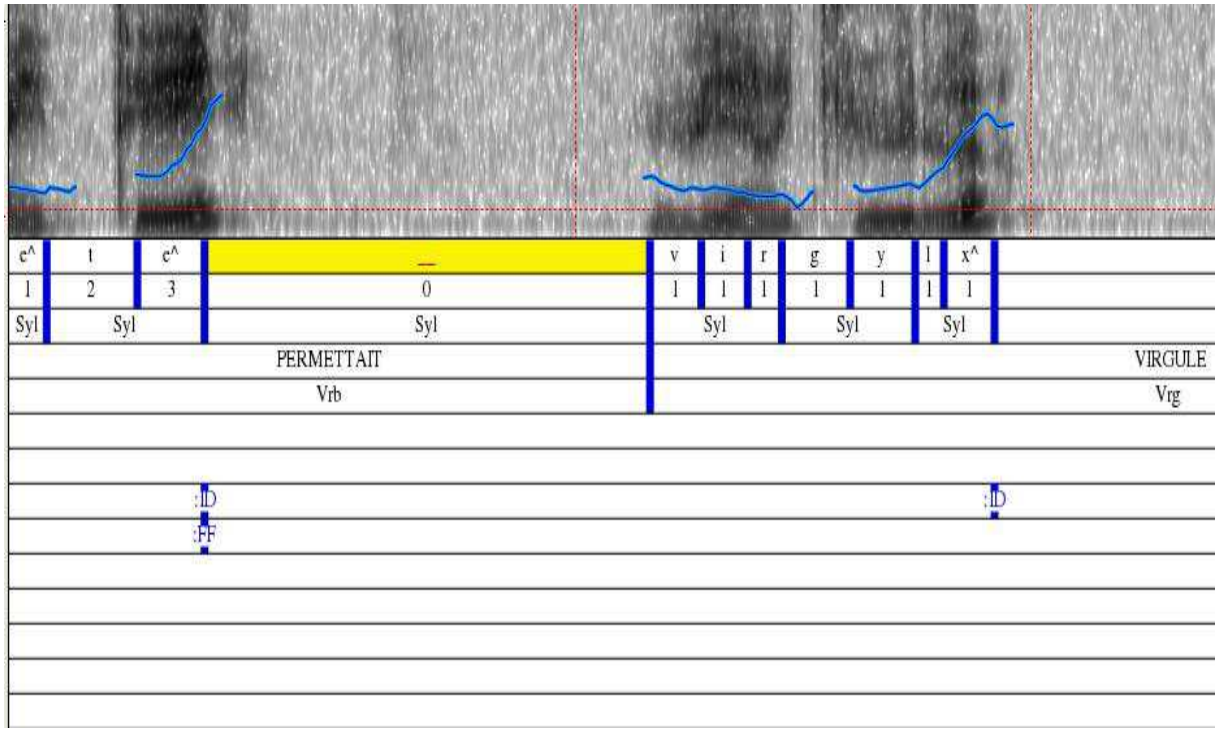



**Figure 42.** **Rime voisée : La montée finale est réalisée sur le schwa (étiqueté comme une voyelle pleine ici) de « virgule » alors que le contour est plat sur la voyelle.**

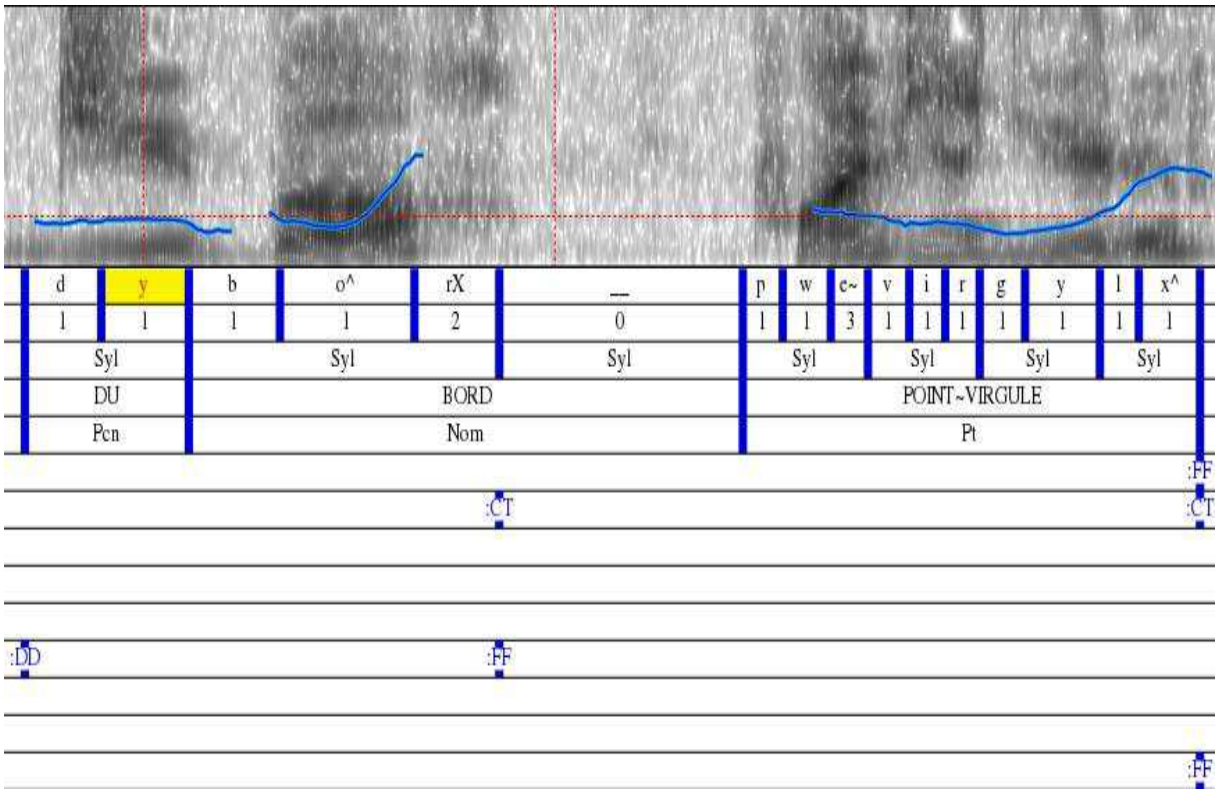



**Figure 43.** **Rime non voisée : la montée finale est réalisée sur la voyelle [□] de « bord ».**

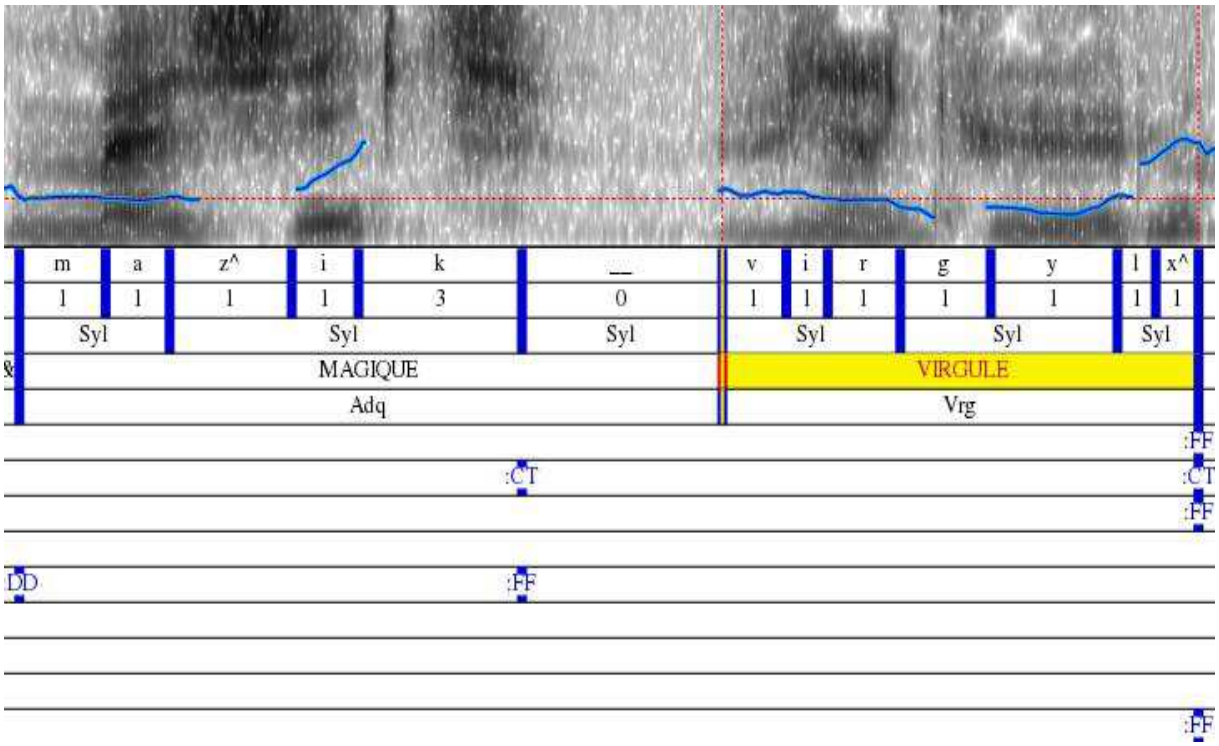



**Figure 44.** **Rime non voisée : la montée finale est réalisée sur la voyelle [i] de « magique ».**

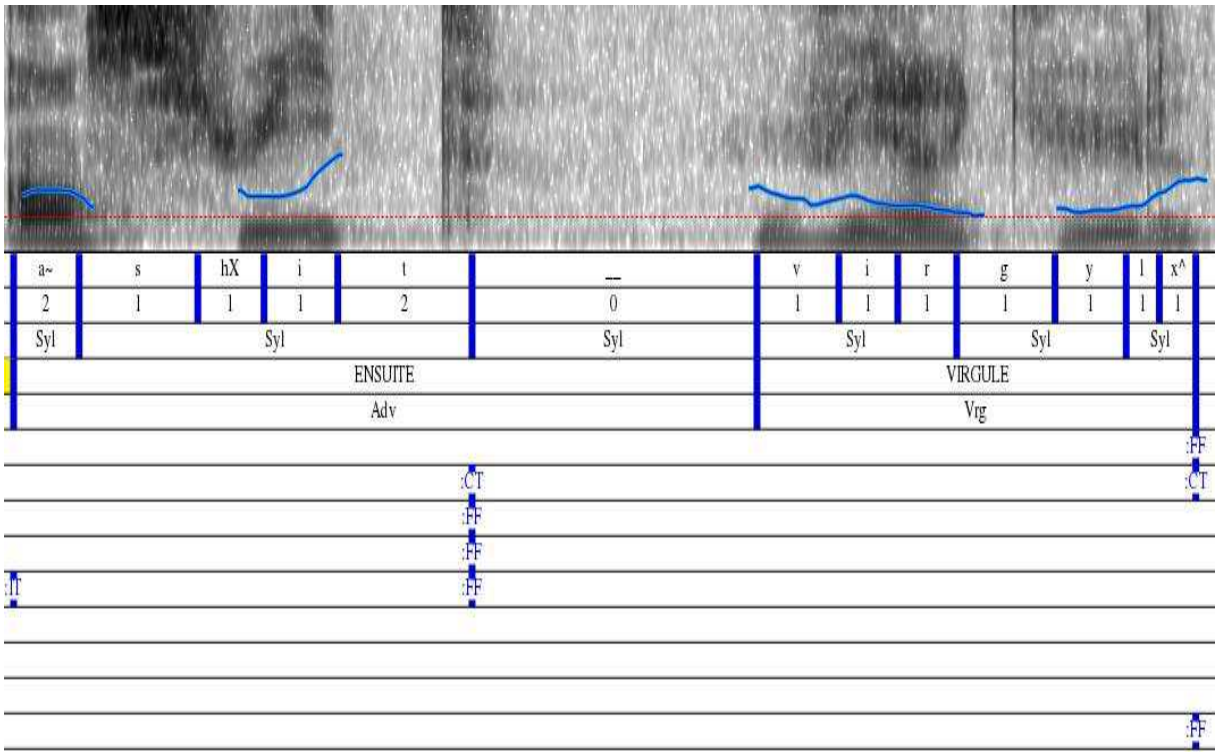



**Figure 45. Rime non voisée : la montée finale est réalisée sur la voyelle [i] de « ensuite ».**

## 6.5 Solution appliquée

Vu que nous n'ayons pas eu le temps pour reprogrammer des scripts, alors la deuxième solution était la plus commode. Ainsi, nous avons remplacé manuellement le schwa, noté « q » sous praat, par un noyau vocalique, noté « x^ » sous Praat, dans le but de capturer l'ensemble de contour.

## 6.6 Conclusion

En guise de conclusion, nous avons analysé les caractéristiques des paramètres acoustiques d'un système de synthèse de la parole afin de mettre en place les stratégies convenables. Puis, nous avons appliqué la formalisation de règles d'énonciation du niveau textuel (découpages, répétitions, etc.) et du niveau sonore (phonétisation, prosodie, etc.). Des techniques d'apprentissage automatique (arbres de régression, réseaux de neurones, chaînes de Markov) sont mises en œuvre pour mettre en correspondance l'arbre syntaxique avec les trajectoires de paramètres prosodiques en fonction des niveaux de lecture du texte. Enfin, nous avons évalué notamment la capacité du modèle SFC à capturer les régularités de cette correspondance. Il est indéniable que nous regrettons de ne pas avoir eu plus de sujets pour pouvoir donner encore plus d'influence à nos résultats, mais nous avons quand même des résultats satisfaisants.

*« Ce ne sont pas les heures qui sont précieuses, ce sont les minutes. »*
*[George Bernard Shaw]*

# 7 CONCLUSIONS & PERSPECTIVES

## 7.1 Conclusions

En guise de conclusion, ce travail ouvre de nombreuses réflexions de par l'aspect novateur du domaine. Très peu d'études existent sur l'étude des modifications des stratégies d'énonciation, notamment des changements de contours prosodiques, de ces divers découpages et relectures du texte, c'est pourquoi les tests effectués durant ce mémoire suscitent beaucoup d'interrogation et/ou de remise en cause des hypothèses. Grâce à ce travail, on voit que l'étude des modifications des stratégies d'énonciation, ces divers découpages et relectures du texte sont possibles et entièrement justifiés, comme domaine à part entière étant donné les spécificités qu'elle présente.

L'étude que nous avons menée, ici, a permis de montrer que les diverses versions d'énonciation des groupes de mots ne modifient que la forme des contours mélodiques (c'est-à-dire une courbe mélodique la plus naturelle et cohérente possible) et rythmiques (de manière plus concrète, la durée de chaque phonème et des pauses) et non pas leur fonction.

Nous avons également pointé le fait que dans la démarche de l'amélioration d'un système prosodique utilisé SFC, bien que la plupart des chercheurs jouent directement sur la durée du phonème comme le HTS, dans notre cas, la génération du coefficient d'allongement du syllabe s'est montrée plus efficace. Par ailleurs, extrêmement peu –à notre connaissance – de chercheurs ont étudié l'aspect « version contrôle » versus « version multi-style». Tout d'abord nous avons vu qu'il est possible de créer un nouveau marqueur nommé **:CT** qui désigne la continuité de la parole, notamment, pour les contours des virgules.

Nos résultats sont satisfaisants bien que les pourcentages d'ordre de préférence pour la « version contrôle » versus « version multi-style » étaient presque égaux. Nous en avons conclu que la reconnaissance des mots était problématique, qu'il s'agisse d'une « version contrôle » ou d'une « version multi-style».

Cette étude a permis, aussi, de montrer une partie des spécificités de la voix synthétisée en comparaison à la voix naturelle, démontrant qu'elle devait être traitée différemment de la voix naturelle.

## 7.2 Perspectives

Durant ces cinq mois de recherche, nous avons utilisé un modèle connexionniste de génération biparamétrique de la prosodie. Ce modèle s'est fondé sur l'approche théorique d'Aubergé et Bailly *(1995)* qui demande un encodage immédiat des informations purement linguistiques et paralinguistiques en passant par un lexique de formes intonatives prototypiques.

Le modèle, dans sa version actuelle, ne réalise pas une génération complète de la prosodie pour la synthèse à partir du texte : intensité, timbre, assimilation de voisement, etc. Le modèle devra prendre en compte des phénomènes de réductions prosodiques au niveau rythmique et qui se traduisent par l'effacement progressif de certaines marques prosodiques en fonction du débit d'élocution comme le « schwa ». Ainsi, l'utilisation du nouveau marqueur **:CT** nous a permis de mener à bien ce projet.

La stylisation prosodique doit mieux prendre en compte la dynamique mélodique (notamment le contour porté par la rime suivant que les consonnes finales sont ou non voisées) et intégrer d'autres paramètres prosodiques, dont notamment les *bruits respiratoires* (cf. ci-dessous).

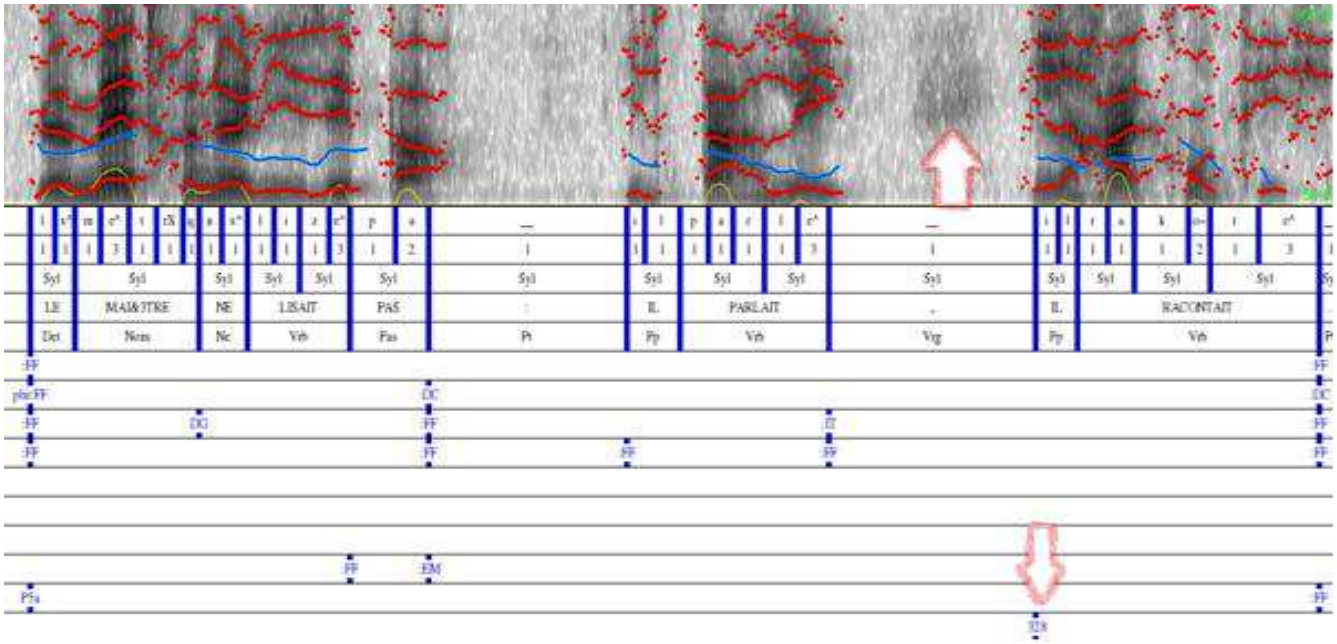



**Figure 46.** **Bruits respiratoires**

# SITOGRAPHIE

**-Dernière consultation : juillet 2014-**

*(Le Megueur, S. 2013)* [tel-00913565, v1] Évaluation Expérimentale D'un Système Statistique de Synthèse de La Parole, HTS, Pour La Langue Française - These.pdf. ***En ligne :*** http://tel.archives-ouvertes.fr/docs/00/91/35/65/PDF/these.pdf, accessed August 6, 2014.

*(D'Alessandro, C. & Richard, G., 2013)* Analyses et Traitements Linguistiques | Techniques de l'Ingénieur. ***En ligne*** **:** http://www.techniques-ingenieur.fr/base-documentaire/technologies-de-l-information-th9/documents-numeriques-technologies-d-acquisition-et-de-restitution-42310210/synthese-de-la-parole-a-partir-du-texte-h7288/analyses-et-traitements-linguistiques-h7288v2niv10002.html, accessed August 5, 2014.

(Berard, L., 2012) Sommaire - These_Berard_Lolita_2012.pdf. **En ligne :** http://www.atilf.fr/IMG/pdf/theses/These_Berard_Lolita_2012.pdf, accessed August 7, 2014.

*(DEGELS, 2011)* ***En ligne :*** www.degels2011.limsi.fr/actes/actes_degels2011.pdf

*(Roekhaut, S. & Beaufort, R., 2011)* taln2011-Beaufort-Roekhaut-Ref.pdf. ***En ligne :*** http://cental.fltr.ucl.ac.be/team/beaufort/file/taln2011-beaufort-roekhaut-ref.pdf, accessed August 4, 2014.

*(Ferré, G. 2011)* ***En ligne :*** www.hal.archives-ouvertes.fr/docs/00/60/91/24/PDF/Multimodalite-Ferre.pdf

*(La Meuse Luxembourg, 2011)* PlatON va Vous Faire Faire Des Dictées : Le Blog de Nassogne - 20110908-Sudpresse-PlatON.pdf. ***En ligne :*** http://cental.fltr.ucl.ac.be/team/beaufort/press/20110908-sudpresse-PlatON.pdf, accessed August 4, 2014.

*(Ferré, G., 2011)* [hal-00609124, v1] Analyse Multimodale de La Parole - Multimodalite-Ferre.pdf. ***En ligne :*** http://hal.archives-ouvertes.fr/docs/00/60/91/24/PDF/Multimodalite-Ferre.pdf, accessed August 4, 2014.

*(Mertens, P., 2009)* ActesIDP09.pdf - Mertens.pdf
***En ligne :*** http://makino.linguist.jussieu.fr/idp09/docs/IDP_actes/Articles/Mertens.pdf, accessed August 6, 2014.

*(Boidin, C. & Boeffard, O., 2008)* jep2008_v4_revised - 102_jep_1576.pdf. ***En ligne :*** http://www.afcpparole.org/doc/Archives_JEP/2008_XXVIIe_JEP_Avignon/PDF/avignon2008_pdf/JEP/102_jep_1576.pdf, accessed August 6, 2014.

*(Richard, G., 2008)* IAD-PS-Parol-Transp5 - IAD_Cours7.pdf. ***En ligne :*** http://www.perso.telecom-paristech.fr/~grichard/Enseignements/IAD/IAD_Cours7.pdf, accessed August 6, 2014.

*(Gibert, G., 2006)* gibert_these06.pdf. ***En ligne :*** http://www.gipsa-lab.grenoble-inp.fr/~gerard.bailly/publis/synthese/_gg/gibert_these06.pdf, accessed August 5, 2014.

*(Ruelle, A. & Beaufort, R., 2006)* Titre Du Document - jep2006_rbar.pdf. ***En ligne :*** http://www.tcts.fpms.ac.be/publications/papers/2006/jep2006_rbar.pdf, accessed August 4, 2014.

*(Boite, R. et al., 2000)* René Boite, Hervé Bourlard, Thierry Dutoit, Joel Hancq et Henri Leich: Traitement de La Parole - Google Livres ***En ligne :*** http://books.google.fr/books?id=kFS-HMNd3HcC&pg=PA357&lpg=PA357&dq=d%27analyseurs+probabilistes+:+les+n-grammes&source=bl&ots=ZAVXKmmWRB&sig=NStVNKJRANKVO8iEUggNdERdIqI&hl=fr&sa=X&ei=PFdFU8jsOoam0AWKhIDIAg&ved=0CDMQ6AEwAQ#v=onepage&q=d%27analyseurs%20probabilistes%20%3A%20les%20n-grammes&f=false, accessed August 5, 2014.

*(Goldman, J.P.)* Microsoft Word - CLF23_9_Goldman_déf.doc - CLF23_9_Goldman.pdf
***En ligne :*** http://www.unige.ch/lettres/linguistique/Auchlin/MA-DiscoursProsodie/docs_son_CLF23/Articles_pdf.pdf/CLF23_9_Goldman.pdf, accessed August 7, 2014.

*(Constant, M.)* Modèles Probabilistes Pour L'ingénierie Linguistique - C4.pdf. ***En ligne :*** http://igm.univ-mlv.fr/~mconstan/enseignement/m1/tal/C4.pdf, accessed August 6, 2014.

*(Beaufort, R. & Colotte, V.)* Synthèse Vocale Par Sélection Linguistiquement Orientée D'unités Non-Uniformes : LiONS | Richard Beaufort - Academia.edu. ***En ligne :*** http://www.academia.edu/305292/Synthese_vocale_par_selection_linguistiquement_orientee_dunites_non-uniformes_LiONS, accessed August 6, 2014.

Traitement Automatique Du Langage Naturel: Tree Tagger : Etiquetage Morpho-Syntaxique et Lemmatisation. ***En ligne :*** http://taln09.blogspot.fr/2009/02/etiquetage-morpho-syntaxique-et.html, accessed August 5, 2014.

*(Didier Schwab)* ***En ligne :*** www.getalp.imag.fr/homepages/schwab/pmwiki/pmwiki.php?n=Recherche.Glossaire
www.xlstat.com/fr/produits-solutions/fonctionnalite/arbres-de-classification-et-de-regression.html

*(Grobet, A.)* La pronctuation prosodique dans les dimensions périodique et informationnelle du discours]. ***En ligne :*** www.clf.unige.ch/display.php?numero=19&idFichier=183

*(Peeters, G.)* ***En ligne :*** www.recherche.ircam.fr/anasyn/peeters/ARTICLES/Peeters_2001_PhDThesisv1.1.pdf

*(Le Beux, S., 2009)* Thèse Le Beux Sylvain; contrôle gestuel de la prosodie et de la qualité vocale ; 2009) ***En ligne :*** www.halshs.archives-ouvertes.fr/docs/00/61/84/27/PDF/These_LeBeux2009.pdf

*(Hirst, D. et al., 2006)* [Hirst Daniel, Martin Philippe & Mertens Piet] ***En ligne :*** http://halshs.archives-ouvertes.fr/docs/00/25/63/95/PDF/2921.pdf

*(Ardaillon, L., 2013)* ***En ligne :*** http://www.atiam.ircam.fr/Archives/Stages1213/ARDAILLON_Luc.pdf

*(Colotte, V. & Beaufort, R., 2004)* ***En ligne :***
http://www.tcts.fpms.ac.be/publications/papers/2004/jep2004_vcrb.pdf

*(Mareuil, P.B.)* ***En ligne :*** http://fox.diwi.org/zen/dea/DEA_ling.pdf

*(Hugues V. et al.)* (Hugues Vinet, Gérard Assayag, Juan José Burred, Grégoire Carpentier, Nicolas Misdariis, Geoffroy Peeters, Axel Roebel, Norbert Schnell, Diemo Schwarz, Damien Tardieu) ***En ligne :*** http://articles.ircam.fr/textes/Vinet11a/index.pdf

*(Touati, R.)* ***En ligne :*** http://gerflint.fr/Base/Algerie13/touati.pdf

*(Brognaux, S.)* Sandrine Brognaux, Sophie Roekhaut, Thomas Drugman, Richard Beaufort ***En ligne :*** http://media.wix.com/ugd/5f9481_419d7996a92c4da38788c207a81ec2c8.pdf

www.link.springer.com/chapter/10.1007%2F978-1-4612-2258-3_15#page-1

www.ra5-telecharger.phpnuke.org/fr/c113438/praat-gratuit#.U3lPinlZrcc

www.moulaytahar.forumparfait.com/la-prosodie-par-slimani-rachida-vt30.html

www.python.sm.u-bordeaux2.fr/ter/2011/sc/ducasse-jeunet-poublang/anat_lang.php

www.makammoun.mypressonline.com/index_fichiers/Publications/Conferences/JTEA_08/JTEA_Med.pdf

www.msagent.free.fr/synth/prosodie.htm, accessed August 6, 2014.

# RÉFÉRENCES BIBLIOGRAPHIQUES

**Nous avons utilisé « Zotero » qui permet de gérer les données bibliographiques et les documents de recherche. *En ligne* : https://www.zotero.org/**


(Polyphone pour Ivan Fónagy) : mélanges offerts en hommage à Iván Fónagy par un groupe de disciples, collègues et admirateurs.

(Bailly, G. & Benoit, C.) : Talking machines, theories, models and designs.

(Chaudiron, S., Choukri, K.) L'évaluation des technologies de traitement de la langue.

*(Roekhaut, S., Beaufort, R. & Fairon, C. 2013)* "PLATON, un outil de dictée automatique au service de l'apprentissage de l'orthographe " Le Langage et l'Homme 48(2): 127-140.

*(Pazos, C.G., 2012)* La dictée sans faute comme moyen d'améliorer l'orthographe des élèves. Genève.

*(Beaufort, R. & S. Roekhaut, 2011)* Le TAL au service de l'ALAO/ELAO. L'exemple des exercices de dictée automatisés. Conférence sur le Traitement Automatique des Langues Naturelles (TALN), Montpellier.

*(BOERSMA, P. & WEENINK, D., 2009)* Praat : doing phonetics by computer (version 5.1.05) [computer program].

*(Van NIEKERK, D. & BARNARD, E., 2009)* Phonetic alignment for speech synthesis in underresourced languages. In Interspeech.

*(ADELL, J., BONAFONTE, A., GOMEZ, J. A. & CASTRO, M. J., 2005)* Comparative study of automatic phone segmentation methods for TTS. In ICASSP.

*(Bailly, G. & Holm, B., 2005)* "SFC: a trainable prosodic model." Speech Communication

46(3-4): 348-364.

*(KAWAI, H. & TODA, T., 2004)* An evaluation of automatic phone segmentation for concatenative speech synthesis. In ICASSP.

*(SCHIEL, F. & DRAXLER, C., 2003)* The production of speech corpora. Rapport technique, Bavarian Archive for Speech Signals.

*(Mercier, G., Guyomard, M., Siroux, J., Bramoullé, A., Gourmelon, H. Guillou, A. & (Dicristo, A., 2000)* Interpréter la prosodie. In XXII émes Journées d'études sur la Parole, Aussois, France, 2000.

*Lavannant, P., 2000)* Courseware for breton spelling pronunciation and intonation learning. Proceedings of InSTIL, Dundee, England 145-148.

*(Ruggia, S., 2000)* "La dictée interactive." Alsic 3 DOI: 10.4000/alsic.1820.

*(Holm, B., G. Bailly & Laborde, C., 1999)* Performance structures of mathematical formulae. International Congress of Phonetic Sciences, San Francisco, USA**:** 1297-1300]).

*(Santiago-Oriola, C., 1999)* Vocal synthesis in a computerized dictation exercise.Eurospeech, Budapest, Hungary: 191-194.

*(Coniam, D., 1998)* "Interactive evaluation of listening comprehension: How the context may help." Computer Assisted Language Learning 11(1): 35-53.

*(Black, A., Lenzo, K. & Pagel, V., 1998)* Issues in building general letter-to-sound rules. ESCA Workshop on Speech Synthesis, Jenolan Caves, Australia: 77-80.

*(Santiago-Oriola, C., 1998)* "Système vocal interactif pour l'apprentissage des langues : la synthèse de la parole au service de la dictée." U. de Toulouse le Mirail. Toulouse. IRIT.

*(Aubergé, V., Grépillat, T. & Rilliard, A., 1997)* Can we perceive attitudes before the end of sentences ? The gating paradigm for prosodic contours. Proceedings of the European Conference on Speech Communication and Technology, Rhodes - Greece**:** 871-874.]

*(Barbosa, P. & Bailly, G., 1997)* Generation of pauses within the z-score model. Progress in Speech Synthesis. J. P. H. V. Santen, R. W. Sproat, J. P. Olive and J. Hirschberg. New York, Springer Verlag**:** 365-381.]

*(Guyomard, M., Siroux, J, Pernici, D. & Royer, C., 1997)* A general public application of pedagogic and linguistic vocations of speech synthesis: ordictée. Research on Computational Linguistics (ROCLING), Taipei, Taiwan: 316-32

*(Van Heuven, V. J., Haan, J. , Janse, E. & van der Torre, E. J., 1997)* Perceptual identification of sentence type and the time-distribution of prosodic interrogativity marker in dutch. ETRW Workshop on Prosody, Athens - Greece**:** 317-320.]

*(Coniam, D., 1996)* "Computerized dictation for assessing listening proficiency." Calico Journal 13: 73-86.

*(Aubergé, V. & Bailly, G., 1995)* Generation of intonation : a global approach. Dans Proceedings of the European Conference on Speech Communication and Technology, pages 2065–2068, Madrid, 1995.

*(Barbosa, P. & Bailly, G., 1994)* "Characterisation of rhythmic patterns for text-to-speech synthesis." Speech Communication 15: 127-137.]

*(Alissali, M. & Bailly, G., 1991)* COMPOST: un serveur de synthèse multilingue. 8 éme Congrès sur la Reconnaissance de Formes et l'Intelligence Artificielle, Lyon-Villeurbanne: 183-192.

*(Bailly, G. & Alissali, M., 1992)*. "COMPOST: a server for multilingual text-to-speech system." Traitement du Signal 9(4): 359-366.

(Aubergé, V., 1991) : Aubergé Véronique. La synthèse de la parole : ”des règles aux lexiques”. Thèse de doctorat, Université Pierre Mendès-France, Grenoble – France, 1991. sous la direction de Rouault, J.

*(Charpentier, F. & Moulines, E., 1989)* Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. In Proceedings of the European Conference on Speech Communication and Technology (Eurospeech), pages 2013-2019, 1989.

*(Catach, N., 1989)* Les délires de l’orthographe: en forme de dictionnaire [sic], préf. de Philippe de Saint-Robert. Paris.

*(Fujisaki, H. & Hirose, K., 1984)* Analysis of voice fundamental frequency contours for declarative sentences of japanese. Journal of the Acoustic Society of Japan (E), 5(4) : 233–241, 1984.

*(Bailly, G., 1983)* Contribution à la détermination automatique de la prosodie du français parlé à partir d'une analyse syntaxique. Etablissement d'un modèle de génération. Thèse de troisième cycle, Institut National Polytechnique de Grenoble, Grenoble, France, 1983.

*(Hirst, 1991 ; Pierrehumbert, J. 1981)* "Synthetizing intonation." Journal of the Acoustical Society of America 70 (4): 985-995).

# 8 TABLE DES ANNEXES

## *ANNEXE A :* Jeu de description proposé & Catégorisation des descripteurs

| | Horizon | Description |
|---|---|---|
| Pho. | | Label phonétique du segment courant |
| | | Labels phonétiques des segments précédent/suivant |
| | | Labels phonétiques des segments précédent-précédent/suivant-suivant |
| Syllabe | P/C/S | Nb. phones + position du phone courant dans la syl. |
| | C | Position de la syllabe dans le mot |
| | C | Position de la syllabe dans le Groupe de Souffle (GS) |
| | P/C/S | Information d'accentuation |
| | C | Nb. de syl. depuis la dernière accent./syl. cour. jusque la syl. cour./prochaine acc. |
| | C | Nb. de syllabes acc. avant/après la syllabe courante dans le GS |
| | C | Voyelle de la syllabe |
| Mot | P/C/S | Nb. de syllabes dans le mot |
| | C | Position du mot dans le GS |
| | P/C/S | Tag grammatical du mot |
| | C | Nb. de mots depuis le dernier sign./mot cour. jusqu'au mot cour./prochain sign. |
| | C | Nb. de mots signifiants avant/après le mot cour. dans le GS |
| GS | P/C/S | Nb. de syl. dans le GS |
| | P/C/S | Nb. de mots dans le GS |
| | C | Position du GS dans l'énoncé |

| | Caté. | Description |
|---|---|---|
| Pho. | | Label phonétique du segment courant |
| | | Labels phonétiques des segments précédent/suivant |
| | | Labels phonétiques des segments précédent-précédent/suivant-suivant |
| Syllabe | | Nb. phones + pos. du phone courant dans la syl. |
| | | Position de la syllabe dans le mot |
| | | Position de la syllabe dans le Groupe de Souffle (GS) |
| | | Information d'accentuation |
| | | Nb. de syl. depuis la dernière accent./syl. cour. jusque la syl. cour./prochaine acc. |
| | | Nb. de syllabes acc. avant/après la syllabe courante dans le GS |
| | | Voyelle de la syllabe |
| Mot | | Nb. de syllabes dans le mot |
| | | Position du mot dans le GS |
| | | Tag grammatical du mot |
| | | Nb. de mots depuis le dernier sign./mot cour. jusqu'au mot cour./prochain sign. |
| | | Nb. de mots signifiant avant/après le mot cour. dans le GS |
| GS | | Nb. de syl. dans le GS |
| | | Nb. de mots dans le GS |
| | | Pos. du GS dans l'énoncé |

## *ANNEXE B :* Le Modèle de marquage de Bailly (ceci n'est qu'une interprétation possible)

Le modèle de marquage est un modèle de transfert suggérant une conversion au niveau de la forme : forme du contenu et forme de l'expression. Les modèles de transfert ainsi suggérés utilisent habituellement une relation non bijective entre un espace bidimensionnel-celui des arborescences ou structures de surface qui sont une représentation de la structure syntactico-morphologique du texte- et un espace unidimensionnel, -projection de l'arbre sur l'axe temporel-. Les liaisons entre les composants dans le plan syntagmatique sont décrites de ce fait par un système de marques prosodico-syntaxiques : marques relationnelles.

Bailly (1986) présente un marquage du texte issu de la projection sur l'axe syntagmatique d'un arbre syntaxique. Cet arbre est construit à partir des informations données par une grammaire de dépendance (Tesnière 1965; Hays 1964; Veillon 1970), signifiant les liaisons syntaxiques que les mots maintiennent entre eux. Dans une telle grammaire, ces liaisons listent les classes lexicales attachées à une autre qui peut les dominer (cette dernière est appelée le gouverneur). La racine de cet arbre est généralement le verbe principal de la phrase (cf. figure 46).

Dans un tel système, tout mot lexical non clitique, accentuable, obtient une marque qui désigne la liaison maintenue par celui-ci et le non clitique suivant. Le système de Bailly est appuyé sur deux approches théoriques. En premier lieu, il s'agit d'une généralisation du concept de frontière majeure et mineure, développée par Di Cristo (1981) : les marques de Bailly sont fortes ou faibles. En second lieu, le système est appuyé sur le fait signalé par Martin (1981) que les liaisons maintenues entre les mots sont quadrivalentes : indépendance **ID** (aucune liaison de dépendance immédiate entre les deux mots), interdépendance **IT** (les mots sont dépendants du même gouverneur), dépendance gauche **DG** (le premier mot est dépendant du deuxième) et dépendance droite **DD** (le deuxième mot est dépendant du premier). Son système est basé sur quatre marques relationnelles organisées en ordre décroissant de force :

**ID** indépendance.

Exemple : Monsieur Bailly, (ID) rappelez Monsieur SLIM jeudi.

**DG** dépendance gauche signale que, de deux groupes de mots celui de gauche est dominé par celui de droite. Comme c'est le cas entre nom sujet et verbe subséquent ou bien un adjectif et son nom.

Exemple 1 : La fille (DG) mange une pomme.

Exemple 2 : Le petit (DG) chien joue au ballon.

**DD** dépendance droite : réfère au cas inverse où le groupe droit est dominé par celui de gauche, comme c'est le cas entre le verbe et le complément subséquent ou bien un nom et son adjectif.

Exemple  : Le chien (DD) noir boit du lait.

**IT** interdépendance.

Exemple : Marie, (IT) Pierre (IT) et Jean boivent du lait.

L'auteur détermine que ces marques sont évocatrices de relations locales entre les éléments de l'axe syntagmatique : la construction prosodique qui sera produite ne contient pas la connaissance à première vue de la construction entière de la phrase. Cela mène à des systèmes pour le marquage automatique moins coûteux et plus réalistes, en particulier sur le plan de la procédure de prononciation des phrases lues qui dépendent beaucoup de sa manière personnelle de lire.

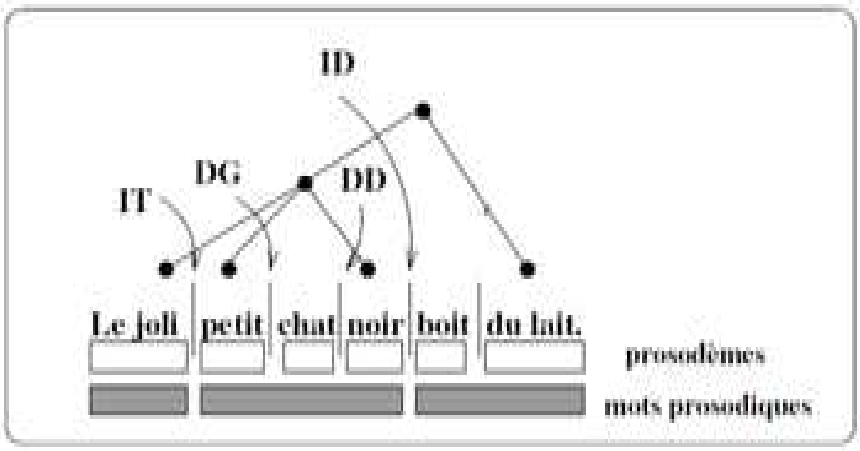



**Figure 47. Exemple d'arbre de dépendance et de marquage prosodique. D'après Bailly (1989, p. 140)**

# *ANNEXE C :* Transcriptions des dictées

## DICTEE_AUDE0000

**Avant l'orage.**

Des nuages violets passaient sur nos têtes, et la lumière bleuâtre baissait de minute en minute, comme celle d'une lampe qui meurt. Je n'avais pas peur, mais je sentais une inquiétude étrange, une angoisse profonde, animale. Les parfums de la colline étaient devenus des odeurs et montaient du sol, presque visibles.

Plusieurs lapins passèrent, aussi pressés que devant les chiens, puis des perdrix surgirent sans bruit du vallon, et se posèrent à trente pas sur notre gauche ...

Texte de Marcel Pagnol.

## DICTEE_AUDE0001

**Dans la jungle.**

Je continue ma marche, une marche toujours pénible.

J'enfonce dans la boue, je trébuche sur des branches.

J'ai mal au dos et aux reins.

Je commence à douter d'avoir pris la bonne direction.

Soudain, je crois distinguer une forme rouge à travers les feuilles: des Indiens !

Pourvu qu'ils ne me prennent pas pour un quelconque gibier.

Je lance un appel et m'approche d'eux.

Trois hommes et une femme sont en train de manger.

Ils me rassurent : je ne me suis pas perdu, mais ils m'engagent à revenir avec eux, car ils affirment que je n'ai aucune chance d'atteindre les sources du Tamouri.

Bien entendu, je refuse et nous nous séparons.

Dictée de Richard Chapelle.

## DICTEE_AUDE0002

**Drame en montagne.**

Toute la nuit, la tempête balaya la vallée avec la même violence.

Les petits ruisseaux devenaient des torrents furieux qui se précipitaient dans la rivière écumante.

Les scieurs de la vallée coururent à leurs scieries pour empêcher leur stock de bois d'être emporté.

Les gardes, postés auprès des ponts, devaient assister, impuissants, à la ruée du flot.

## DICTEE_AUDE0003

**Incendie dans les Landes.**

Le feu progressait sur un large front, et un épais nuage noir visible à plus de cent kilomètres à la ronde couvrait la forêt.

À Bordeaux, c'était la nuit en plein a (T., 2000)près-midi, tandis que parvenait jusqu'à Cognac l'odeur des résines en flammes.

De tout le sud-ouest, les sauveteurs affluaient.

**DICTEE_AUDE0004**

**La biche.**

La semaine passée, nous avons marché longtemps, sous un soleil matinal, dans le bois.

Au détour d'une allée déserte, nous nous arrêtâmes, nez à museau, devant une biche toute jeune qui s'arrête au lieu de s'enfuir.

Elle haletait d'émotion et ses jambes fines tremblaient;

mais ses longs yeux exprimaient plus d'embarras que de peur.

J'aurais voulu toucher ses oreilles pelucheuses et son doux museau de velours cotonneux.

Quand j'étendis la main, elle tourna le front d'un mouvement sauvage et disparut.

Texte de Colette.

**DICTEE_AUDE0005**

**La chasse du renard.**

Depuis deux mois, les petits renards sont nés.

Alors, commence pour le père une vie dure et dangereuse.

Au petit jour, il quitte sa retraite et passe la rivière pour venir rôder près d'une maison isolée, au milieu d'une sapinière.

Caché dans les fougères, il attend patiemment le passage des volailles imprudentes.

**DICTEE_AUDE0006**

**La petite fille et les oiseaux.**

Tous les oiseaux aimaient la petite fille parce qu'elle était bonne pour eux.

Ainsi un soir, elle put voir la belette qui rampait dans le sillon, vers le nid de l'alouette.

Elle se mit aussitôt à crier, à taper dans ses mains, pour mettre en fuite la bête au poil jaune.

Elle la chassa, la poursuivit au loin et les petites alouettes du nid furent sauvées.

Un autre jour, elle sauva le rossignol que la grande couleuvre allait dévorer.

Texte de Maurice Genevoix.

**DICTEE_AUDE0007**

**La rivière.**

Ton nom aimable et léger, je ne le connaissais pas, mais toi, je te connaissais si bien !

Tu viens d'une source vive, sous la roche, là-haut ; tu te hâtes de descendre et puis tu entres dans nos prés.
Alors, tu es nôtre pendant près d'une lieue ; à tous les détours de ta fuite, je t'ai accompagnée.
Je sais maintenant pourquoi tu coules d'abord dans une large prairie, pourquoi tu t'étrangles ensuite dans l'étroit vallon que dominent nos maisons.

## DICTEE_AUDE0008

**Le loup.**

Vers le milieu de l'hiver, les froids furent excessifs et les loups devinrent féroces.
Ils attaquaient même les paysans attardés, rôdaient la nuit autour des maisons, hurlaient du coucher du soleil à son lever et dépeuplaient les étables.
Et bientôt une rumeur circula.
On parlait d'un loup colossal, au pelage gris, presque blanc, qui avait mangé deux enfants, dévoré le bras d'une femme, étranglé tous les chiens de garde du pays, et qui pénétrait sans peur dans les enclos pour venir flairer sous la porte.
Une panique courut par toute la province.
Personne n'osait sortir dès que tombait le soir.
Les ténèbres semblaient hantées par l'image de cette bête.

## DICTEE_AUDE0009

**Le repas d'un paysan.**

La barrière de bois s'ouvrit; un homme entra, âgé de quarante ans peut-être, mais qui semblait vieux de soixante, marchant à grands pas alourdis par le poids de ses lourds sabots, pleins de paille.
L'homme entra dans la cuisine, ouvrit le buffet, prit un pain de six livres, en coupa soigneusement une tranche, recueillit dans le creux de sa main les miettes tombées sur la tablette et se les jeta dans la bouche pour ne rien perdre.
Puis il se mit à manger son pain lentement.
Texte de Guy de Maupassan

## DICTEE_AUDE0010

**Le vieux noyer et les oiseaux.**

Dans mon jardin, il y a un vieux noyer presque mort qui fait peur aux petits oiseaux.
Seul, un oiseau noir habite ses dernières feuilles.
Mais le reste du jardin est plein de jeunes arbres fleuris où nichent des oiseaux gais, vifs et de toutes les couleurs.
Et il semble que ces jeunes arbres se moquent du vieux noyer.

## DICTEE_AUDE0011

**Pauvre minet.**

Un de nos petits chats s'est endormi sur le foin de la grange.

Il ne s'est pas réveillé au retour des hommes.

Et ils ont déchargé la charrette sur lui sans faire attention.

On l'entend miauler de détresse, la voix étouffée par trois mètres de foin.

## DICTEE_AUDE0012

**Un coiffeur novice.**

Paul fut donc installé sur une chaise surmontée d'une petite caisse.

On lui mit la serviette au cou.

J'avais été chargé d'aller voler à la cuisine une casserole d'une taille convenable et, pour plus de sûreté, j'en avais pris deux.

Je lui mis la plus juste comme un chapeau et j'en tins le manche : pendant ce temps, avec une paire de ciseaux, mon père trancha les boucles au ras du bord ; ce fut fait avec une rapidité magique, mais le résultat ne fut pas satisfaisant, car, la casserole ôtée, la chevelure du patient apparut curieusement crénelée.

Texte de Marcel Pagnol.

## DICTEE_AUDE0013

**La petite école de campagne.**

Une seule maîtresse, jeune fille de vingt ans environ, s'occupait des enfants de tous âges.

La discipline était souple.

Après le déjeuner, si le temps le permettait, l'institutrice descendait avec les enfants jusqu'à la rivière.

Parfois, un inspecteur montait jusqu'à la petite école et interrogeait les enfants.

On voyait de loin venir sa voiture et, si l'on s'était attardé, on rentrait vite et chacun se trouvait à sa place quand l'inspecteur entrait.

L'inspection se passait bien ; la maîtresse était souriante ensuite, ce qui prouvait qu'elle avait reçu des compliments.

Texte d'André Maurois.

## DICTEE_AUDE0014

**La leçon d'histoire.**

Les classes d'histoire avaient toujours lieu l'après-midi.

On entendait parfois des soupirs de satisfaction et les bras se croisaient tout seuls sur les tables.

Le maître descendait de son estrade et venait s'asseoir parmi nous.

Le maître ne lisait pas : il parlait, il racontait.
Il lui arrivait de se lever, de tirer d'un coffre une carte qu'il suspendait au mur pour nous montrer les endroits où s'étaient passées les grandes choses qu'il nous enseignait.
Nous étions suspendus à ses lèvres.
Texte de Louis Guilloux.

# *ANNEXE D :* Les différentes versions des dictées

## DICTEE_AUDE0000

Avant l'orage PT.

Avant l'orage PT ##

Des nuages violets passaient sur nos têtes VRG.

Des nuages violets passaient sur nos têtes VRG ##

Et la lumière bleuâtre baissait de minute en minute VRG.

Et la lumière bleuâtre baissait de minute en minute VRG ##

Comme celle d'une lampe qui meurt PT.

Comme celle d'une lampe qui meurt PT ##

Je relis la phrase en entier.

Des nuages violets passaient sur nos têtes, et la lumière bleuâtre baissait de minute en minute, comme celle d'une lampe qui meurt ##

Je n'avais pas peur VRG.

Je n'avais pas peur VRG ##

Mais je sentais une inquiétude étrange VRG.

Mais je sentais une inquiétude étrange VRG ##

Une angoisse profonde VRG.

Une angoisse profonde VRG ##

Animale PT.

Animale PT ##

Je relis la phrase.

Je n'avais pas peur VRG mais je sentais une inquiétude étrange VRG une angoisse profonde VRG animale PT ##

Les parfums de la colline.

Les parfums de la colline ##

Étaient devenus des odeurs.

Étaient devenus des odeurs ##

Et montaient du sol VRG.

Et montaient du sol VRG ##

Presque visibles PT.

Presque visibles PT ##

Je relis la phrase.

Les parfums de la colline étaient devenus des odeurs et montaient du sol VRG presque visibles PT ##

Plusieurs lapins passèrent VRG.

Plusieurs lapins passèrent VRG ##.

Aussi pressés que devant les chiens VRG.

Aussi pressés que devant les chiens VRG ##

Puis des perdrix surgirent sans bruit du vallon VRG

Puis des perdrix surgirent sans bruit du vallon VRG ##

Et se posèrent à trente pas sur notre gauche TROISPT.

Et se posèrent à trente pas sur notre gauche TROISPT ##

Je relis la phrase en entier.

Plusieurs lapins passèrent VRG aussi pressés que devant les chiens VRG puis des perdrix surgirent sans bruit du vallon VRG et se posèrent à trente pas sur notre gauche TROISPT ##

Texte de Marcel Pagnol ##

**DICTEE_AUDE0001**

Dans la jungle

Dans la jungle ##

Je continue ma marche VRG

Je continue ma marche VRG

Une marche toujours pénible PT

Une marche toujours pénible PT ##

Je relis la phrase ##

Je continue ma marche VRG une marche toujours pénible PT ##

J'enfonce dans la boue VRG

J'enfonce dans la boue VRG

Je trébuche sur des branches

Je trébuche sur des branches PT ##

Je relis la phrase ##

J'enfonce dans la boue VRG je trébuche sur des branches PT ##

J'ai mal au dos et aux reins PT

J'ai mal au dos et aux reins PT ##

Je relis la phrase ##

J'ai mal au dos et aux reins PT ##

Je commence à douter d'avoir pris la bonne direction PT

Je commence à douter d'avoir pris la bonne direction PT ##

Je relis la phrase ##

Je commence à douter d'avoir pris la bonne direction PT ##

Soudain VRG

Soudain VRG

Je crois distinguer une forme rouge à travers les feuilles DEUXPT

Je crois distinguer une forme rouge à travers les feuilles DEUXPT ##

Je relis ##

Soudain VRG je crois distinguer une forme rouge à travers les feuilles DEUXPT ##

Des Indiens PTE

Des Indiens PTE

Pourvu qu'ils ne me prennent pas pour un quelconque gibier PT

Pourvu qu'ils ne me prennent pas pour un quelconque gibier PT ##

Je relis la phrase ##

Pourvu qu'ils ne me prennent pas pour un quelconque gibier PT ##

Je lance un appel et m'approche d'eux PT

Je lance un appel et m'approche d'eux PT ##

Je relis ##

Je lance un appel et m'approche d'eux PT ##

Trois hommes et une femme sont en train de manger PT

Trois hommes et une femme sont en train de manger PT ##

Je relis la phrase ##

Trois hommes et une femme sont en train de manger PT ##

Ils me rassurent DEUXPT

Ils me rassurent DEUXPT

Je ne me suis pas perdu VRG

Je ne me suis pas perdu VRG ##

Je relis ##

Ils me rassurent DEUXPT je ne me suis pas perdu VRG ##

Mais ils m'engagent à revenir avec eux VRG

Mais ils m'engagent à revenir avec eux VRG

Car ils affirment que je n'ai aucune chance d'atteindre les sources du Tamouri Tamouri TAMOURI TAMOURI PT ##

Je relis la phrase ##

Ils me rassurent DEUXPT je ne me suis pas perdu VRG mais ils m'engagent à revenir avec eux VRG car ils affirment que je n'ai aucune chance d'atteindre les sources du Tamouri TAMOURI PT ##

Bien entendu VRG

Bien entendu VRG

Je refuse et nous nous séparons PT

Je refuse et nous nous séparons PT ##

Je relis la phrase ##

Bien entendu VRG je refuse et nous nous séparons PTF ##

Dictée de Richard Chapelle ##

La dictée est désormais terminée. Si vous souhaitez une relecture, vous pouvez rejouer la dictée en appuyant sur le bouton lecture. Et vous pouvez aussi retrouver d'autres dictées sur le site la dictée PT fr. A bientôt ##

**DICTEE_AUDE0002**

Drame en montage.

Drame en montage ##

Toute la nuit VRG

Toute la nuit VRG

La tempête balaya la vallée

La tempête balaya la vallée ##

Avec la même violence PT

Avec la même violence PT ##

Je relis la phrase en entier.

Toute la nuit VRG la tempête balaya la vallée avec la même violence PT

Les petits ruisseaux

Les petits ruisseaux ##

Devenaient des torrents furieux

Devenaient des torrents furieux ##

Qui se précipitaient dans la rivière écumante PT

Qui se précipitaient dans la rivière écumante PT

Je relis la phrase en entier.
Les petits ruisseaux devenaient des torrents furieux qui se précipitaient dans la rivière écumante PT ##
Les scieurs de la vallée
Les scieurs de la vallée ##
Coururent à leurs scieries
Coururent à leurs scieries ##
Pour empêcher leur stock de bois d'être emporté PT
pour empêcher leur stock de bois d'être emporté PT ##
Je relis la phrase en entier.
Les scieurs de la vallée coururent à leurs scieries pour empêcher leur stock de bois d'être emporté PT ##
Les gardes VRG
Les gardes VRG ##
Postés auprès des ponts VRG
Postés auprès des ponts VRG ##
Devaient assister VRG
Devaient assister VRG ##
Impuissants VRG
Impuissants VRG ##
À la ruée du flot PTF
À la ruée du flot PTF ##
Je relis la phrase.
Les gardes VRG postés auprès des ponts VRG devaient assister VRG impuissants VRG à la ruée du flot PTF ##

**DICTEE_AUDE0003**

Incendie dans les Landes.
Incendie dans les Landes ##
Landes L majuscule A N D E S
L majuscule A N D E S ##
Le feu progressait sur un large front VRG
Le feu progressait sur un large front VRG ##
Et un épais nuage noir

Et un épais nuage noir ##

Visible à plus de cent kilomètres à la ronde

Visible à plus de cent kilomètres à la ronde ##

Couvrait la forêt PT

Couvrait la forêt PT ##

Je relis la phrase en entier.

Le feu progressait sur un large front VRG et un épais nuage noir visible à plus de cent kilomètres à la ronde couvrait la forêt PT ##

À Bordeaux VRG

Bordeaux B majuscule O R D E A U X

À Bordeaux VRG ##

C’était la nuit en plein après-midi VRG

C’était la nuit en plein après-midi VRG ##

Tandis que parvenait jusqu'à Cognac

Cognac C majuscule O G N A C

Tandis que parvenait jusqu'à Cognac C majuscule O G N A C ##

L’odeur des résines en flammes

L’odeur des résines en flammes PT ##

Je relis la phrase en entier.

À Bordeaux VRG c'était la nuit en plein après-midi VRG tandis que parvenait jusqu'à Cognac l'odeur des résines en flammes PT ##

De tout le sud-ouest VRG

De tout le sud-ouest VRG ##

Les sauveteurs affluaient PTF

Les sauveteurs affluaient PTF ##

Je relis la phrase.

De tout le sud-ouest VRG les sauveteurs affluaient PTF ##

**DICTEE_AUDE0004**

La biche

La biche ##

La semaine passée VRG

La semaine passée VRG ##

Nous avons marché longtemps VRG

Nous avons marché longtemps VRG ##
Sous un soleil matinal VRG
Sous un soleil matinal VRG ##
Dans le bois PT
Dans le bois PT ##
Je relis la phrase en entier ##
La semaine passée VRG nous avons marché longtemps VRG sous un soleil matinal VRG dans le bois PT ##
Au détour d'une allée déserte VRG
Au détour d'une allée déserte VRG ##
Nous nous arrêtâmes VRG
Nous nous arrêtâmes VRG ##
Nez à museau VRG
Nez à museau VRG ##
Devant une biche toute jeune
Devant une biche toute jeune ##
Qui s'arrête au lieu de s'enfuir PT
Qui s'arrête au lieu de s'enfuir PT ##
Je relis la phrase en entier:
Au détour d'une allée déserte VRG nous nous arrêtâmes VRG nez à museau VRG devant une biche toute jeune qui s'arrête au lieu de s'enfuir PT ##
Elle haletait d'émotion
Elle haletait d'émotion ##
Et ses jambes fines tremblaient PVRG
Et ses jambes fines tremblaient PVRG ##
Mais ses longs yeux
Mais ses longs yeux ##
Exprimaient plus d'embarras
Exprimaient plus d'embarras ##
Que de peur PT
Que de peur PT ##
Je relis la phrase en entier ##
Elle haletait d'émotion et ses jambes fines tremblaient PVRG mais ses longs yeux exprimaient plus d'embarras que de peur PT ##

J'aurais voulu toucher

J'aurais voulu toucher ##

Ses oreilles pelucheuses

Ses oreilles pelucheuses ##

Et son doux museau

Et son doux museau ##

De velours cotonneux PT

De velours cotonneux PT ##

Je relis la phrase en entier ##

J'aurais voulu toucher ses oreilles pelucheuses et son doux museau de velours cotonneux PT ##

Quand j'étendis la main VRG

Quand j'étendis la main VRG##

Elle tourna le front

Elle tourna le front ##

D’un mouvement sauvage

D’un mouvement sauvage ##

Et disparut PTF

Et disparut PTF ##

Je relis la phrase en entier ##

Quand j'étendis la main VRG elle tourna le front d'un mouvement sauvage et disparut PTF ##

Texte de Colette ##

**DICTEE_AUDE0005**

La chasse du renard.

La chasse du renard ##

Depuis deux mois VRG

Depuis deux mois VRG ##

Les petits renards sont nés PT

Les petits renards sont nés PT ##

Je relis la phrase ##

Depuis deux mois VRG les petits renards sont nés PT ##

Alors VRG

Alors VRG ##

Commence pour le père

Commence pour le père ##

Une vie dure et dangereuse PT

Une vie dure et dangereuse PT ##

Je répète la phrase ##

Alors VRG commence pour le père une vie dure et dangereuse PT ##

Au petit jour VRG

Au petit jour VRG ##

Il quitte sa retraite

Il quitte sa retraite ##

Et passe la rivière

Et passe la rivière ##

Pour venir rôder

Pour venir rôder ##

Près d'une maison isolée VRG

Près d'une maison isolée VRG ##

Au milieu d'une sapinière PT

Au milieu d'une sapinière PT ##

Je relis la phrase ##

Au petit jour VRG il quitte sa retraite et passe la rivière pour venir rôder près d'une maison isolée VRG au milieu d'une sapinière PT ##

Caché dans les fougères VRG

Caché dans les fougères VRG ##

Il attend patiemment

Il attend patiemment ##

Le passage des volailles imprudentes PT

Le passage des volailles imprudentes PTF ##

Je relis la phrase ##

Caché dans les fougères VRG il attend patiemment le passage des volailles imprudentes PTF ##

**DICTEE_AUDE0006**

La petite fille et les oiseaux.

La petite fille et les oiseaux ##

Tous les oiseaux aimaient la petite fille.
Tous les oiseaux aimaient la petite fille ##
Parce qu'elle était bonne pour eux PT.
Parce qu'elle était bonne pour eux PT ##
Je relis la phrase en entier ##
Tous les oiseaux aimaient la petite fille parce qu'elle était bonne pour eux PT ##
Ainsi un soir VRG
Ainsi un soir VRG ##
Elle put voir la belette.
Elle put voir la belette ##
Qui rampait dans le sillon VRG.
Qui rampait dans le sillon VRG ##
Vers le nid de l'alouette PT.
Vers le nid de l'alouette PT ##
Je relis la phrase en entier ##
Ainsi un soir VRG elle put voir la belette qui rampait dans le sillon VRG vers le nid de l'alouette PT ##
Elle se mit aussitôt à crier VRG.
Elle se mit aussitôt à crier VRG ##
À taper dans ses mains VRG.
À taper dans ses mains VRG ##
Pour mettre en fuite la bête au poil jaune PT.
Pour mettre en fuite la bête au poil jaune PT ##
Je relis la phrase en entier ##
Elle se mit aussitôt à crier VRG à taper dans ses mains VRG pour mettre en fuite la bête au poil jaune PT ##
Elle la chassa VRG.
Elle la chassa VRG ##
La poursuivit au loin.
La poursuivit au loin ##
Et les petites alouettes.
Et les petites alouettes ##
Du nid furent sauvées.
Du nid furent sauvées PT ##

Je relis la phrase en entier ##

Elle la chassa VRG la poursuivit au loin et les petites alouettes du nid furent sauvées PT ##

Un autre jour VRG.

Un autre jour VRG ##

Elle sauva le rossignol.

Elle sauva le rossignol ##

Que la grande couleuvre allait dévorer PTF.

Que la grande couleuvre allait dévorer PTF ##

Je relis la phrase en entier ##

Un autre jour VRG elle sauva le rossignol que la grande couleuvre allait dévorer PTF ##

Texte de Maurice Genevoix ##

**DICTEE_AUDE0007**

La rivière.

La rivière ##

Ton nom aimable et léger VRG.

Ton nom aimable et léger VRG ##

Je ne le connaissais pas VRG.

Je ne le connaissais pas VRG ##

Mais toi VRG.

Mais toi VRG ##

Je te connaissais si bien PTE.

Je te connaissais si bien PTE ##

Je relis la phrase en entier ##

Ton nom aimable et léger VRG je ne le connaissais pas VRG mais toi VRG je te connaissais si bien PTE ##

Tu viens d'une source vive VRG.

Tu viens d'une source vive VRG ##

Sous la roche VRG.

Sous la roche VRG ##

Là-haut PTVRG.

Là-haut PTVRG ##

Tu te hâtes de descendre.

Tu te hâtes de descendre ##

Et puis tu entres dans nos prés PT.
Et puis tu entres dans nos prés PT ##
Je relis la phrase ##
Tu viens d'une source vive VRG sous la roche VRG là-haut PTVRG tu te hâtes de descendre et puis tu entres dans nos prés PT ##
Alors VRG.
Alors VRG ##
Tu es nôtre pendant près d'une lieue PTVRG.
Tu es nôtre pendant près d'une lieue PTVRG ##
À tous les détours de ta fuite VRG.
À tous les détours de ta fuite VRG ##
Je t'ai accompagnée PT.
Je t'ai accompagnée PT ##
Je relis la phrase ##
Alors VRG tu es nôtre pendant près d'une lieue PTVRG à tous les détours de ta fuite VRG je t'ai accompagnée PT ##
Je sais maintenant.
Je sais maintenant ##
Pourquoi tu coules d'abord dans une large prairie VRG .
Pourquoi tu coules d'abord dans une large prairie VRG ##
Pourquoi tu t'étrangles ensuite.
Pourquoi tu t'étrangles ensuite ##
Dans l'étroit vallon que dominent nos maisons PTF.
Dans l'étroit vallon que dominent nos maisons PTF ##
Je relis la phrase ##
Je sais maintenant pourquoi tu coules d'abord dans une large prairie VRG pourquoi tu t'étrangles ensuite dans l'étroit vallon que dominent nos maisons PTF ##

## DICTEE_AUDE0008

Le loup ##
Vers le milieu de l'hiver VRG.
Vers le milieu de l'hiver VRG ##
Les froids furent excessifs.

Les froids furent excessifs ##
Et les loups devinrent féroces PT.
Et les loups devinrent féroces PT ##
Je relis la phrase en entier ##
Vers le milieu de l'hiver VRG les froids furent excessifs et les loups devinrent féroces PT ##
Ils attaquaient même les paysans attardés VRG.
Ils attaquaient même les paysans attardés VRG ##
Rôdaient la nuit autour des maisons VRG.
Rôdaient la nuit autour des maisons VRG ##
Hurlaient du coucher du soleil.
Hurlaient du coucher du soleil ##
À son lever et dépeuplaient les étables PT.
À son lever et dépeuplaient les étables PT ##
Je relis la phrase en entier ##
Ils attaquaient même les paysans attardés VRG rôdaient la nuit autour des maisons VRG hurlaient du coucher du soleil à son lever et dépeuplaient les étables PT ##
Et bientôt une rumeur circula PT.
Et bientôt une rumeur circula PT ##
Je relis.
Et bientôt une rumeur circula PT ##
On parlait d'un loup colossal VRG.
On parlait d'un loup colossal VRG ##
Au pelage gris VRG.
Au pelage gris VRG ##
Presque blanc VRG.
Presque blanc VRG ##
Qui avait mangé deux enfants VRG.
Qui avait mangé deux enfants VRG ##
Dévoré le bras d'une femme VRG.
Dévoré le bras d'une femme VRG ##
Étranglé tous les chiens de garde du pays VRG.
Étranglé tous les chiens de garde du pays VRG ##
Et qui pénétrait sans peur dans les enclos.
Et qui pénétrait sans peur dans les enclos ##

Pour venir flairer sous la porte PT.

Pour venir flairer sous la porte PT ##

Je relis la phrase en entier ##

On parlait d'un loup colossal VRG au pelage gris VRG presque blanc VRG qui avait mangé deux enfants VRG dévoré le bras d'une femme VRG étranglé tous les chiens de garde du pays VRG et qui pénétrait sans peur dans les enclos pour venir flairer sous la porte PT ##

Une panique courut par toute la province PT .

Une panique courut par toute la province PT ##

Je relis ##

Une panique courut par toute la province PT ##

Personne n'osait sortir dès que tombait le soir PT.

Personne n'osait sortir dès que tombait le soir PT ##

Je relis ##

Personne n'osait sortir dès que tombait le soir PT ##

Les ténèbres semblaient hantées par l'image de cette bête PTF.

Les ténèbres semblaient hantées par l'image de cette bête PTF ##

Je relis.

Les ténèbres semblaient hantées par l'image de cette bête PTF ##

**DICTEE_AUDE0009**

Le repas d'un paysan.

Le repas d'un paysan ##

La barrière de bois s'ouvrit PTVRG.

La barrière de bois s'ouvrit PTVRG ##

Un homme entra VRG.

Un homme entra VRG ##

Âgé de quarante ans peut-être VRG.

Âgé de quarante ans peut-être VRG ##

Mais qui semblait vieux de soixante VRG.

Mais qui semblait vieux de soixante VRG ##

Marchant à grands pas alourdis.

Marchant à grands pas alourdis ##

Par le poids de ses lourds sabots VRG

Par le poids de ses lourds sabots VRG ##
Pleins de paille PT.
Pleins de paille PT ##
Je relis la phrase en entier ##
La barrière de bois s'ouvrit PTVRG un homme entra VRG âgé de quarante ans peut-être VRG mais qui semblait vieux de soixante VRG marchant à grands pas alourdis par le poids de ses lourds sabots VRG pleins de paille PT ##
L'homme entra dans la cuisine VRG.
L'homme entra dans la cuisine VRG ##
Ouvrit le buffet VRG.
Ouvrit le buffet VRG ##
Prit un pain de six livres VRG.
Prit un pain de six livres VRG ##
En coupa soigneusement une tranche VRG.
En coupa soigneusement une tranche VRG ##
Recueillit dans le creux de sa main.
Recueillit dans le creux de sa main ##
Les miettes tombées sur la tablette.
Les miettes tombées sur la tablette ##
Et se les jeta dans la bouche pour ne rien perdre PT.
Et se les jeta dans la bouche pour ne rien perdre PT ##
Je relis la phrase en entier ##
L'homme entra dans la cuisine VRG ouvrit le buffet VRG prit un pain de six livres VRG en coupa soigneusement une tranche VRG recueillit dans le creux de sa main les miettes tombées sur la tablette et se les jeta dans la bouche pour ne rien perdre PT ##
Puis il se mit à manger son pain lentement PTF.
Puis il se mit à manger son pain lentement PTF ##
Je relis la phrase ##
Puis il se mit à manger son pain lentement PTF ##
Texte de Guy de Maupassant ##

**DICTEE_AUDE0010**

Le vieux noyer et les oiseaux.
Le vieux noyer et les oiseaux ##

Dans mon jardin VRG.
Dans mon jardin VRG ##
Il y a un vieux noyer presque mort.
Il y a un vieux noyer presque mort ##
Qui fait peur aux petits oiseaux PT
Qui fait peur aux petits oiseaux PT ##
Je relis la phrase en entier ##
Dans mon jardin VRG il y a un vieux noyer presque mort qui fait peur aux petits oiseaux PT ##
Seul VRG.
Seul VRG ##
Un oiseau noir habite.
Un oiseau noir habite ##
Ses dernières feuilles PT.
Ses dernières feuilles PT ##
Je relis la phrase en entier ##
Seul VRG un oiseau noir habite ses dernières feuilles PT ##
Mais le reste du jardin.
Mais le reste du jardin ##
Est plein de jeunes arbres fleuris
Est plein de jeunes arbres fleuris ##
Où nichent des oiseaux gais VRG
Où nichent des oiseaux gais VRG ##
Vifs et de toutes les couleurs PT
Vifs et de toutes les couleurs PT ##
Je relis la phrase en entier ##
Mais le reste du jardin est plein de jeunes arbres fleuris où nichent des oiseaux gais VRG vifs et de toutes les couleurs PT ##
Et il semble.
Et il semble ##
Que ces jeunes arbres.
Que ces jeunes arbres ##
Se moquent du vieux noyer PTF.
Se moquent du vieux noyer PTF ##

Je relis la phrase ##

Et il semble que ces jeunes arbres se moquent du vieux noyer PT ##

**DICTEE_AUDE0011**

Pauvre minet

Pauvre minet ##

Un de nos petits chats.

Un de nos petits chats ##

S’est endormi.

S’est endormi ##

Sur le foin de la grange PT.

Sur le foin de la grange PT ##

Je relis la phrase en entier.

Un de nos petits chats s'est endormi sur le foin de la grange PT ##

Il ne s'est pas réveillé.

Il ne s'est pas réveillé ##

Au retour des hommes PT.

Au retour des hommes PT ##

Je relis la phrase.

Il ne s'est pas réveillé au retour des hommes PT ##

Et ils ont déchargé.

Et ils ont déchargé ##

La charrette sur lui.

La charrette sur lui ##

Sans faire attention PT.

Sans faire attention PT ##

Je relis la phrase.

Et ils ont déchargé la charrette sur lui sans faire attention PT ##

On l'entend miauler de détresse VRG.

On l'entend miauler de détresse VRG ##

La voix étouffée par trois mètres de foin PTF.

La voix étouffée par trois mètres de foin PTF ##

Je relis la phrase ##

On l'entend miauler de détresse VRG la voix étouffée par trois mètres de foin PTF ##

**DICTEE_AUDE0012**

Un coiffeur novice.

Un coiffeur novice ##

Paul fut donc installé.

Paul fut donc installé ##

Sur une chaise surmontée d'une petite caisse PT.

Sur une chaise surmontée d'une petite caisse PT ##

Je relis la phrase en entier.

Paul fut donc installé sur une chaise surmontée d'une petite caisse PT ##

On lui mit la serviette au cou PT.

On lui mit la serviette au cou PT.

Je relis.

On lui mit la serviette au cou PT ##

J'avais été chargé d'aller voler à la cuisine.

J'avais été chargé d'aller voler à la cuisine ##

Une casserole d'une taille convenable et VRG.

Une casserole d'une taille convenable et VRG ##

Pour plus de sûreté VRG.

Pour plus de sûreté VRG ##

J’en avais pris deux PT.

J’en avais pris deux PT ##

Je relis la phrase.

J'avais été chargé d'aller voler à la cuisine une casserole d'une taille convenable et VRG pour plus de sûreté VRG j'en avais pris deux PT ##

Je lui mis la plus juste.

Je lui mis la plus juste ##

Comme un chapeau et j'en tins le manche DEUXPT.

Comme un chapeau et j'en tins le manche DEUXPT ##

Pendant ce temps VRG.

Pendant ce temps VRG ##

Avec une paire de ciseaux VRG

Avec une paire de ciseaux VRG ##

Mon père trancha les boucles.

Mon père trancha les boucles ##

Au ras du bord PTVRG.

Au ras du bord PTVRG ##

Ce fut fait avec une rapidité magique VRG.

Ce fut fait avec une rapidité magique VRG ##

Mais le résultat ne fut pas satisfaisant VRG.

Mais le résultat ne fut pas satisfaisant VRG ##

Car VRG.

Car VRG ##

La casserole ôtée VRG.

La casserole ôtée VRG ##

La chevelure du patient apparut curieusement crénelée PTF.

La chevelure du patient apparut curieusement crénelée PTF ##

Je relis la phrase tout entier ##

Je lui mis la plus juste comme un chapeau et j'en tins le manche DEUXPT pendant ce temps VRG avec une paire de ciseaux VRG mon père trancha les boucles au ras du bord PTVRG ce fut fait avec une rapidité magique VRG mais le résultat ne fut pas satisfaisant VRG car VRG la casserole ôtée VRG la chevelure du patient apparut curieusement crénelée PTF ##

Texte de Marcel Pagnol ##

**DICTEE_AUDE0013**

La petite école de campagne

La petite école de campagne ##

Une seule maîtresse VRG

Une seule maîtresse VRG ##

Jeune fille de vingt ans environ VRG

Jeune fille de vingt ans environ VRG ##

S’occupait des enfants de tous âges PT

S’occupait des enfants de tous âges PT ##

Je relis la phrase en entier.

Une seule maîtresse VRG jeune fille de vingt ans environ VRG s'occupait des enfants de tous âges PT ##

La discipline était souple PT

La discipline était souple PT ##

Je relis.
La discipline était souple PT ##
Après le déjeuner VRG
Après le déjeuner VRG ##
Si le temps le permettait VRG
Si le temps le permettait VRG ##
L’institutrice descendait avec les enfants
L’institutrice descendait avec les enfants ##
Jusqu’à la rivière PT
Jusqu’à la rivière PT ##
Je relis la phrase ##
Après le déjeuner VRG si le temps le permettait VRG l'institutrice descendait avec les enfants jusqu'à la rivière PT ##
Parfois VRG
Parfois VRG ##
Un inspecteur montait jusqu'à la petite école
Un inspecteur montait jusqu'à la petite école ##
Et interrogeait les enfants PT
Et interrogeait les enfants PT ##
Je relis la phrase.
Parfois VRG un inspecteur montait jusqu'à la petite école et interrogeait les enfants PT ##
On voyait de loin venir sa voiture et VRG
On voyait de loin venir sa voiture et VRG ##
Si l'on s'était attardé VRG
Si l'on s'était attardé VRG ##
On rentrait vite
On rentrait vite ##
Et chacun se trouvait à sa place
Et chacun se trouvait à sa place ##
Quand l'inspecteur entrait PT
Quand l'inspecteur entrait PT ##
Je relis la phrase en entier ##
On voyait de loin venir sa voiture et VRG si l'on s'était attardé VRG on rentrait vite et chacun se trouvait à sa place quand l'inspecteur entrait PT ##

L'inspection se passait bien PTVRG

L'inspection se passait bien PTVRG ##

La maîtresse était souriante ensuite VRG

La maîtresse était souriante ensuite VRG ##

Ce qui prouvait qu'elle avait reçu des compliments PTF

Ce qui prouvait qu'elle avait reçu des compliments PTF ##

Je relis la phrase en entier ##

L'inspection se passait bien PTVRG la maîtresse était souriante ensuite VRG ce qui prouvait qu'elle avait reçu des compliments PTF ##

Texte d'André Maurois.

**DICTEE_AUDE0014**

La leçon d’histoire.

La leçon d'histoire ##

Les classes d'histoire avaient toujours lieu l'après-midi PT ##

Les classes d'histoire avaient toujours lieu l'après-midi PT ##

Je relis la phrase.

Les classes d'histoire avaient toujours lieu l'après-midi PT ##

On entendait parfois.

On entendait parfois ##

Des soupirs de satisfaction.

Des soupirs de satisfaction ##

Et les bras.

Et les bras ##

Se croisaient tout seuls sur les tables PT.

Se croisaient tout seuls sur les tables PT ##

Je relis la phrase en entier ##

On entendait parfois des soupirs de satisfaction et les bras se croisaient tout seuls sur les tables PT ##

Le maître descendait de son estrade.

Le maître descendait de son estrade ##

Et venait s'asseoir parmi nous PT.

Et venait s'asseoir parmi nous PT ##

Je relis la phrase.

Le maître descendait de son estrade et venait s'asseoir parmi nous PT ##

Le maître ne lisait pas DEUXPT.

Le maître ne lisait pas DEUXPT ##

Il parlait VRG.

Il parlait VRG ##

Il racontait PT.

Il racontait PT ##

Je relis la phrase ##

Le maître ne lisait pas DEUXPT il parlait VRG il racontait PT ##

Il lui arrivait de se lever VRG.

Il lui arrivait de se lever VRG ##

De tirer d'un coffre.

De tirer d'un coffre ##

Une carte qu'il suspendait au mur.

Une carte qu'il suspendait au mur ##

Pour nous montrer.

Pour nous montrer ##

Les endroits où s'étaient passées.

Les endroits où s'étaient passées ##

Les grandes choses qu'il nous enseignait PT.

Les grandes choses qu'il nous enseignait PT ##

Je relis la phrase en entier ##

Il lui arrivait de se lever VRG de tirer d'un coffre une carte qu'il suspendait au mur pour nous montrer les endroits où s'étaient passées les grandes choses qu'il nous enseignait PT ##

Nous étions suspendus à ses lèvres PTF.

Nous étions suspendus à ses lèvres PTF ##

Je relis la phrase.

Nous étions suspendus à ses lèvres PTF.

Texte de Louis Guilloux.

# *ANNEXE E :* Quelques expressions régulières utilisées

- **Accent circonflexe**

```
$ligne=~s/â/a&3/gi;
$ligne=~s/î/i&3/gi;
 $ligne=~s/û/u&3/gi;
$ligne=~s/ê/e&3/gi;
$ligne=~s/ô/o&3/gi;
```

- **Accent grave**

```
 $ligne=~s/à/a&2/gi;
 $ligne=~s/è/e&2/gi;
$ligne=~s/ù/u&2/gi;
```

- **Accent aigu**

```
$ligne=~s/á/a&6/gi;
$ligne=~s/é/e&6/gi;
$ligne=~s/ú/u&6/gi;
$ligne=~s/í/i&6/gi;
```

- **Recherche pattern Verbe+Verbe => Aux+Verbe**

```
 $ligne=~s/(VER)(\s\([A-Za-z0-9&\-]+\)\s)(VER\s+\([A-Za-z0-9&\-]+\)\s)/Aux$2$3/g;
```

- **Le temps conjugué du verbe est simple et le sujet est un Pronom Personnel; recherche du pattern syntagme+verbe+syntagme => syntagme verbe :DD syntagme**

```
$ligne=~s/^(Pp\s)(.*?)(Vrb\s+\([A-Za-z0-9&\-]+\))(.*?)(Vrg\s\(,\))/Gn($1$2
Gn($3<Gn;grp=$DD>($4$5)))/g;
$ligne=~s/^(Cco\s)(.*?)(Pp\s)(.*?)(Vrb\s+\([A-Za-z0-9&\-]+\))(.*?)(Vrg\s\(,\))/Gn($1$2$3$4
Gn($5<Gn;grp=$DD>($6$7)))/g;
 $ligne=~s/^(Pp\s)(.*?)(Vrb\s+\([A-Za-z0-9&\-]+\))(.*?)(Vrg\s\(,\))/Gn($1$2
Gn($3<Gn;grp=$DD>($4$5)))/g;
```

- **Le sujet est un NOM; recherche du pattern syntagme+verbe+syntagme => syntagme :DG verbe :DD syntagme**

```
 $ligne=~s/(Pre\s)(.*?)(\sNom\s)(.*?)(Vrb\s+\([A-Za-z0-9&\-
]+\))(.*?)(Vrg\s\(,\))/Gn($1$2$3$4)<Gn;grp=$DG>($5<Gn;grp=$DD>($6$7))/g;
```

- **Le sujet est un NOM; recherche du pattern syntagme+verbe+syntagme => syntagme :DG verbe :DD syntagme**

$ligne=~s/(\sDet\s)(.*?)(\sNom\s)(.*?)(Vrb\s+\([A-Za-z0-9&\-]+\)\s)(.*?)(Vrg\s\(,\))/Gn($1$2$3$4)<Gn;grp=$DG>($5<Gn;grp=$DD>($6$7))/g;

- **Suppression de DD dans le contexte d'un verbe + ponctuation (virgule)**

$ligne=~s/(.*?)(Vrb\s+\([A-Za-z0-9&\-]+\)\s*?)(<Gn;grp=:DD>\s*?\()(\s+Vrg\s+\(,\))(\))/$1$2$4/g;

- **Suppression de DD dans le contexte d'un verbe + ponctuation (un point)**

$ligne=~s/(.*?)(Vrb\s+\([A-Za-z0-9&\-]+\)\s*?)(<Gn;grp=:DD>\s*?\()(\sPt\s\(\.\))(\))/$1$2$4/g;

$ligne=~s/(.*?)(Vrb\s+\([A-Za-z0-9&\-]+\)\s*?)(<Gn;grp=:DD>\s*?\()(\sPt\s\(\.\))(\))/$1$2$4/g;

- **Recherche du pattern ADJ+NOM => ADJ : DG NOM**

$ligne=~s/(Adq\s+\([A-Za-z0-9&\-]+\)\s)(Nom\s+\([A-Za-z0-9&]+\)\s)/$1<Gadj;grp=$DG>($2)/g;

- **Recherche du pattern NOM+ADJ => NOM : DD ADJ**

$ligne=~s/(Nom\s+\([A-Za-z0-9&\-]+\)\s)(Adq\s+\([A-Za-z0-9&]+\)\s)/$1<Gadj;grp=$DD>($2)/g;

- **Recherche du pattern Pre +NOM PROPRE=> Pre :DD NOM PROPRE**

$ligne=~s/(Pre\s\(DE\)\s)(Npr\s\([A-Za-z0-9&\-]+\)\s)(Npr\s+\([A-Za-z0-9&]+\))/<Gn;grp=$DD>($1$2$3)/g;

- **Recherche du pattern NOM PROPRE+NOM PROPRE=> NOM PROPRE :DD NOM PROPRE**

$ligne=~s/(Npr\s\([A-Za-z0-9&\-]+\)\s)(Npr\s\([A-Za-z0-9&\-]+\))/$1<Gn;grp=$DD>($2)/g;

- **Recherche du pattern Pre (DE|DU)+Nom => Pre :DD Nom**

$ligne=~s/(Det\s+\([A-Za-z0-9&\-]+\)\s+)(Nom\s+\([A-Za-z0-9&\-]+\)\s+)(Pre\s+\(DE\)\s+)(Det\s+\([A-Za-z0-9&\-]+\)\s+)(Nom\s+\([A-Za-z0-9&\-]+\)\s+)/$1$2<Gn;grp=$DD>($3$4$5)/g;

- **Recherche du pattern Pre+Nom+Pre+Nom => Pre :DD Nom**

$ligne=~s/(Pre\s+\([A-Za-z0-9&\-]+\)\s+)(Nom\s+\([A-Za-z0-9&\-]+\)\s+)(Pre\s+\([A-Za-z0-9&\-]+\)\s+)(Nom\s+\([A-Za-z0-9&\-]+\)\s+)/Gn($1$2<Gn;grp=$IT>($3$4))/g;

## *ANNEXE F :* Systéme de marquage de Bailly appliqué aux dictées

**DICTEE_AUDE0000**

```
MACRO(#DC)=attitude=10
MACRO(#CT)=attitude=17
MACRO(#DD)=grp=2
MACRO(#DG)=grp=3
MACRO(#IT)=grp=4
MACRO(#ID)=grp=5

##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEa/"
>
Gn (
	Pre(AVANT)
	Det(L\\\')
	Nom(ORAGE)
 )
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEa/"
>
Gn(
	Gn (
		Pre(DES)
		Nom(NUAGES)
		<Gadj;#DD>(
			Adq(VIOLETS)
		)
	)
	<Gv;#DG>(
		Vrb(PASSAIENT)
		<Gn;#DD>(
			Pre (SUR)
			Det(NOS)
			Nom(TE&3TES)
			Vrg(\,)
		)
	)
)
<Gn;#ID>(
	Cco(ET)
	Gn (
		Det (LA)
		Nom(LUMIE&2RE)
		<Gadj;#DD>(
			Adq(BLEUA&3TRE)
		)
```

```
        )
        <Gv;#DG>(
                Vrb(BAISSAIT)
                <Gn;#DD>(
                        Gn(
                                Pre(DE)
                                Nom(MINUTE)
                        )
                        <Gn;#IT>(
                                Pre(EN)
                                Nom(MINUTE)
                        )
                )
        )
        Vrg(\,)
)
<Gn;#ID>(
        Cco(COMME)
        Pp(CELLE)
        <Gp;#DD> (
                Pre(D\\\')
                Det(UNE)
                Nom(LAMPE)
                <Gn;#DD>(
                        Ccs(QUI)
                        Vrb(MEURT)
                )
        )
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEa/"
>
Gv (
        Pp(JE)
        Ne(N\\\')
        Vrb(AVAIS)
        <Gn;#DD>(
                Ne(PAS)
                Nom(PEUR)
                Vrg(\,)
        )
)
<Gn;#ID>(
        Cco(MAIS)
        Gv (
                Pp(JE)
                Vrb(SENTAIS)
                <Gn;#DD>(
                        Det (UNE)
```

```
                        Nom(INQUIE&6TUDE)
                        <Gadj;#DD>(
                                Adq(E&6TRANGE)
                        )
                        Vrg(\,)
                )
                <Gn;#ID>(
                        Det(UNE)
                        Nom(ANGOISSE)
                        <Gadj;#DD>(
                                Adq(PROFONDE)
                                Vrg(\,)
                        )
                        <Gadj;#ID>(
                                Adq(ANIMALE)
                        )
                )
        )
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEa/"
>
Gn (
        Det(LES)
        Nom(PARFUMS)
        Gn (
                Pre(DE)
                Det(LA)
                Nom(COLLINE)
        )
)
<Gv;#DG>(
        Aux(E&6TAIENT)
        Pps(DEVENUS)
        <Gn;#DD>(
                Pre (DES)
                Nom(ODEURS)
        )
)
<Gv;#DG>(
        Cco(ET)
        Vrb(MONTAIENT)
        <Gn;#DD>(
                Pre(DU)
                Nom(SOL)
        )
        Vrg(\,)
)
<Gn;#ID>(
```

```
        Adv(PRESQUE)
        <Gadj;#DG>(
                Adq(VISIBLES)
        )
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEa/"
>
Gn (
        Det(PLUSIEURS)
        Nom(LAPINS)
)
<Gv;#DG>(
        Vrb(PASSE&2RENT)
        Vrg(\,)
)
<Gn;#ID>(
        Adv(AUSSI)
        Pps(PRESSE&6S)
        <Gn;#DD>(
                Ccs(QUE)
                Pre(DEVANT)
                Det(LES)
                Nom(CHIENS)
        )
        Vrg(\,)
)
<Gn;#ID>(
        Cco(PUIS)
        <Gn;#IT>(
                Det(DES)
                Nom(PERDRIX)
        )
        <Gv;#DG>(
                Vrb(SURGIRENT)
                <Gn;#DD>(
                        Pre (SANS)
                        Nom(BRUIT)
                )
                <Gn;#IT>(
                        Pre(DU)
                        Nom(VALLON)
                )
        )
        Vrg(\,)
)
<Gn;#ID>(
        Cco(ET)
        Pp(SE)
```

```
        Vrb(POSE&2RENT)
        <Gn;#DD>(
                Pre(A&2)
                Num(TRENTE)
                Nom(PAS)
        )
        <Gn;#IT>(
                Pre(SUR)
                Det(NOTRE)
                Nom(GAUCHE)
        )
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEa/"
>
Gn (
        Nom(TEXTE)
        <Gn;#DD>(
                Pre(DE)
                Npr(MARCEL)
                <Gn;#DG>(
                        Npr(PAGNOL)
                )
        )
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/
">
Gn (
        Pre(AVANT)
        Det(L\\\')
        Nom(ORAGE)
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/
">
Gn (
        Pt(POINT)
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
Gn (
        Pre(AVANT)
        Det(L\\\')
        Nom(ORAGE)
```

```
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
Gn (
	Pt(POINT)
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/"
>
Gn(
	Gn (
		Pre(DES)
		Nom(NUAGES)
		<Gadj;#DD>(
			Adq(VIOLETS)
		)
	)
	<Gv;#DG>(
		Vrb(PASSAIENT)
		<Gn;#DD>(
			Pre (SUR)
			Det(NOS)
			Nom(TE&3TES)
		)
	)
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/"
>
Gn (
	Vrg(VIRGULE)
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
Gn(
	Gn (
		Pre(DES)
		Nom(NUAGES)
		<Gadj;#DD>(
			Adq(VIOLETS)
		)
	)
	<Gv;#DG>(
		Vrb(PASSAIENT)
```

```
                <Gn;#DD>(
                        Pre (SUR)
                        Det(NOS)
                        Nom(TE&3TES)
                )
        )
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
Gn (
        Vrg(VIRGULE)
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/"
>
Gn(
        Cco(ET)
        Gn (
                Det (LA)
                Nom(LUMIE&2RE)
                <Gadj;#DD>(
                        Adq(BLEUA&3TRE)
                )
        )
        <Gv;#DG>(
                Vrb(BAISSAIT)
                <Gn;#DD>(
                        Gn(
                                Pre(DE)
                                Nom(MINUTE)
                        )
                        <Gn;#IT>(
                                Pre(EN)
                                Nom(MINUTE)
                        )
                )
        )
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/"
>
Gn (
        Vrg(VIRGULE)
)
##
##
```

```
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
Gn(
	Cco(ET)
	Gn (
		Det (LA)
		Nom(LUMIE&2RE)
		<Gadj;#DD>(
			Adq(BLEUA&3TRE)
		)
	)
	<Gv;#DG>(
		Vrb(BAISSAIT)
		<Gn;#DD>(
			Gn(
				Pre(DE)
				Nom(MINUTE)
			)
			<Gn;#IT>(
				Pre(EN)
				Nom(MINUTE)
			)
		)
	)
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
Gn (
	Vrg(VIRGULE)
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/
">
<Gn;#ID>(
	Cco(COMME)
	Pp(CELLE)
	Gp (
		Pre(D\\\')
		Det(UNE)
		Nom(LAMPE)
		<Gn;#DD>(
			Ccs(QUI)
			Vrb(MEURT)
		)
	)
)
##
```

```
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/
">
Gn (
	Pt(POINT)
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
<Gn;#ID>(
	Cco(COMME)
	Pp(CELLE)
	Gp (
		Pre(D\\\')
		Det(UNE)
		Nom(LAMPE)
		<Gn;#DG>(
			Ccs(QUI)
			Vrb(MEURT)
		)
	)
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
Gn (
	Pt(POINT)
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDE/"
>
Gv (
	Pp(JE)
	Vrb(RELIS)
	<Gn;#DD> (
		Det(LA)
		Nom(PHRASE)
	)
	<Gn;#IT> (
		Pre(EN)
		Nom(ENTIER)
	)
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEd/
">
```

```
Gn(
	Gn (
		Pre(DES)
		Nom(NUAGES)
		<Gadj;#DD>(
			Adq(VIOLETS)
		)
	)
	<Gv;#DG>(
		Vrb(PASSAIENT)
		<Gn;#DD>(
			Pre (SUR)
			Det(NOS)
			Nom(TE&3TES)
			Vrg(\,)
		)
	)
)
<Gn;#ID>(
	Cco(ET)
	Gn (
		Det (LA)
		Nom(LUMIE&2RE)
		<Gadj;#DD>(
			Adq(BLEUA&3TRE)
		)
	)
	<Gv;#DG>(
		Vrb(BAISSAIT)
		<Gn;#DD>(
			Gn(
				Pre(DE)
				Nom(MINUTE)
			)
			<Gn;#IT>(
				Pre(EN)
				Nom(MINUTE)
			)
		)
	)
	Vrg(\,)
)
<Gn;#ID>(
	Cco(COMME)
	Pp(CELLE)
	Gp (
		Pre(D\\\')
		Det(UNE)
		Nom(LAMPE)
		<Gn;#DD>(
```

```
                        Ccs(QUI)
                        Vrb(MEURT)
                )
        )
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEd/
">
Gn(
        Pt(POINT)
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/"
>
Gv (
        Pp(JE)
        Ne(N\\\')
        Vrb(AVAIS)
        <Gn;#DD>(
                Pas(PAS)
                Nom(PEUR)
        )
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/"
>
Gn (
        Vrg(VIRGULE)
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
Gv (
        Pp(JE)
        Ne(N\\\')
        Vrb(AVAIS)
        <Gn;#DD>(
                Pas(PAS)
                Nom(PEUR)
        )
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
Gn (
```

```
        Vrg(VIRGULE)
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/"
>
<Gn;#ID>(
        Cco(MAIS)
        Gv (
                Pp(JE)
                Vrb(SENTAIS)
                <Gn;#DD>(
                        Det (UNE)
                        Nom(INQUIE&6TUDE)
                        <Gadj;#DD>(
                                Adq(E&6TRANGE)
                        )
                )
        )
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/"
>
Gn(
        Vrg(VIRGULE)
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
<Gn;#ID>(
        Cco(MAIS)
        Gv (
                Pp(JE)
                Vrb(SENTAIS)
                <Gn;#DD>(
                        Det (UNE)
                        Nom(INQUIE&6TUDE)
                        <Gadj;#DD>(
                                Adq(E&6TRANGE)
                        )
                )
        )
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
Gn(
```

```
        Vrg(VIRGULE)
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/"
>
<Gn;#ID> (
        Det(UNE)
        Nom(ANGOISSE)
        <Gadj;#DD>(
                Adq(PROFONDE)
        )
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/"
>
Gn (
        Vrg(VIRGULE)
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
<Gn;#ID> (
        Det(UNE)
        Nom(ANGOISSE)
        <Gadj;#DD>(
                Adq(PROFONDE)
        )
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
Gn (
        Vrg(VIRGULE)
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/
">
<Gn;#ID> (
        Adq(ANIMALE)
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/
">
Gn (
```

```
        Pt(POINT)
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
<Gn;#ID> (
        Adq(ANIMALE)
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
Gn (
        Pt(POINT)
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDE/"
>
Gv (
        Pp(JE)
        Vrb(RELIS)
        <Gn;#DD> (
                Det(LA)
                Nom(PHRASE)
        )
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEd/
">
Gv (
        Pp(JE)
        Ne(N\\\')
        Vrb(AVAIS)
        <Gn;#DD>(
                Pas(PAS)
                Nom(PEUR)
                Vrg(\,)
        )
)
<Gn;#ID>(
        Cco(MAIS)
        Gv (
                Pp(JE)
                Vrb(SENTAIS)
                <Gn;#DD>(
                        Det (UNE)
                        Nom(INQUIE&6TUDE)
```

```
                        <Gadj;#DD>(
                                Adq(E&6TRANGE)
                        )
                        Vrg(\,)
                )
                <Gn;#ID>(
                        Det(UNE)
                        Nom(ANGOISSE)
                        <Gadj;#DD>(
                                Adq(PROFONDE)
                                Vrg(\,)
                        )
                        <Gadj;#ID>(
                                Adq(ANIMALE)
                        )
                )
        )
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEd/
">
Gn(
        Pt(POINT)
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/"
>
Gn (
        Det(LES)
        Nom(PARFUMS)
        <Gn;#DD> (
                Pre(DE)
                Det(LA)
                Nom(COLLINE)
        )
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
Gn (
        Det(LES)
        Nom(PARFUMS)
        <Gn;#DD> (
                Pre(DE)
                Det(LA)
                Nom(COLLINE)
        )
```

```
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/"
>
<Gv;#DG>(
        Aux(E&6TAIENT)
        Pps(DEVENUS)
        <Gn;#DD>(
                Pre (DES)
                Nom(ODEURS)
        )
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
<Gv;#DG>(
        Aux(E&6TAIENT)
        Pps(DEVENUS)
        <Gn;#DD>(
                Pre (DES)
                Nom(ODEURS)
        )
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/"
>
<Gv;#DG>(
        Cco(ET)
        Vrb(MONTAIENT)
        <Gn;#DD>(
                Pre(DU)
                Nom(SOL)
        )
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/"
>
Gn(
        Vrg(VIRGULE)
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
Gv(
        Cco(ET)
```

```
        Vrb(MONTAIENT)
        <Gn;#DD>(
                Pre(DU)
                Nom(SOL)
        )
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
Gn(
        Vrg(VIRGULE)
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/"
>
Gn(
        Adv(PRESQUE)
        <Gadj;#DG>(
                Adq(VISIBLES)
        )
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/"
>
Gn(
        Pt(POINT)
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
Gn(
        Adv(PRESQUE)
        <Gadj;#DG>(
                Adq(VISIBLES)
        )
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
Gn(
        Pt(POINT)
)
##
##
```

```
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDE/"
>
Gv (
	Pp(JE)
	Vrb(RELIS)
	<Gn;#DD> (
		Det(LA)
		Nom(PHRASE)
	)
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEd/
">
Gn (
	Det(LES)
	Nom(PARFUMS)
	Gn (
		Pre(DE)
		Det(LA)
		Nom(COLLINE)
	)
)
<Gv;#DG>(
	Aux(E&6TAIENT)
	Pps(DEVENUS)
	<Gn;#DD>(
		Pre (DES)
		Nom(ODEURS)
	)
)
<Gv;#DG>(
	Cco(ET)
	Vrb(MONTAIENT)
	<Gn;#DD>(
		Pre(DU)
		Nom(SOL)
	)
	Vrg(\,)
)
<Gn;#IT>(
	Adv(PRESQUE)
	<Gadj;#DG>(
		Adq(VISIBLES)
	)
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEd/
">
```

```
Gn(
        Pt(POINT)
)
##
##

<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/
">
Gn (
        Det(PLUSIEURS)
        Nom(LAPINS)
)
<Gv;#DG>(
        Vrb(PASSE&2RENT)
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/"
>
Gn (
        Vrg(VIRGULE)
)
##
##

<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
Gn (
        Det(PLUSIEURS)
        Nom(LAPINS)
)
<Gv;#DG>(
        Vrb(PASSE&2RENT)
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
Gn (
        Vrg(VIRGULE)
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/"
>
<Gn;#ID>(
        Adv(AUSSI)
        Pps(PRESSE&6S)
        <Gn;#DD>(
                Ccs(QUE)
```

```
                Pre(DEVANT)
                Det(LES)
                Nom(CHIENS)
        )
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/"
>
Gn (
        Vrg(VIRGULE)
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
<Gn;#ID>(
        Adv(AUSSI)
        Pps(PRESSE&6S)
        <Gn;#DD>(
                Ccs(QUE)
                Pre(DEVANT)
                Det(LES)
                Nom(CHIENS)
        )
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
Gn (
        Vrg(VIRGULE)
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/"
>

<Gn;#ID>(
        Cco(PUIS)
        <Gn;#IT>(
                Det(DES)
                Nom(PERDRIX)
        )
        <Gv;#DG>(
                Vrb(SURGIRENT)
                <Gn;#DD>(
                        Pre (SANS)
                        Nom(BRUIT)
                )
```

```
                <Gn;#IT>(
                        Pre(DU)
                        Nom(VALLON)
                )
        )

)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/"
>
Gn (
        Vrg(VIRGULE)
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>

<Gn;#ID>(
        Cco(PUIS)
        <Gn;#IT>(
                Det(DES)
                Nom(PERDRIX)
        )
        <Gv;#DG>(
                Vrb(SURGIRENT)
                <Gn;#DD>(
                        Pre (SANS)
                        Nom(BRUIT)
                )
                <Gn;#IT>(
                        Pre(DU)
                        Nom(VALLON)
                )
        )

)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
Gn (
        Vrg(VIRGULE)
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/"
>
```

```
<Gn;#ID>(
	Cco(ET)
	Pp(SE)
	Vrb(POSE&2RENT)
	<Gn;#DD>(
		Pre(A&2)
		Num(TRENTE)
		Nom(PAS)
	)
	<Gn;#IT>(
		Pre(SUR)
		Det(NOTRE)
		Nom(GAUCHE)
	)
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEb/
">
Gn (
	Pt(TROIS~POINT)
)
##
##
<intonation;nouv_pro=278,#CT,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>

<Gn;#ID>(
	Cco(ET)
	Pp(SE)
	Vrb(POSE&2RENT)
	<Gn;#DD>(
		Pre(A&2)
		Num(TRENTE)
		Nom(PAS)
	)
	<Gn;#IT>(
		Pre(SUR)
		Det(NOTRE)
		Nom(GAUCHE)
	)
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEc/"
>
Gn (
	Pt(TROIS~POINT)
)
##
```

```
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDE/"
>
Gv (
	Pp(JE)
	Vrb(RELIS)
	<Gn;#DD> (
		Det(LA)
		Nom(PHRASE)
	)
	<Gn;#IT> (
		Pre(EN)
		Nom(ENTIER)
	)
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEd/
">
Gn (
	Det(PLUSIEURS)
	Nom(LAPINS)
)
<Gv;#DG>(
	Vrb(PASSE&2RENT)
	Vrg(\,)
)
<Gn;#ID>(
	Adv(AUSSI)
	Pps(PRESSE&6S)
	<Gn;#DD>(
		Ccs(QUE)
		Pre(DEVANT)
		Det(LES)
		Nom(CHIENS)
	)
	Vrg(\,)
)
<Gn;#ID>(
	Cco(PUIS)
	<Gn;#IT>(
		Det(DES)
		Nom(PERDRIX)
	)
	<Gv;#DG>(
		Vrb(SURGIRENT)
		<Gn;#DD>(
			Pre (SANS)
			Nom(BRUIT)
		)
```

```
                <Gn;#IT>(
                        Pre(DU)
                        Nom(VALLON)
                )
        )
        Vrg(\,)
)
<Gn;#ID>(
        Cco(ET)
        Pp(SE)
        Vrb(POSE&2RENT)
        <Gn;#DD>(
                Pre(A&2)
                Num(TRENTE)
                Nom(PAS)
        )
        <Gn;#IT>(
                Pre(SUR)
                Det(NOTRE)
                Nom(GAUCHE)
        )
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEd/
">
Gn(
        Pt(TROIS~POINT)
)
##
##
<intonation;nouv_pro=278,#DC,nom_prosodie="$FRANCAIS/_livres_sonores/Net/AUDEa/"
>
Gn (
        Nom(TEXTE)
        <Gn;#DD>(
                Pre(DE)
                Npr(MARCEL)
                <Gn;#DG>(
                        Npr(PAGNOL)
                )
        )
)
```

# *ANNEXE G :* Questionnaire

**Renseignements généraux**

**Nom/prénom :** Cliquez ici pour taper du texte. **Age :** Cliquez ici pour taper du texte. ans
**Niveau d'étude :** Doctorat **Domaine d'études :** Choisissez un élément.

| Questions | Beaucoup | Un peu | Indifférent | Très peu | Pas du tout |
|---|---|---|---|---|---|
| Aimez-vous les dictées | ☐ | ☐ | ☐ | ☐ | ☐ |
| Avez-vous déjà fait des dictées automatiques ? | ☐ | ☐ | ☐ | ☐ | ☐ |
| Vous préférez écrire à l'ordinateur plutôt qu'au stylo ? | ☐ | ☐ | ☐ | ☐ | ☐ |

Vous allez transcrire trois textes sous la dictée de trois systèmes de synthèse automatique. Chaque dictée est organisée de la manière suivante :

1. Une lecture du texte dans son entier
2. La dictée est ensuite effectuée phrase par phrase en trois passes :
    a. Une lecture en groupes de mots répétée deux fois
    b. Une lecture entière de la phrase qui vous permettra de vérifier votre orthographe

Merci de taper le texte dans l'encart réservé à cet effet (une phrase par ligne. Naviguez entre les lignes avec les flèches haut/bas), puis de remplir le questionnaire à la fin de la dictée.

**Questionnaire 1. Dictée 1**

| Questions | Très facile | Facile | Moyen | Difficile | Très difficile |
|---|---|---|---|---|---|
| Le niveau de la dictée était | ☐ | ☐ | ☐ | ☐ | ☐ |
| Donner le nombre de mots qui ont été difficiles à orthographier | 0 | | | | |

| Questions | Excellent(e) | Bon(ne) | Moyen(ne) | Mauvais(e) | Médiocre |
|---|---|---|---|---|---|
| L'intelligibilité de la dictée est | ☐ | ☐ | ☐ | ☐ | ☐ |
| L'agrément de la voix est | ☐ | ☐ | ☐ | ☐ | ☐ |
| La manière de lire la dictée est | ☐ | ☐ | ☐ | ☐ | ☐ |
| La vitesse générale de lecture est | ☐ | ☐ | ☐ | ☐ | ☐ |
| Le rendu des ponctuations est | ☐ | ☐ | ☐ | ☐ | ☐ |
| La première lecture du texte est | ☐ | ☐ | ☐ | ☐ | ☐ |
| La deuxième lecture en groupes est | ☐ | ☐ | ☐ | ☐ | ☐ |
| La troisième lecture en mots est | ☐ | ☐ | ☐ | ☐ | ☐ |
| La relecture finale pour vérifier votre transcription est | ☐ | ☐ | ☐ | ☐ | ☐ |

**Questionnaire 2. Dictée 2**

| Questions | Très facile | Facile | Moyen | Difficile | Très difficile |
|---|---|---|---|---|---|
| Le niveau de la dictée était | ☐ | ☐ | ☐ | ☐ | ☐ |
| Donner le nombre de mots qui ont été difficiles à orthographier | 0 | | | | |

| Questions | Excellent(e) | Bon(ne) | Moyen(ne) | Mauvais(e) | Médiocre |
|---|---|---|---|---|---|
| L'intelligibilité de la dictée est | ☐ | ☐ | ☐ | ☐ | ☐ |
| L'agrément de la voix est | ☐ | ☐ | ☐ | ☐ | ☐ |
| La manière de lire la dictée est | ☐ | ☐ | ☐ | ☐ | ☐ |
| La vitesse générale de lecture est | ☐ | ☐ | ☐ | ☐ | ☐ |
| Le rendu des ponctuations est | ☐ | ☐ | ☐ | ☐ | ☐ |
| La première lecture du texte est | ☐ | ☐ | ☐ | ☐ | ☐ |
| La deuxième lecture en groupes est | ☐ | ☐ | ☐ | ☐ | ☐ |
| La troisième lecture en mots est | ☐ | ☐ | ☐ | ☐ | ☐ |
| La relecture finale pour vérifier votre transcription est | ☐ | ☐ | ☐ | ☐ | ☐ |

**Questionnaire 3. Dictée 3**

| Questions | Très facile | Facile | Moyen | Difficile | Très difficile |
|---|---|---|---|---|---|
| Le niveau de la dictée était | ☐ | ☐ | ☐ | ☐ | ☐ |
| Donner le nombre de mots qui ont été difficiles à orthographier | 0 | | | | |

| Questions | Excellent(e) | Bon(ne) | Moyen(ne) | Mauvais(e) | Médiocre |
|---|---|---|---|---|---|
| L'intelligibilité de la dictée est | ☐ | ☐ | ☐ | ☐ | ☐ |
| L'agrément de la voix est | ☐ | ☐ | ☐ | ☐ | ☐ |
| La manière de lire la dictée est | ☐ | ☐ | ☐ | ☐ | ☐ |
| La vitesse générale de lecture est | ☐ | ☐ | ☐ | ☐ | ☐ |
| Le rendu des ponctuations est | ☐ | ☐ | ☐ | ☐ | ☐ |
| La première lecture du texte est | ☐ | ☐ | ☐ | ☐ | ☐ |
| La deuxième lecture en groupes est | ☐ | ☐ | ☐ | ☐ | ☐ |
| La troisième lecture en mots est | ☐ | ☐ | ☐ | ☐ | ☐ |
| La relecture finale pour vérifier votre transcription est | ☐ | ☐ | ☐ | ☐ | ☐ |

## Questionnaire final

| Questions | Meilleur | Moyen | Moins |
|---|---|---|---|
| Donner votre ordre de préférence des systèmes | 1 | 1 | 1 |

| Questions | Excellent(e) | Bon(ne) | Moyen(ne) | Mauvais(e) | Médiocre |
|---|---|---|---|---|---|
| **Pour le meilleur système :** | | | | | |
| Pour faire des dictées à la maison, cet outil | ☐ | ☐ | ☐ | ☐ | ☐ |
| Pour faire des dictées en classe, cet outil est | ☐ | ☐ | ☐ | ☐ | ☐ |
| **Pour le deuxième système :** | | | | | |
| Pour faire des dictées à la maison, cet outil | ☐ | ☐ | ☐ | ☐ | ☐ |
| Pour faire des dictées en classe, cet outil est | ☐ | ☐ | ☐ | ☐ | ☐ |
| **Pour le système le moins bon:** | | | | | |
| Pour faire des dictées à la maison, cet outil | ☐ | ☐ | ☐ | ☐ | ☐ |
| Pour faire des dictées en classe, cet outil est | ☐ | ☐ | ☐ | ☐ | ☐ |

## *ANNEXE H :* Tableaux présentant les résultats d'évaluation des différentes versions « originale », « contrôle » et « multi-style »

| | |
|---|---|
| Q1 | Aimez-vous les dictées |
| Q2 | Avez-vous déjà fait des dictées |
| Q3 | Préférez-vous écrire à l'ordinateur plutôt qu'au stylo |
| Q4 | Niveau de la dictée |
| Q5 | L'intelligibilité de la dictée est |
| Q6 | L'agrément de la voix est |
| Q7 | La manière de lire la dictée est |
| Q8 | La vitesse générale de lecture est |
| Q9 | Le rendu des ponctuations est |
| Q10 | La première lecture du texte est |
| Q11 | La deuxième lecture en groupes est |
| Q12 | La troisième lecture en mots est |
| Q13 | La relecture finale pour vérifier votre transcription est |
| Q14 | Ordre de préférence |

**Figure 48.** **Les différentes questions d'évaluation.**

| | | | | | | | | Version originale | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sujet | Sexe | Age | Etudes | Domaine | Q1 | Q2 | Q3 | ordre | num | Q4 | nbm | Q5 | Q6 | Q7 | Q8 | Q9 | Q10 | Q11 | Q12 | Q13 | rang |
| 1 | F | 21 | Master | Ingénieur | 3 | 1 | 1 | 2 | 5 | 5 | 0 | 4 | 4 | 3 | 3 | 4 | 4 | 4 | 3 | 4 | 3 |
| 2 | F | 22 | Licence | Ingénieur | 4 | 1 | 4 | 2 | 8 | 3 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 5 | 4 | 3 |
| 3 | F | 24 | Doctorat | Ingénieur | 5 | 1 | 2 | 2 | 6 | 5 | 0 | 4 | 4 | 4 | 5 | 4 | 4 | 5 | 5 | 5 | 3 |
| 4 | F | 43 | Master | SHS | 3 | 1 | 5 | 3 | 13 | 5 | 0 | 4 | 4 | 5 | 5 | 5 | 4 | 5 | 5 | 5 | 2 |
| 5 | M | 21 | Master | Informatique | 3 | 1 | 3 | 2 | 9 | 3 | 4 | 4 | 4 | 4 | 5 | 5 | 4 | 5 | 3 | 4 | 3 |
| 6 | M | 21 | Master | Ingénieur | 4 | 1 | 5 | 2 | 2 | 4 | 2 | 3 | 2 | 5 | 5 | 5 | 4 | 4 | 3 | 4 | 3 |
| 7 | M | 22 | Master | Ingénieur | 3 | 1 | 1 | 1 | 12 | 4 | 0 | 3 | 2 | 2 | 3 | 4 | 3 | 3 | 3 | 3 | 2 |
| 8 | M | 22 | Master | Ingénieur | 4 | 1 | 3 | 2 | 5 | 5 | 1 | 4 | 4 | 5 | 5 | 4 | 4 | 3 | 4 | 4 | 3 |
| 9 | M | 22 | Master | Ingénieur | 1 | 1 | 2 | 3 | 0 | 3 | 6 | 4 | 3 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 3 |
| 10 | M | 22 | Master | Ingénieur | 5 | 1 | 4 | 2 | 7 | 4 | 1 | 4 | 3 | 3 | 5 | 5 | 2 | 4 | 4 | 3 | 1 |
| 11 | M | 23 | Master | Ingénieur | 5 | 1 | 5 | 3 | 10 | 5 | 0 | 3 | 4 | 4 | 4 | 4 | 4 | 4 | 3 | 4 | 3 |
| 12 | M | 24 | Doctorat | Ingénieur | 1 | 1 | 3 | 1 | 10 | 4 | 0 | 4 | 2 | 4 | 4 | 5 | 4 | 4 | 4 | 4 | 3 |
| 13 | M | 41 | Doctorat | Ingénieur | 4 | 1 | 2 | 3 | 11 | 5 | 0 | 5 | 4 | 5 | 5 | 5 | 4 | 4 | 4 | 4 | 3 |
| 14 | M | 45 | Doctorat | Ingénieur | 2 | 1 | 4 | 2 | 5 | 3 | 4 | 3 | 3 | 4 | 5 | 4 | 4 | 4 | 4 | 3 | 3 |
| | | | | | | | | | | | | | | | | | | | | | |
| **Moyennes** | | | | | | | | **2,1** | | **4** | **1,6** | **4** | **3** | **4** | **4** | **4** | **3,8** | **4** | **4** | **4** | **2,7** |

**Figure 49.** **Tableau présentant les résultats d'évaluation de la version « originale ».**

| | | | | | | | | Version contrôle | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sujet | Sexe | Age | Étude | Domaine | Q1 | Q2 | Q3 | ordre | num | Q4 | nbm | Q5 | Q6 | Q7 | Q8 | Q9 | Q10 | Q11 | Q12 | Q13 | rang |
| 1 | F | 21 | Master | Ingénieur | 3 | 1 | 1 | 1 | 0 | 5 | 1 | 4 | 3 | 3 | 3 | 3 | 4 | 3 | 3 | 3 | 2 |
| 2 | F | 22 | Licence | Ingénieur | 4 | 1 | 4 | 1 | 7 | 4 | 2 | 2 | 3 | 2 | 3 | 3 | 2 | 3 | 3 | 4 | 1 |
| 3 | F | 24 | Doctorat | Ingénieur | 5 | 1 | 2 | 1 | 5 | 5 | 0 | 3 | 2 | 4 | 2 | 4 | 4 | 3 | 3 | 4 | 2 |
| 4 | F | 43 | Master | SHS | 3 | 1 | 5 | 2 | 12 | 5 | 0 | 1 | 3 | 3 | 4 | 4 | 2 | 3 | 4 | 3 | 1 |
| 5 | M | 21 | Master | Informatique | 3 | 1 | 3 | 1 | 8 | 4 | 1 | 3 | 2 | 4 | 4 | 5 | 4 | 4 | 3 | 4 | 2 |
| 6 | M | 21 | Master | Ingénieur | 4 | 1 | 5 | 1 | 1 | 4 | 2 | 3 | 1 | 3 | 3 | 4 | 4 | 2 | 3 | 4 | 2 |
| 7 | M | 22 | Master | Ingénieur | 3 | 1 | 1 | 2 | 11 | 4 | 0 | 3 | 3 | 4 | 5 | 4 | 2 | 3 | 3 | 3 | 1 |
| 8 | M | 22 | Master | Ingénieur | 4 | 1 | 3 | 1 | 4 | 4 | 1 | 4 | 2 | 3 | 4 | 5 | 4 | 3 | 4 | 4 | 2 |
| 9 | M | 22 | Master | Ingénieur | 1 | 1 | 2 | 2 | 2 | 2 | 6 | 2 | 1 | 3 | 4 | 4 | 3 | 3 | 2 | 3 | 3 |
| 10 | M | 22 | Master | Ingénieur | 5 | 1 | 4 | 1 | 6 | 4 | 1 | 4 | 2 | 3 | 3 | 4 | 4 | 3 | 3 | 4 | 3 |
| 11 | M | 23 | Master | Ingénieur | 5 | 1 | 5 | 2 | 4 | 4 | 1 | 3 | 2 | 2 | 3 | 3 | 4 | 2 | 2 | 3 | 3 |
| 12 | M | 24 | Doctorat | Ingénieur | 1 | 1 | 3 | 3 | 9 | 3 | 0 | 2 | 3 | 4 | 4 | 4 | 3 | 2 | 2 | 3 | 1 |
| 13 | M | 41 | Doctorat | Ingénieur | 4 | 1 | 2 | 2 | 10 | 5 | 0 | 4 | 3 | 3 | 5 | 5 | 3 | 4 | 4 | 4 | 2 |
| 14 | M | 45 | Doctorat | Ingénieur | 2 | 1 | 4 | 1 | 0 | 4 | 1 | 1 | 2 | 2 | 5 | 3 | 2 | 2 | 3 | 3 | 1 |
| | | | | | | | | | | | | | | | | | | | | | |
| **Moyennes** | | | | | | | | **1,5** | | **4** | **1,1** | **3** | **2** | **3** | **4** | **4** | **3,2** | **3** | **3** | **4** | **1,9** |

**Figure 50. Tableau présentant les résultat d'évaluation de la version « contrôle ».**

| | | | | | | | | Version multi-style | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sujet | Sexe | Age | Étude | Domaine | Q1 | Q2 | Q3 | ordre | num | Q4 | nbm | Q5 | Q6 | Q7 | Q8 | Q9 | Q10 | Q11 | Q12 | Q13 | rang |
| 1 | F | 21 | Master | Ingénieur | 3 | 1 | 1 | 3 | 10 | 4 | 0 | 2 | 3 | 1 | 3 | 1 | 3 | 2 | 1 | 2 | 1 |
| 2 | F | 22 | Licence | Ingénieur | 4 | 1 | 4 | 3 | 9 | 3 | 2 | 4 | 3 | 3 | 2 | 2 | 3 | 4 | 3 | 4 | 2 |
| 3 | F | 24 | Doctorat | Ingénieur | 5 | 1 | 2 | 3 | 7 | 4 | 0 | 3 | 3 | 3 | 3 | 4 | 2 | 3 | 3 | 3 | 1 |
| 4 | F | 43 | Master | SHS | 3 | 1 | 5 | 1 | 14 | 4 | 0 | 2 | 3 | 2 | 4 | 4 | 4 | 3 | 2 | 4 | 3 |
| 5 | M | 21 | Master | Informatique | 3 | 1 | 3 | 3 | 10 | 4 | 1 | 3 | 3 | 3 | 4 | 5 | 3 | 4 | 4 | 4 | 1 |
| 6 | M | 21 | Master | Ingénieur | 4 | 1 | 5 | 3 | 3 | 4 | 0 | 2 | 1 | 2 | 3 | 3 | 3 | 2 | 2 | 3 | 1 |
| 7 | M | 22 | Master | Ingénieur | 3 | 1 | 1 | 3 | 13 | 4 | 0 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 3 | 3 |
| 8 | M | 22 | Master | Ingénieur | 4 | 1 | 3 | 3 | 6 | 4 | 1 | 3 | 3 | 4 | 4 | 3 | 3 | 4 | 4 | 4 | 1 |
| 9 | M | 22 | Master | Ingénieur | 1 | 1 | 2 | 1 | 1 | 3 | 6 | 4 | 2 | 3 | 4 | 4 | 3 | 3 | 2 | 4 | 2 |
| 10 | M | 22 | Master | Ingénieur | 5 | 1 | 4 | 3 | 8 | 3 | 3 | 2 | 1 | 3 | 3 | 4 | 3 | 2 | 2 | 3 | 2 |
| 11 | M | 23 | Master | Ingénieur | 5 | 1 | 5 | 1 | 11 | 5 | 0 | 4 | 3 | 2 | 2 | 2 | 4 | 2 | 2 | 2 | 3 |
| 12 | M | 24 | Doctorat | Ingénieur | 1 | 1 | 3 | 2 | 11 | 4 | 0 | 3 | 2 | 3 | 3 | 4 | 4 | 2 | 2 | 4 | 2 |
| 13 | M | 41 | Doctorat | Ingénieur | 4 | 1 | 2 | 1 | 12 | 4 | 2 | 3 | 2 | 3 | 3 | 5 | 3 | 2 | 3 | 4 | 1 |
| 14 | M | 45 | Doctorat | Ingénieur | 2 | 1 | 4 | 3 | 10 | 3 | 4 | 1 | 1 | | 3 | 2 | 2 | 2 | 2 | 4 | 2 |
| | | | | | | | | | | | | | | | | | | | | | |
| **Moyennes** | | | | | | | | **2,4** | | **4** | **1,4** | **3** | **2** | **3** | **3** | **3** | **3,1** | **3** | **3** | **3** | **1,8** |

**Figure 51. Tableau présentant les résultat d'évaluation de la version « multi-style ».**

# ***ANNEXE I :*** **Le schwa et son élision**

D'après Goldman (Goldman, J.P., de l'analyse syntaxique à la synthèse de la parole dans le système FipsVox) « Le schwa, appelé également e-muet ou e-caduc, est un phonème instable pouvant disparaître dans certaines conditions. Ce schwa peut se trouver en début de mot (*f**e**nêtre*), en milieu de mot (*sam**e**di*) ou en fin de mot (*j**e** m**e** l**e** demande*). Une séquence de schwas, comme dans le dernier exemple, peut donner plusieurs prononciations possibles (*j'me le d'mande, je m'le d'mande, j'me l'demande*,…) ainsi que dans le mot *genevois* (*gen'vois* ou *g'nevois*). Les liens entre l'élision du schwa et la linguistique étant limités, nous nous contentons de décrire brièvement la problématique, les études portant sur ce sujet et le traitement du schwa dans notre système.

Les facteurs influençant la chute, ou amuïssement, du schwa sont d'ordre phonétique (contexte phonétique), articulatoire (selon Malécot 1955, la chute du schwa dépend de la différence d'aperture entre les consonnes accolées), prosodique (accentuation du mot), linguistique (nature du mot, sa fréquence – comparez *cerise* et *merise*) et phonostylistiques (conversation spontanée ou lecture oralisée, langage familier ou soutenu, débit de parole, accent régional). En plus de ces domaines, la chute du schwa soulève d'autres questions en traitement automatique du langage ainsi qu'en psycholinguistique comme la représentation lexicale du schwa et l'influence sur l'accès au lexique mental (Racine 2001), (Boula de Mareuil 1997), (Dard 2001).

Les études les plus connues visant à modéliser ce phénomène phonologique sont celles de Grammont dès 1894, puis celles de Malécot qui reprennent les premières. Le résultat principal est l'établissement de la règles dite des 3 consonnes, qui consiste à interdire la chute du schwa si celle-ci provoque une séquence de 3 consonnes (sa**m'd**i mais *mer**cr'd**i) à moins que la troisième ne soit une liquide ou une semi-voyelle. De plus si la deuxième consonne est une liquide, l'élision est également interdite (ex : *donnerions* et *Richelieu*). De nombreux cas particuliers existent et demandent des traitements spécifiques comme les paires minimales de mots (le schwa de *belette* ne s'élide pas à cause de l'existence de *blette*, tout comme *peluche* et *pluche* malgré leur éthymologie commune), le h aspiré qui bloque l'élision (**l'hibou*) sauf dans les nombres (*quarant'huit*), le 'de' nobiliaire qui n'est jamais élidé (*Madame de Fontenay*)… »

# *ANNEXE J:* Erreurs commises en fonction de chaque système

## Système «contrôle »

SCORING FOR SPEAKER: ctrl
of /tmp/tst

SENTENCE RECOGNITION PERFORMANCE

| | | |
|---|---|---|
| sentences | 14 | |
| with errors | 78.6% | ( 11) |
| with substitions | 78.6% | ( 11) |
| with deletions | 21.4% | ( 3) |
| with insertions | 21.4% | ( 3) |

WORD RECOGNITION PERFORMANCE

Percent Total Error = 5.7% ( 75)

Percent Correct = 94.6% (1247)

Percent Substitution = 4.9% ( 65)
Percent Deletions = 0.5% ( 6)
Percent Insertions = 0.3% ( 4)
Percent Word Accuracy = 94.3%

Ref. words = (1318)
Hyp. words = (1316)
Aligned words = (1322)

CONFUSION PAIRS Total (56)
With >= 1 occurances (56)

1: 6 -> ; ==> ,
2: 2 -> ! ==> .
3: 2 -> : ==> ,
4: 2 -> minute ==> minutes
5: 2 -> velours ==> velour
6: 1 -> , ==> .
7: 1 -> : ==> .
8: 1 -> accompagnée ==> accompagné
9: 1 -> alourdis ==> allourdis

10: 1 -> animale ==> animal
11: 1 -> arrêtâmes ==> arretâmes
12: 1 -> arrêtâmes ==> arrétâmes
13: 1 -> au ==> aux
14: 1 -> balaya ==> baleya
15: 1 -> bleuâtre ==> ?
16: 1 -> casserole ==> casserol
17: 1 -> cotonneux ==> cotoneux
18: 1 -> coucher ==> couché
19: 1 -> deux ==> des
20: 1 -> devenus ==> devenu
21: 1 -> disparut ==> disparu
22: 1 -> disparut ==> disparût
23: 1 -> drame ==> drâme
24: 1 -> déserte ==> deserte
25: 1 -> elle ==> il
26: 1 -> enclos ==> enclots
27: 1 -> haletait ==> hâletait
28: 1 -> hantées ==> hantés
29: 1 -> inquiétude ==> inquiètude
30: 1 -> j'étendis ==> tendis
31: 1 -> jaune ==> jaunes
32: 1 -> jeta ==> jetta
33: 1 -> l'hiver ==> l'hivers
34: 1 -> la ==> les
35: 1 -> leurs ==> leur
36: 1 -> lever ==> levé
37: 1 -> nos ==> notre
38: 1 -> notre ==> nore
39: 1 -> pelage ==> peulage
40: 1 -> pleins ==> plein
41: 1 -> poil ==> poils
42: 1 -> prit ==> pris
43: 1 -> province ==> provinces
44: 1 -> put ==> pû
45: 1 -> pénétrait ==> pénétetrait
46: 1 -> recueillit ==> receuilli
47: 1 -> rôdaient ==> rodaient
48: 1 -> s'ouvrit ==> s'ouvrait
49: 1 -> sauvées ==> sauver
50: 1 -> scieries ==> scierie
51: 1 -> tamouri ==> tamoori
52: 1 -> toute ==> toutes
53: 1 -> têtes ==> tete
54: 1 -> vieux ==> vieu
55: 1 -> âgé ==> agé
56: 1 -> étouffée ==> étauffée
---
65

INSERTIONS Total (3)
With >= 1 occurances (3)

1: 2 -> ,
2: 1 -> a
3: 1 -> je
---
4

DELETIONS Total (4)
With >= 1 occurances (4)

1: 3 -> .
2: 1 -> -
3: 1 -> biche
4: 1 -> la
---
6

SUBSTITUTIONS Total (53)
With >= 1 occurances (53)

1: 6 -> ;
2: 3 -> :
3: 2 -> !
4: 2 -> arrêtâmes
5: 2 -> disparut
6: 2 -> minute
7: 2 -> velours
8: 1 -> ,
9: 1 -> accompagnée
10: 1 -> alourdis
11: 1 -> animale
12: 1 -> au
13: 1 -> balaya
14: 1 -> bleuâtre
15: 1 -> casserole
16: 1 -> cotonneux
17: 1 -> coucher
18: 1 -> deux
19: 1 -> devenus
20: 1 -> drame
21: 1 -> déserte
22: 1 -> elle

23: 1 -> enclos
24: 1 -> haletait
25: 1 -> hantées
26: 1 -> inquiétude
27: 1 -> j'étendis
28: 1 -> jaune
29: 1 -> jeta
30: 1 -> l'hiver
31: 1 -> la
32: 1 -> leurs
33: 1 -> lever
34: 1 -> nos
35: 1 -> notre
36: 1 -> pelage
37: 1 -> pleins
38: 1 -> poil
39: 1 -> prit
40: 1 -> province
41: 1 -> put
42: 1 -> pénétrait
43: 1 -> recueillit
44: 1 -> rôdaient
45: 1 -> s'ouvrit
46: 1 -> sauvées
47: 1 -> scieries
48: 1 -> tamouri
49: 1 -> toute
50: 1 -> têtes
51: 1 -> vieux
52: 1 -> âgé
53: 1 -> étouffée

---

65

* NOTE: The 'Substitution' words are those reference words for which the recognizer supplied an incorrect word.

FALSELY RECOGNIZED Total (53)
With >= 1 occurances (53)

1: 8 -> ,
2: 4 -> .
3: 2 -> minutes
4: 2 -> velour
5: 1 -> ?
6: 1 -> accompagné
7: 1 -> agé
8: 1 -> allourdis

9: 1 -> animal
10: 1 -> arretâmes
11: 1 -> arrétâmes
12: 1 -> aux
13: 1 -> baleya
14: 1 -> casserol
15: 1 -> cotoneux
16: 1 -> couché
17: 1 -> des
18: 1 -> deserte
19: 1 -> devenu
20: 1 -> disparu
21: 1 -> disparût
22: 1 -> drâme
23: 1 -> enclots
24: 1 -> hantés
25: 1 -> hâletait
26: 1 -> il
27: 1 -> inquiètude
28: 1 -> jaunes
29: 1 -> jetta
30: 1 -> l'hivers
31: 1 -> les
32: 1 -> leur
33: 1 -> levé
34: 1 -> nore
35: 1 -> notre
36: 1 -> peulage
37: 1 -> plein
38: 1 -> poils
39: 1 -> pris
40: 1 -> provinces
41: 1 -> pénétetrait
42: 1 -> pû
43: 1 -> receuilli
44: 1 -> rodaient
45: 1 -> s'ouvrait
46: 1 -> sauver
47: 1 -> scierie
48: 1 -> tamoori
49: 1 -> tendis
50: 1 -> tete
51: 1 -> toutes
52: 1 -> vieu
53: 1 -> étauffée
---
65

* NOTE: The 'Falsely Recognized' words are those hypothesis words

which the recognizer incorrectly substituted for a reference word.

## Système « originale »

```
SCORING FOR SPEAKER: org
    of /tmp/tst

SENTENCE RECOGNITION PERFORMANCE

 sentences                          13
 with errors                   84.6%   (  11)

   with substitions             76.9%   (  10)
   with deletions               53.8%   (   7)
   with insertions               0.0%   (   0)


WORD RECOGNITION PERFORMANCE

Percent Total Error      =    3.5%   (  42)

Percent Correct          =   96.5%   (1172)

Percent Substitution     =    2.7%   (  33)
Percent Deletions        =    0.7%   (   9)
Percent Insertions       =    0.0%   (   0)
Percent Word Accuracy    =   96.5%


Ref. words               =          (1214)
Hyp. words               =          (1205)
Aligned words            =          (1214)

CONFUSION PAIRS                Total              (31)
                         With >=  1 occurances (31)

   1:    2  ->  minute ==> minutes
   2:    2  ->  vieux ==> vieu
   3:    1  ->  accompagnée ==> accompagné
   4:    1  ->  alourdis ==> allourdis
   5:    1  ->  au ==> aux
   6:    1  ->  circula ==> circulât
   7:    1  ->  crénelée ==> crenelée
   8:    1  ->  d'une ==> d'un
   9:    1  ->  dangereuse ==> dangeureuse
```

10: 1 -> dominent ==> domine
11: 1 -> enclos ==> enclôts
12: 1 -> furent ==> fûrent
13: 1 -> jaune ==> jaunes
14: 1 -> leurs ==> leur
15: 1 -> meurt ==> meure
16: 1 -> nos ==> le
17: 1 -> nôtre ==> neutre
18: 1 -> perdu ==> perdue
19: 1 -> poil ==> poils
20: 1 -> presque ==> preque
21: 1 -> pressés ==> préssés
22: 1 -> prés ==> pré
23: 1 -> rôdaient ==> rodaient
24: 1 -> s'ouvrit ==> s'ouvrait
25: 1 -> scieries ==> scierie
26: 1 -> sûreté ==> sureté
27: 1 -> t'ai ==> j'étais
28: 1 -> tins ==> teint
29: 1 -> trébuche ==> trébûche
30: 1 -> âgé ==> agé
31: 1 -> ôtée ==> otée

---
33

INSERTIONS Total (0)
With >= 1 occurances (0)

---
0

DELETIONS Total (4)
With >= 1 occurances (4)

1: 5 -> .
2: 2 -> ,
3: 1 -> arbres
4: 1 -> je

---
9

SUBSTITUTIONS Total (31)
With >= 1 occurances (31)

1: 2 -> minute
2: 2 -> vieux
3: 1 -> accompagnée
4: 1 -> alourdis
5: 1 -> au
6: 1 -> circula
7: 1 -> crénelée
8: 1 -> d'une
9: 1 -> dangereuse
10: 1 -> dominent
11: 1 -> enclos
12: 1 -> furent
13: 1 -> jaune
14: 1 -> leurs
15: 1 -> meurt
16: 1 -> nos
17: 1 -> nôtre
18: 1 -> perdu
19: 1 -> poil
20: 1 -> presque
21: 1 -> pressés
22: 1 -> prés
23: 1 -> rôdaient
24: 1 -> s'ouvrit
25: 1 -> scieries
26: 1 -> sûreté
27: 1 -> t'ai
28: 1 -> tins
29: 1 -> trébuche
30: 1 -> âgé
31: 1 -> ôtée
---
33

* NOTE: The 'Substitution' words are those reference words
for which the recognizer supplied an incorrect word.

FALSELY RECOGNIZED Total (31)
With >= 1 occurances (31)

1: 2 -> minutes
2: 2 -> vieu
3: 1 -> accompagné
4: 1 -> agé
5: 1 -> allourdis
6: 1 -> aux
7: 1 -> circulât
8: 1 -> crenelée

```
 9:    1  ->  d'un
10:    1  ->  dangeureuse
11:    1  ->  domine
12:    1  ->  enclôts
13:    1  ->  fûrent
14:    1  ->  j'étais
15:    1  ->  jaunes
16:    1  ->  le
17:    1  ->  leur
18:    1  ->  meure
19:    1  ->  neutre
20:    1  ->  otée
21:    1  ->  perdue
22:    1  ->  poils
23:    1  ->  preque
24:    1  ->  pré
25:    1  ->  préssés
26:    1  ->  rodaient
27:    1  ->  s'ouvrait
28:    1  ->  scierie
29:    1  ->  sureté
30:    1  ->  teint
31:    1  ->  trébûche
   -------
      33
```

* NOTE: The 'Falsely Recognized' words are those hypothesis words which the recognizer incorrectly substituted for a reference word.

## Système «multi-style »

```
SCORING FOR SPEAKER: multi
    of /tmp/tst

SENTENCE RECOGNITION PERFORMANCE

 sentences                          13
 with errors                     84.6%   (  11)

   with substitions              84.6%   (  11)
   with deletions                38.5%   (   5)
   with insertions                7.7%   (   1)
```

WORD RECOGNITION PERFORMANCE

Percent Total Error = 3.8% ( 45)

Percent Correct = 96.3% (1150)

Percent Substitution = 3.0% ( 36)
Percent Deletions = 0.7% ( 8)
Percent Insertions = 0.1% ( 1)
Percent Word Accuracy = 96.2%

Ref. words = (1194)
Hyp. words = (1187)
Aligned words = (1195)

CONFUSION PAIRS Total (29)
With >= 1 occurances (29)

1: 5 -> ; ==> ,
2: 3 -> : ==> ,
3: 2 -> ! ==> .
4: 1 -> : ==> .
5: 1 -> charrette ==> charette
6: 1 -> chats ==> chat
7: 1 -> coururent ==> courûrent
8: 1 -> courut ==> courût
9: 1 -> deux ==> des
10: 1 -> détours ==> détour
11: 1 -> détresse ==> detresse
12: 1 -> furent ==> fûrent
13: 1 -> hantées ==> hanter
14: 1 -> la ==> leur
15: 1 -> les ==> le
16: 1 -> m'approche ==> m'aproche
17: 1 -> miauler ==> mioler
18: 1 -> mètres ==> metres
19: 1 -> passées ==> passé
20: 1 -> petits ==> petit
21: 1 -> prit ==> pris
22: 1 -> put ==> pu
23: 1 -> recueillit ==> receuillit
24: 1 -> réveillé ==> reveillé
25: 1 -> scieurs ==> sieurs
26: 1 -> seuls ==> seul
27: 1 -> tous ==> tout
28: 1 -> étouffée ==> étoufée
29: 1 -> ôtée ==> otée
---
36

INSERTIONS Total (1)
With >= 1 occurances (1)

1: 1 -> ,
---
1

DELETIONS Total (5)
With >= 1 occurances (5)

1: 4 -> ,
2: 1 -> .
3: 1 -> drame
4: 1 -> minet
5: 1 -> pauvre
---
8

SUBSTITUTIONS Total (28)
With >= 1 occurances (28)

1: 5 -> ;
2: 4 -> :
3: 2 -> !
4: 1 -> charrette
5: 1 -> chats
6: 1 -> coururent
7: 1 -> courut
8: 1 -> deux
9: 1 -> détours
10: 1 -> détresse
11: 1 -> furent
12: 1 -> hantées
13: 1 -> la
14: 1 -> les
15: 1 -> m'approche
16: 1 -> miauler
17: 1 -> mètres
18: 1 -> passées
19: 1 -> petits
20: 1 -> prit
21: 1 -> put
22: 1 -> recueillit

23: 1 -> réveillé
24: 1 -> scieurs
25: 1 -> seuls
26: 1 -> tous
27: 1 -> étouffée
28: 1 -> ôtée
---
36

* NOTE: The 'Substitution' words are those reference words
for which the recognizer supplied an incorrect word.

FALSELY RECOGNIZED Total (27)
With >= 1 occurances (27)

1: 8 -> ,
2: 3 -> .
3: 1 -> charette
4: 1 -> chat
5: 1 -> courûrent
6: 1 -> courût
7: 1 -> des
8: 1 -> detresse
9: 1 -> détour
10: 1 -> fûrent
11: 1 -> hanter
12: 1 -> le
13: 1 -> leur
14: 1 -> m'aproche
15: 1 -> metres
16: 1 -> mioler
17: 1 -> otée
18: 1 -> passé
19: 1 -> petit
20: 1 -> pris
21: 1 -> pu
22: 1 -> receuillit
23: 1 -> reveillé
24: 1 -> seul
25: 1 -> sieurs
26: 1 -> tout
27: 1 -> étoufée
---
36

* NOTE: The 'Falsely Recognized' words are those hypothesis words
which the recognizer incorrectly substituted for a reference word.

*Fin.*