

Stylisation des Contours Intonatifs et Transcription Tonale

NICOLAS OBIN

TABLE DES MATIÈRES

1	Introduction	1
2	Formats d'entrée/sorties	1
3	Présentation de l'outil	2
3.1	Registre et Contour	2
3.2	Transcription	3
3.2.1	Contours	3
3.2.2	Registre	4
4	Unités traitées	4
4.1	Unités prosodiques	5
4.2	Unités syntaxiques	5

1 INTRODUCTION

Nous présentons ici une courte documentation de l'outil de stylisation des contours intonatifs et de transcription tonale développés dans le cadre du projet Rhapsodie.

2 FORMATS D'ENTRÉE/SORTIES

L'outil prend au choix un fichier de segmentation au format TextGrid ou XML et un fichier d'analyse de F0 au format PITCH ou autre.

- TextGrid : dans le cas d'un fichier d'entrée TextGrid, la transcription tonale est ajoutée sous forme de nouvelles Tiers avec les identifiants `<basename_contour>` et `<basename_register>`.
- XML : dans le cadre de Rhapsodie, les fichiers d'annotation sont sous format XML (prosodie et syntaxe). Pour chaque fichier d'annotation, la transcription tonale est déterminée pour chaque balise linguistique pourvue de tags de temps de début et temps de fin (`<time_start>`, `<time_end>`). Chaque balise est alors enrichie de tag `<basename_contour>` et `<basename_register>`.

La traitement différencié des locuteurs nécessite la présence d'une tier `<speaker>` pour les TextGrids, ou d'une balise `<speaker>` pourvu des tags `<speaker_id>`, `<time_start>`,

<time_end> dans le cas d'un fichier XML.

Dans Rhapsodie, le traitement des locuteurs est réalisé par l'intermédiaire des fichiers *pivots* (fichiers contenant l'alignement des mots et des locuteurs).

Dans le cas de chevauchement de parole :

- le locuteur principal est spécifié en première position (« \$L1 \$L2 » signifie que L1 est le locuteur principal)
- les contours sont déterminés en prenant le locuteur principal comme référence, et en excluant les segments de tour de parole des autres locuteurs – si présent.

3 PRÉSENTATION DE L'OUTIL

L'outil se décompose en 3 parties :

- Stylisation des contours de F0 par décomposition sur une base de fonctions lentement variable en temps afin de filtrer les variations micro-prosodiques mais également les éventuelles erreurs d'estimation de la F0 (optionnel)
- Détermination des caractéristiques acoustiques de F0 sur un segment linguistique arbitraire (ex : syllabe, mot, pile)
- Transcription des contours de F0 sur un alphabet tonal

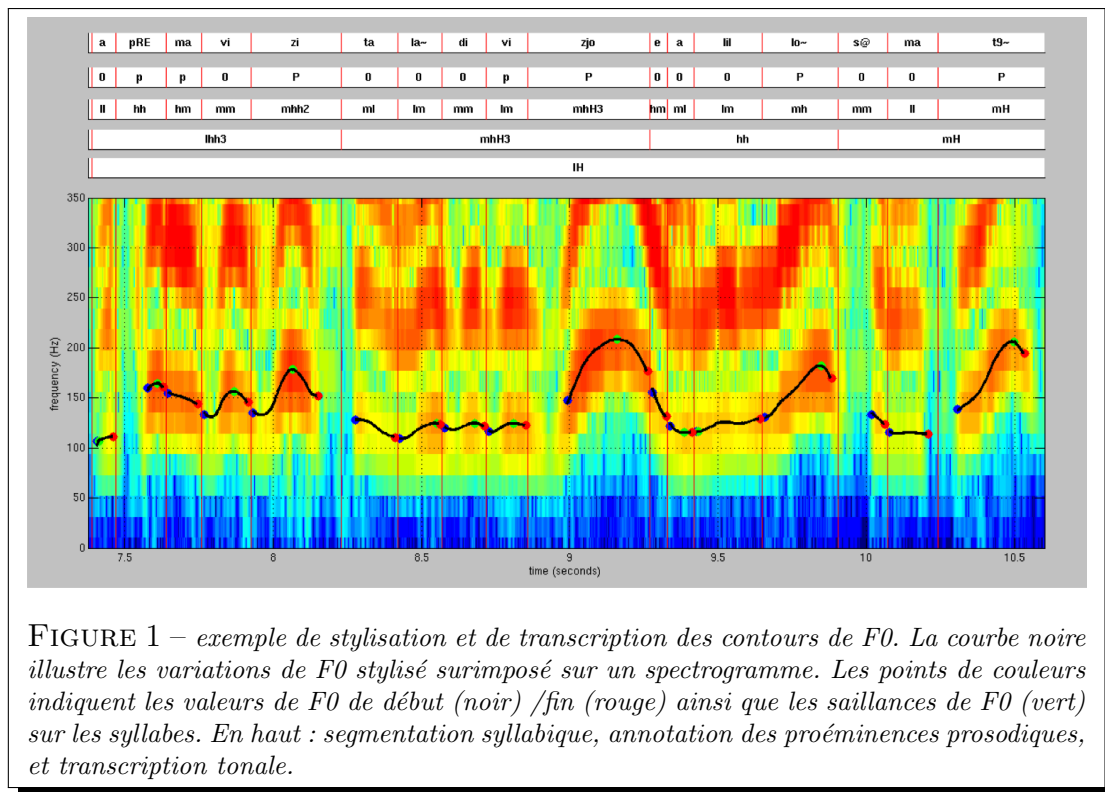
3.1 REGISTRE ET CONTOUR

Le registre illustre le registre local du locuteur par rapport à sa tessiture.

Les contours sont déterminés sur les points de F0 suivants :

1. niveau initial
2. niveau et position des saillances intermédiaires (si présence)
3. niveau final

Le niveau est par ailleurs normalisé par considération du registre moyen du locuteur.



3.2 TRANSCRIPTION

3.2.1 CONTOURS

Lors de la transcription des contours, les niveaux et les positions prennent les valeurs suivantes :

Niveau	
L	niveau bas extrême
l	niveau bas
m	niveau medium
h	niveau haut
H	niveau haut extrême

TABLE 1 – Niveaux de base utilisé pour la stylisation des contours et du registre

Les valeurs de niveaux sont déterminées à partir d'une discrétisation du niveau de hauteur est 5 intervalles égaux en échelle logarithmique, et sont utilisés pour déterminer les contours et les registres.

La position des saillances permet de décrire la structure dynamique des contours, en spécifiant la position des saillances de F0 à l'intérieur d'une unité donnée (e.g., syllabe, mot, etc...)

Position	
1	Premier tiers de l'unité considérée
2	Deuxième tiers de l'unité considérée
3	Dernier tiers de l'unité considérée

TABLE 2 – Valeurs de position utilisées pour la stylisation des contours

Exemples :

- Un contour « mm » indique un plateau dans le registre medium du locuteur.
- Un contour « lH2h » indique un contour initié dans le niveau bas (« l »), terminé dans le registre haut (« h »), et avec une saillance dans le registre haut-extrême (« H ») réalisé dans le cœur de l'unité (« 2 »)

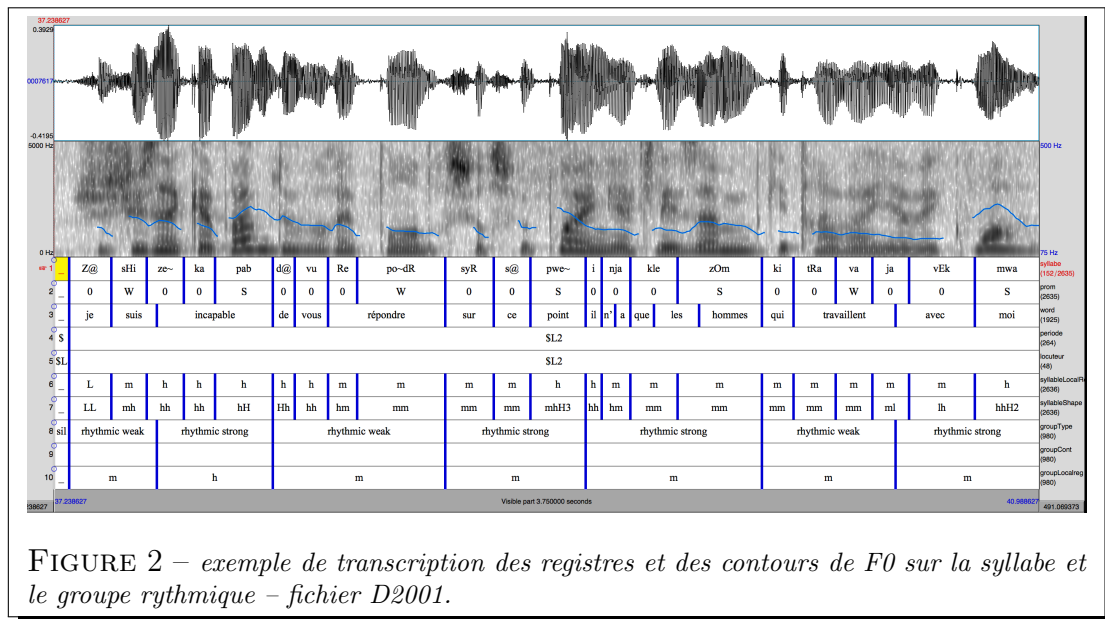


FIGURE 2 – exemple de transcription des registres et des contours de F0 sur la syllabe et le groupe rythmique – fichier D2001.

3.2.2 REGISTRE

Le registre local est décrit à partir des 5 valeurs de niveau de base, soit « L », « l », « m », « h », et « H ».

4 UNITÉS TRAITÉES

Les contours et registres sont déterminés pour chaque élément de XML accompagné de balise de temps *<time_start>* et *<time_end>*, et des balises *<shape>* et *<localreg>* – ceci à l'exception des unités de type phone qui ne sont pas considérés ici comme une unité pertinente pour la stylisation.

4.1 UNITÉS PROSODIQUES

Unité
syllable
foot
word
group
package
period

TABLE 3 – Unités prosodiques

4.2 UNITÉS SYNTAXIQUES

Unité
discursivemarker
embedded
graft
inkernel
integratedinkernel
integratedpostkernel
integratedprekernel
intro
iu
kernel
layer
lexeme
parenthesis
pile
postkernel
prekernel

TABLE 4 – Unités syntaxiques