# Protocole de codage prosodique

annotation des proéminences, disfluences et périodes dans le projet Rhapsodie

## Table des matières


**1 Présentation** **3**

**2 Pré-requis techniques** **3**
2.1 Le logiciel Praat . . . 3
2.1.1 Installer Praat . . . 3
2.1.2 Configurer Praat . . . 3
2.1.3 Charger et éditer un fichier . . . 4
2.1.4 Comment se présente le fichier ? . . . 7
2.1.5 Quelques raccourcis . . . 8
2.2 Le logiciel Analor . . . 9
2.2.1 Comment installer Analor ? . . . 9
2.2.2 À quoi sert Analor ? . . . 9
2.2.3 Le menu de la fenêtre principale . . . 9
2.2.3.1 Documents . . . 9
2.2.3.2 Affichage . . . 9
2.2.3.3 Statistiques . . . 10
2.2.3.4 Paramétrage . . . 10
2.2.3.5 Stylisations . . . 10
2.2.3.6 Périodes . . . 10
2.2.3.7 Proéminences . . . 10
2.2.4 Les autres fonctionnalités . . . 10
2.2.4.1 Sélection d'un intervalle temporel et écoute du son . . . 10
2.2.4.2 Modifications de la fondamentale . . . 11
2.2.4.3 Opérations sur les tires d'annotation . . . 11

**3 Repères définitoires** **11**
3.1 Proéminences et principe de contraste . . . 11
3.2 Disfluence et principe de non linéarité . . . 13
3.3 Les périodes, unités prosodiques majeures . . . 13

**4 Consignes pour le codage des proéminences et des disfluences** **13**
4.1 Les conditions d'un bon codage . . . 13
4.2 Les étapes de codage . . . 13
4.2.1 Principes généraux de codage . . . 14
4.2.2 Etape 0 : Choisir sa fenêtre de codage . . . 14
4.2.3 Etape 1 : Codage des proéminences fortes : intervalle de confiance ou empan de codage . . . 17
4.2.4 Etape 2 : Codage des proéminences faibles (« W ») et des disfluences (« H ») . . . 19
4.2.4.1 Codage des « W » . . . 19
4.2.4.2 Codage des « H » . . . 20
4.2.5 Etape 3 : Empan suivant . . . 23
4.2.6 Etape 4 : Finalisation du codage . . . 25

**5 Surcodage vs proéminences contiguës effectives** **27**
5.1 Contextes terminaux : syllabe pénultième + syllabe finale . . . 28
5.2 Contiguité d'un accent final et initial . . . 29
5.3 Contextes d'insistance ou de scansion rythmique . . . 29
5.4 La part de l'acoustique . . . 30

**6 Remarques et précisions pour le codage** **30**
6.1 Les proéminences . . . 30
6.1.1 Que faire lorsqu'on hésite entre « W » et « S » ? . . . 30
6.1.2 Que faire lorsqu'on hésite à annoter une syllabe « W » ou non ? . . . 30
6.1.3 Comment fixer le degré d'une proéminence par rapport à son entourage ? . . . 30
6.1.4 Faut-il tenir compte des allongements syllabiques perçus ou des variations mélodiques pour déterminer la proéminence ? . . . 31
6.1.5 Comment traiter le « e » extramétrique ou post-tonique ? . . . 31
6.1.6 Comment annoter l'attaque consonantique en initiale de mot ? . . . 32
6.2 Les disfluences . . . 33
6.2.1 Que faire lorsqu'on hésite à annoter une syllabe « H » ou non ? . . . 33
6.2.2 Une séquence inachevée est-elle disfluente ? . . . 33
6.2.3 Le schwa post-tonique ou extra-métrique peut-il être annoté disfluent ? . . . 33
6.2.4 Les marqueurs de discours sont-ils tous disfluents ? . . . 33
6.2.5 Les reformulations sont-elles des disfluences ? . . . 33
6.2.6 Les fausses hésitations sont-elles des disfluences ? . . . 33
6.3 Les proéminences et/ou les dyfluences . . . 33
6.3.1 Une syllabe peut-elle être proéminente ET disfluente ? . . . 33
6.3.2 Que faire lorsqu'on hésite entre proéminence et disfluence pour une syllabe allongée ? 34
6.4 Divers . . . 35
6.4.1 Comment coder les fichiers multi-locuteurs ? . . . 35

**7 Suivi d'annotations (pour les experts uniquement)** **35**
7.1 Accord inter-annotateur . . . 35
7.2 Conseils pour la création de l'annotation de référence . . . 35
7.3 Remarques pouvant aider au suivi d'annotation . . . 36

**8 La segmentation en périodes intonatives** **36**
8.1 à venir . . . 36
8.2 à venir . . . 36

**Glossary** **37**

# 1 Présentation

Ce guide propose une méthode d'annotation prosodique unique pour des novices et des experts telle qu'elle a été mise en place dans le cadre du projet Rhapsodie : les étapes 0 à 4 permettent de réaliser une annotation prosodique par des novices comme par des experts. Le suivi d'annotation (cf. section 7) est une étape à suivre uniquement par les experts souhaitant réaliser un suivi d'annotation et éventuellement dériver une annotation de référence.

Ce tutoriel va vous guider dans votre tâche de codage : **il est primordial de le lire en entier avant de vous lancer dans le codage et de vous y reporter au moindre doute.**
Vous trouverez tout d'abord des conseils sur les aspects techniques, puis une description du processus de codage, la section 6 contient tous les remarques, conseils, détails et principes à suivre selon les situations rencontrées. Cette section n'est donc en aucun cas falcultative : il convient de la lire tout comme le reste du guide avant de se lancer.

# 2 Pré-requis techniques

L'annotation des proéminences et des disfluences est réalisée manuellement sous PRAAT.
La segmentation en périodes est effectuée automatiquement sous Analor [1] et vérifiée manuellement sous Analor ou sous PRAAT [2]. Une petite prise en main préalable des deux logiciels pour réaliser ces deux types de tâches s'avère donc nécessaire.

## 2.1 Le logiciel Praat

### 2.1.1 Installer Praat

Tout d'abord il faut télécharger Praat à l'adresse suivante : http://www.fon.hum.uva.nl/praat/.
Praat ne nécessite aucune installation, le fichier exécutable est le logiciel.
Il est possible d'installer des polices permettant de visualiser les caractères phonétiques disponibles à ces adresses : http://www.fon.hum.uva.nl/praat/CharisSIL-4.110.zip ou http://www.fon.hum.uva.nl/praat/DoulosSIL-4.110.zip. Il suffit alors d'extraire les fichiers et de les installer en cliquant-droit et en sélectionnant « installer ». Néanmoins, dans le protocole décrit ici, les caractères utilisés sont les caractères SAMPA qui ne requièrent pas de police particulière.

### 2.1.2 Configurer Praat

Une fois l'installation terminée, il convient de configurer Praat au format *UTF-8*, l'encodage est un élément **très important** qui facilite l'interopérabilité, condition nécessaire à un travail en équipe.
Pour cela il suffit de cliquer sur le menu *Praat* dans la fenêtre *Praat Objects*, puis aller dans *Preferences* puis *Text reading preferences* et enfin sélectionner l'encodage désiré (ici UTF-8), cela ouvrira les fichiers dans l'encodage sélectionné.

1. A noter qu'Analor dispose également d'une fonction de détection automatique des proéminences
2. Un tutoriel Praat écrit par Jean-Philippe Goldman en français est disponible à l'adresse suivante : http://latlcui.unige.ch/phonetique/easyalign/tutorielpraat.pdf

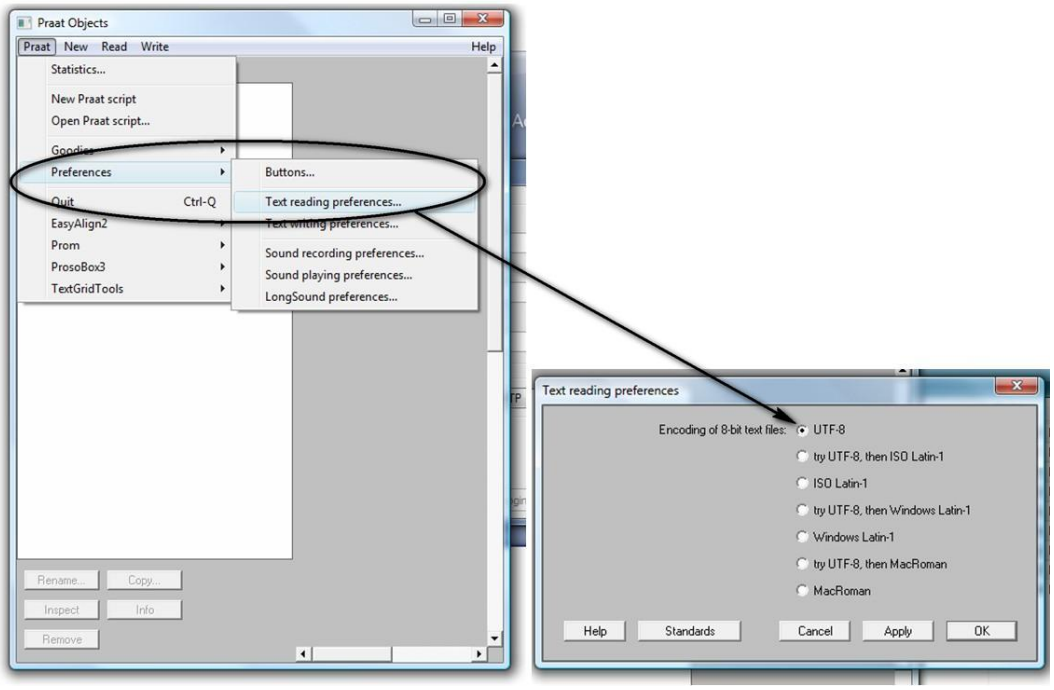


FIGURE 1 – Sélection de Text reading preferences

On répète l'opération pour enregistrer les fichiers dans l'encodage UTF-8, il suffit cette fois-ci de cliquer sur *Text writing preferences* au lieu de *Text reading preferences* et de sélectionner à nouveau l'encodage UTF-8.

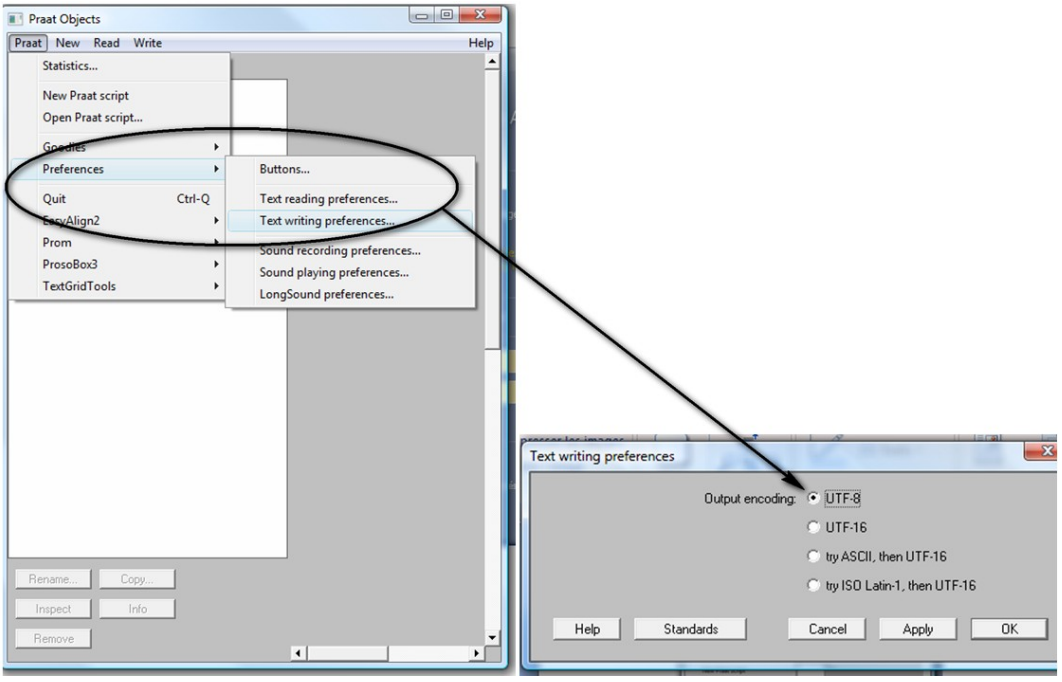


FIGURE 2 – Sélection de Text writing preferences

### 2.1.3 CHARGER ET ÉDITER UN FICHIER

Pour coder la prosodie, il nous faut charger le fichier son (.wav) *et* le fichier d'annotation (.TextGrid). Pour ouvrir le fichier son dans Praat avec le fichier d'alignement, il suffit de cliquer sur *Read* puis de

sélectionner *Read from file...* afin de parcourir les répertoires et de choisir le bon fichier son qui s'affichera alors en bleu dans la fenêtre *Praat Objects*, puis on répète l'opération pour le fichier d'annotation.

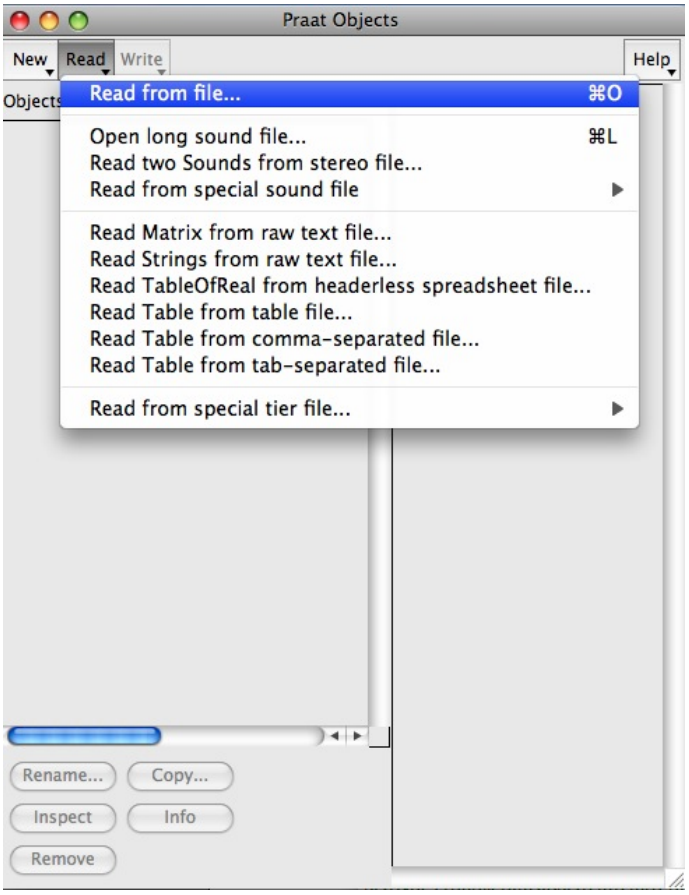


FIGURE 3 – Sélection de Read from file...

Une fois que les deux fichiers sont chargés, on les sélectionne (avec *Shift ou CMD* selon le système d'exploitation) et on clique sur *Edit*.

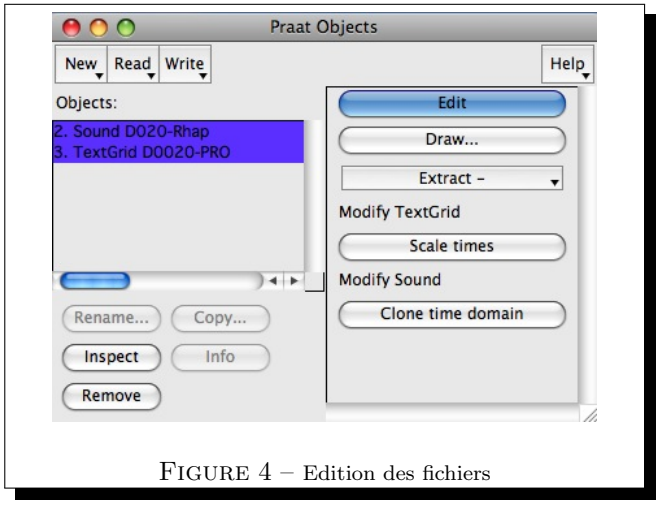


FIGURE 4 – Edition des fichiers

Une nouvelle fenêtre s'ouvre. Il faut éditer l'ensemble sans afficher les informations acoustiques. Pour ce faire :

- Cliquez sur l'onglet *Spectrum* et décocher « show spectrogram » ;
- Cliquez sur l'onglet *Pitch* et décocher « show Pitch » ;

- Cliquez sur l'onglet *Intensity* et décocher « show Intensity » ;
- Cliquez sur l'onglet *Formants* et décocher « show Formants ».

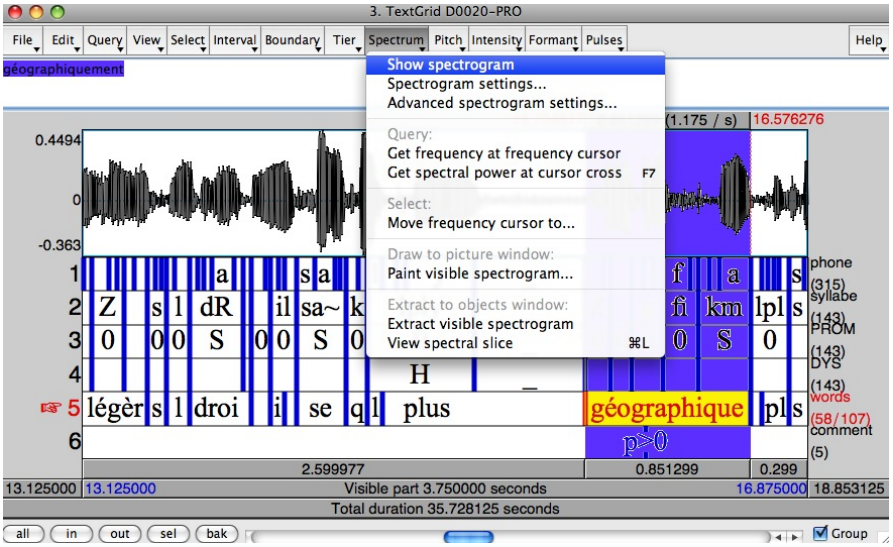


FIGURE 5 – On décoche « show spectrogram »

Tout au long de votre tâche de codage, vous aurez besoin de sélectionner des extraits du fichier en cours d'annotation : pour sélectionner un extrait il suffit simplement de placer le curseur sur le spectrogramme à l'endroit désiré, de cliquer et glisser vers la droite jusqu'à la fin de l'extrait désiré : l'extrait sélectionné apparaît alors en rose.

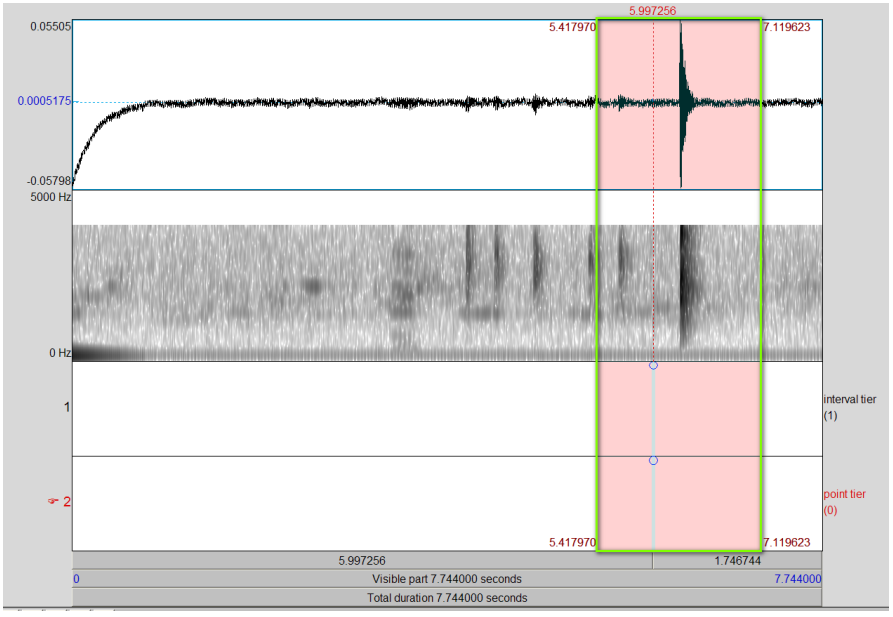


FIGURE 6 – Sélection

Pour travailler sur la fenêtre sélectionnée, il suffit alors de cliquer sur « sel » (ou Ctrl-N) afin de zoomer sur la sélection.

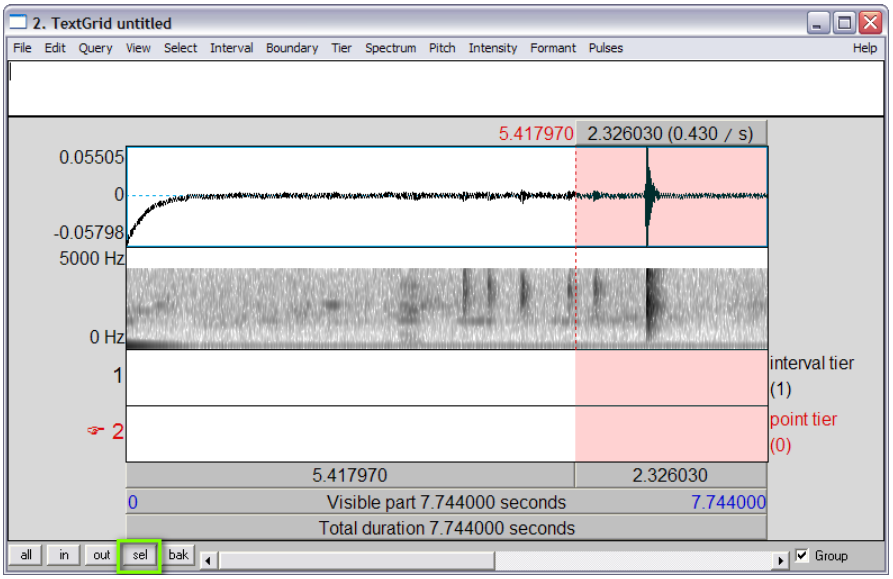


FIGURE 7 – On zoome avec « sel »

« in » (Ctrl-I) permet de zoomer aussi mais ne se restreint pas à la zone sélectionnée et « out » (Ctrl-O) permet de dézoomer.

### 2.1.4 COMMENT SE PRÉSENTE LE FICHIER ?

6 tires ou couches d'annotation sont visibles :

1. Phones
2. Syllabes
3. PROM
4. DYSF
5. Words
6. COMMENTS

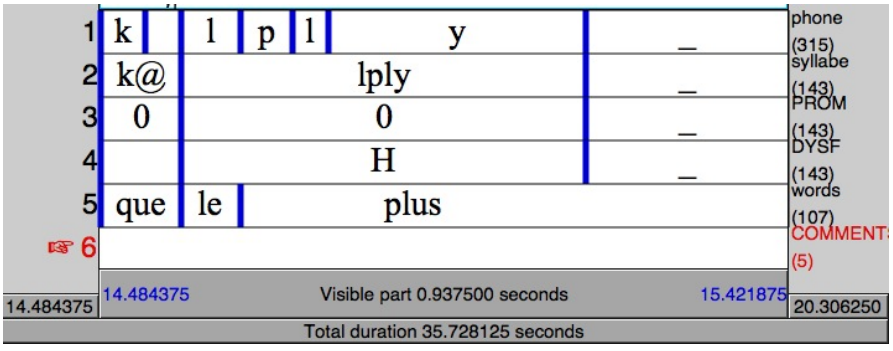


FIGURE 8 – Les différentes tires

Les tires *Phones (1)*, *Syllabes (2)* et *Words (5)* sont issues d'un alignement automatique. Les tires *PROM (3)* et *DYSF (4)* sont des duplicatas vidés de la tire syllabes, seuls demeurent remplis les intervalles que vous n'aurez pas à annoter, soit parce que ce sont des pauses (« _ »), soit parce qu'il s'agit de segments incompréhensibles, inexploitables pour l'annotation prosodique, soit parce qu'il s'agit de chevauchements qu'on ne peut pas non plus traiter prosodiquement. Pour tous ces contextes, vous devez voir apparaître le symbole % . Si ce n'est pas le cas (on peut retrouver notamment les caractères *** ou xxx), il est impératif de le signaler dans le tire *COMMENTS*.

A noter que la dénomination des tires donnée ici est purement indicative, il peut y en avoir d'autres, notamment une tire orthographique segmentée en unités de taille et de nature variable selon les objectifs de l'utilisateur (proposition, syntagme, etc) et les contraintes imposées par l'outil d'alignement automatique, l'exemple proposé ici vous permet donc de visualiser une configuration possible.

La transcription phonétique dans les tires phone et syll est effectuée en caractères SAMPA (voir le tableau ci-dessous pour les correspondances API-SAMPA).

| SAMPA | API | Exemples |
|---|---|---|
| p | p | *pont* |
| b | b | *bon* |
| t | t | *temps* |
| d | d | *dans* |
| k | k | *coût, quand, koala* |
| g | g | *gant* |
| f | f | *femme* |
| v | v | *vent* |
| s | s | *sans, dessus, cerise* |
| z | z | *zone, rose* |
| S | ʃ | *champ* |
| Z | ʒ | *gens, jambon* |
| j | j | *ion* [jo~] |
| m | m | *mont* |
| n | n | *nom* |
| J | ɲ | *oignon* |
| N | ŋ | *camping* |
| l | l | *long* |
| R | ʁ | *rond* |
| w | w | *quoi* [kwa] |
| H | ɥ | *juin* [ZHe~] |

FIGURE 9 – Correspondance API - SAMPA pour les consonnes du français

| SAMPA | API | Exemples |
|---|---|---|
| i | i | *si* |
| e | e | *Blé* |
| E | ɛ | *seize* |
| a | a | *patte* |
| A | ɑ | *pâte* |
| O | ɔ | *comme* |
| o | o | *gros* |
| u | u | *doux* |
| y | y | *du* |
| 2 | ø | *deux* |
| 9 | œ | *neuf* |
| @ | ə | *justement* |
| e~ | ɛ̃ | *vin* |
| a~ | ɑ̃ | *vent* |
| o~ | ɔ̃ | *bon* |
| 9~ | œ̃ | *brun* |

FIGURE 10 – Correspondance API - SAMPA pour les voyelles du français

### 2.1.5 QUELQUES RACCOURCIS

Afin de faciliter la navigation dans le fichier, il est possible d'utiliser les raccourcis suivants :

- Alt + Up/Down pour changer de tire
- Alt + flèche gauche/ flèche droite pour reculer, avancer dans les intervalles
- PageUp/PageDown pour avancer/reculer dans le signal (de 80% de la partie visible)
- Ctrl + I pour zoomer dans la zone affichée à l'écran (correspond à *in* dans Praat)
- Ctrl + N pour zoomer seulement sur la zone sélectionnée (correspond à *sel* dans Praat)
- Ctrl + O pour dézoomer (correspond à *out* dans Praat)

- Ctrl + S pour sauver le textgrid
- Ctrl + F pour chercher une motif dans une tire
- Ctrl + G pour chercher à nouveau un motif précédemment recherché
- flèche droite : avancer dans le fichier
- TAB : lire un son sélectionné (le son est surligné en bleu)
- ESC/ECHAP : arrêter un son

## 2.2 Le logiciel Analor

### 2.2.1 Comment installer Analor ?

Le logiciel est téléchargeable sur le site du Lattice à l'adresse : http://www.lattice.cnrs.fr/Analor
Le fichier exécutable s'appelle "analor.jar". C'est un programme java, qui fonctionne avec n'importe quel système d'exploitation quand le moteur java est présent (c'est le cas pour la plupart des installations).
Pour lancer Analor, il faut double-cliquer sur le fichier "analor.jar", ou exécuter la ligne de commande "java –jar analor.jar".
Si cela ne fonctionne pas, c'est sans doute que le moteur java n'est pas installé. Il s'appelle JRE (Java Runtime Environment) ou aussi JVM (Java Virtual Machine). Il est téléchargeable gratuitement à l'adresse : http://www.java.com/fr/download

### 2.2.2 À quoi sert Analor ?

Analor est un logiciel d'annotation et d'analyse de l'oral. Ce n'est pas un logiciel de traitement du signal : il ne calcule pas la fondamentale, l'intensité, etc. Il faut donc d'abord traiter ses documents par un logiciel qui offre ces fonctionnalités (comme Praat, WinPitch, IRCAM ?, etc.) avant de pouvoir utiliser Analor. Les principales fonctionnalités d'Analor sont les suivantes :

- Importation et exportation de données au format Praat
- Chargement et écoute du son
- Visualisation de la fondamentale, de l'intensité, des tires d'annotation
- Création et/ou modification de tires d'annotation
- Alignement d'une tire d'annotation sur une autre
- Corrections manuelles spécifiques de la fondamentale (sauts d'octave et suppression de valeurs parasites)
- Lissage de la fondamentale
- Détection automatique des coupures de période suivant des critères paramétrables
- Détection automatique de syllabes proéminentes suivant des critères paramétrables

### 2.2.3 Le menu de la fenêtre principale

#### 2.2.3.1 Documents

Ce menu permet d'importer et d'exporter des documents au format Praat (données TextGrid, PitchTier, Pitch, IntensityTier, Intensity, ou un ensemble de ces données sous forme de Collection). Il permet aussi de charger un fichier son (format .wav). Enfin il permet d'enregistrer et de charger des documents au format Analor (extension : ".or"). *Il est fortement recommander de sauvegarder régulièrement et fréquemment ses données (raccourci clavier : Ctrl S), pour ne pas perdre tout son travail en cas de bug...*

#### 2.2.3.2 Affichage

Ce menu permet d'ajuster de manière précise la portion visible du document, si nécessaire, mais le plus souvent, la glissière horizontale et les deux boutons de zoom (+ et –) dans le bas de la fenêtre suffisent pour se déplacer dans le document. Il permet aussi de spécifier les courbes et les tires à visualiser (le menu contextuel sur le nom d'une tire permet aussi de la masquer, mais seul ce menu-ci permet de la démasquer). Enfin, c'est aussi ce menu qui permet de créer une nouvelle tire.

#### 2.2.3.3 Statistiques

Ce menu permet d'afficher un histogramme des valeurs de la fondamentale ou de l'intensité. Un menu contextuel sur l'histogramme permet de sélectionner dans la fenêtre principale les portions de courbe ayant les valeurs choisies. Les flèches gauche et droite du clavier permettent alors de se déplacer de portion en portion.

#### 2.2.3.4 Paramétrage

Ce menu permet de gérer tous les paramètres utilisés dans les calculs (lissage, périodes, proéminences). Il s'agit d'une part des noms des tires (tire syllabique, phonémique, etc.) et des marques (de pause, d'hésitation, etc.) utilisées dans les calculs, et d'autre part des paramètres de calcul proprement dit (valeurs de seuil, etc.). On peut sauvegarder et charger des fichiers de paramètres (extension : ".orp"), ainsi que créer un fichier texte listant les valeurs de tous les paramètres)

#### 2.2.3.5 Stylisations

Ce menu permet de créer une courbe lissée approximant la fondamentale. Le lissage utilise deux paramètres de calcul : la fenêtre temporelle de moyennage et l'écart maximal autorisé entre la courbe lissée et la fondamentale. On peut sauvegarder au format Praat la courbe lissée. Pa ailleurs, il permet aussi d'importer, de visualiser et d'exporter (en format texte) des prosogrammes calculés dans Praat avec le script de Piet Mertens.

#### 2.2.3.6 Périodes

Ce menu permet de détecter automatiquement quelles pauses marquent des coupures de périodes en fonction de 4 critères (durée de la pause, amplitude du geste final avant la pause, saut après la pause, présence d'une hésitation ; voir Lacheret et Victorri ? pour le détail de l'algorithme utilisé). Un menu contextuel (clic droit) sur un losange situé sous le graphique à l'emplacement de chaque pause permet de visualiser tout le détail des calculs concernant cette pause, ainsi que de corriger manuellement la décision prise automatiquement. Les flèches gauche et droite du clavier permettent de se déplacer de pause en pause. D'autre part, se crée automatiquement une tire "periode" (ce nom est modifiable) qui comporte tous les résultats (les périodes sont numérotées, et des symboles à l'emplacement des pauses indiquent comment la décision a été prise).
Ce menu permet aussi d'exporter sous forme de fichier texte (lisible avec n'importe quel tableur) tous les calculs.

#### 2.2.3.7 Proéminences

Ce menu permet de calculer automatiquement le caractère plus ou moins proéminent de chaque syllabe, en fonction de quatre critères (durée relative, hauteur relative, montée intrasyllabique, présence d'une pause après la syllabe ; voir Avanzi, Lacheret et Victorri ? pour les détails de l'algorithme utilisé). Un menu contextuel (clic droit) sur les informations inscrites pour chaque syllabe sous le graphique permet de visualiser tout le détail des calculs concernant cette syllabe. D'autre part se crée automatiquement une tire indiquant les syllabes faiblement et fortement proéminentes (le nom de la tire et les marques de proéminence sont modifiables). Ce menu permet aussi d'exporter sous forme de fichier texte (lisible avec n'importe quel tableur) tous les calculs.

### 2.2.4 Les autres fonctionnalités

#### 2.2.4.1 Sélection d'un intervalle temporel et écoute du son

On peut sélectionner un intervalle temporel sur la bande son (en maintenant le bouton gauche de la souris enfoncé), sur le graphique (idem, ou clic sur une portion de la fondamentale), ou sur une tire (clic dans un intervalle ; en maintenant enfoncée la touche majuscule, on peut sélectionner plusieurs intervalles à la fois).
Une fois un intervalle temporel sélectionné, on peut entendre le son sur cet intervalle en cliquant dans

la partie sélectionnée de la bande son. Si l'on clique sur la partie non sélectionnée de la bande son (et bien sûr aussi quand il n'y a pas de sélection), on entend le son sur toute la durée de la fenêtre visible. Enfin, en cliquant sur le titre "Son" (à gauche de la bande son), on entend le son sur toute la durée du document. On peut recliquer au même endroit pour arrêter l'écoute.

#### 2.2.4.2 Modifications de la fondamentale

Après avoir sélectionné une portion de la fondamentale (soit par sélection d'un intervalle temporel –cf. ci-dessus, soit à partir de l'histogramme –cf. menu Statistiques), on dispose d'un menu contextuel (clic droit sur la portion sélectionnée de la fondamentale) qui permet de supprimer cette portion ou de la déplacer d'une octave.

#### 2.2.4.3 Opérations sur les tires d'annotation

On dispose d'un menu contextuel pour chaque tire d'annotation (clic droit sur le nom de la tire située sur la gauche) qui permet diverses opérations (renommage, déplacement, masquage, etc.). Notamment, on peut ainsi « solidariser » une tire avec une autre : cela consiste à faire en sorte que chaque borne de la première soient systématiquement alignée avec une borne de la seconde : on peut ainsi, par exemple, solidariser une tire syllabique sur une tire phonémique. Ce menu contextuel permet aussi d'exporter en format texte les annotations d'une tire avec leurs bornes temporelles, et, dans le cas où cette tire est solidarisée avec une autre, le contenu de cette autre tire.
Un clic gauche sur le nom d'une tire permet de sélectionner cette tire comme tire principale : les bornes des annotations de cette tire sont alors reproduites sur le graphique.
On peut déplacer une borne d'une annotation dans une tire en faisant glisser avec la souris la « poignée » située dans la petite bande située en haut de la tire (la poignée est la partie plus épaisse de la borne ; cette partie est grisée quand la tire est solidarisée avec une autre : cela signifie qu'on ne peut déplacer la borne que pour l'aligner avec une autre borne de cette autre tire). On peut aussi créer ou supprimer une borne en double-cliquant à l'emplacement voulu dans l'espace réservé aux poignées. Enfin un clic droit dans l'espace réservé aux annotations permet d'éditer une annotation de cette tire.

# 3 Repères définitoires

**Principe** : les objets à coder reposent uniquement sur des bases perceptives, ils sont de deux types :

## 3.1 Proéminences et principe de contraste

Une syllabe peut **se détacher dans le flux verbal *par rapport à ce qui précède*** et ainsi contraster avec son environnement. On dit qu'elle se détache perceptivement comme une figure sur un fond sonore : cette syllabe n'a pas le même poids que son entourage. Les raisons (phonétiques et linguistiques) de ce détachement peuvent être multiples et la position de la proéminence très variable. Peu importe pour vous, ce n'est pas votre problème. Dès que vous avez l'impression de percevoir un contraste sur une syllabe donnée, vous l'annotez proéminente. Cependant, il convient de ne pas oublier que **les syllabes proéminentes ne sont pas proéminentes uniquement grâce à la hauteur !** D'autres paramètres peuvent rentrer en ligne de compte. Notamment : un ton perçu dans l'infra-bas suivi d'une pause, un allongement syllabique, des variations de qualité vocale (voix craqueé ou « creaky voice »), et a fortiori une combinaison de ces différents indices jouent un rôle dans la perception de la proéminence.
Des exemples : cf. l'allongement syllabique dans les figures 11 et 12.

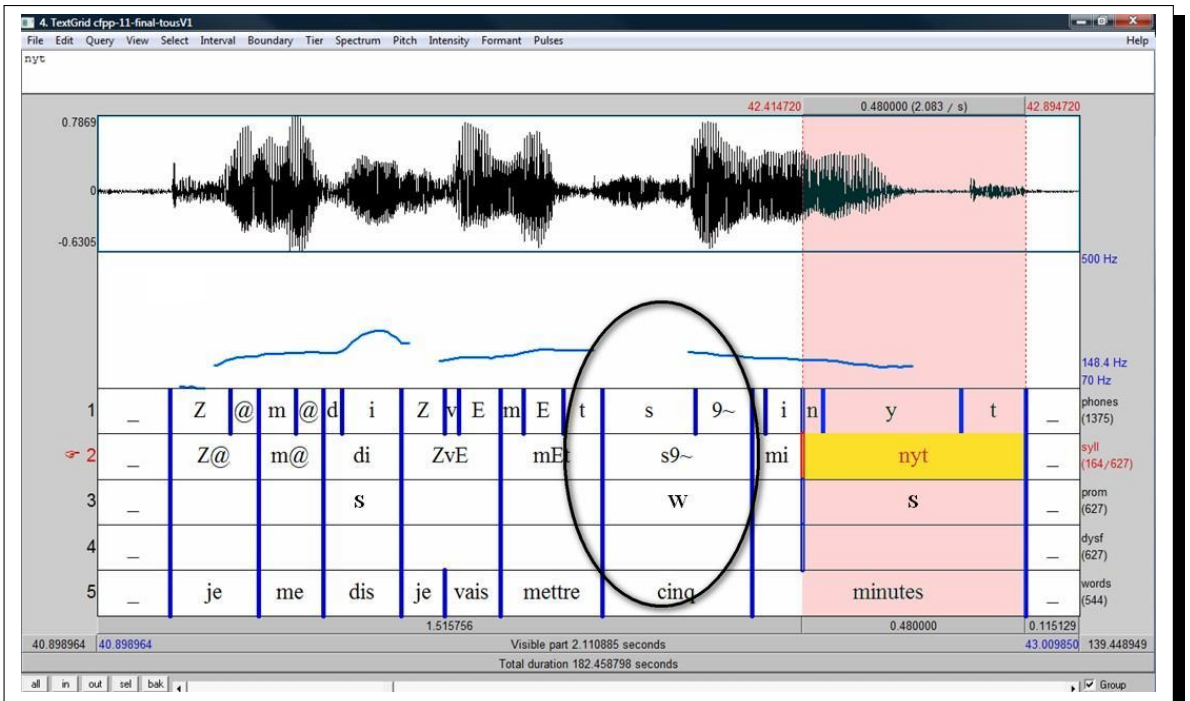


FIGURE 11 – D'autres paramètres que la hauteur rentrent en compte (exemple 1)

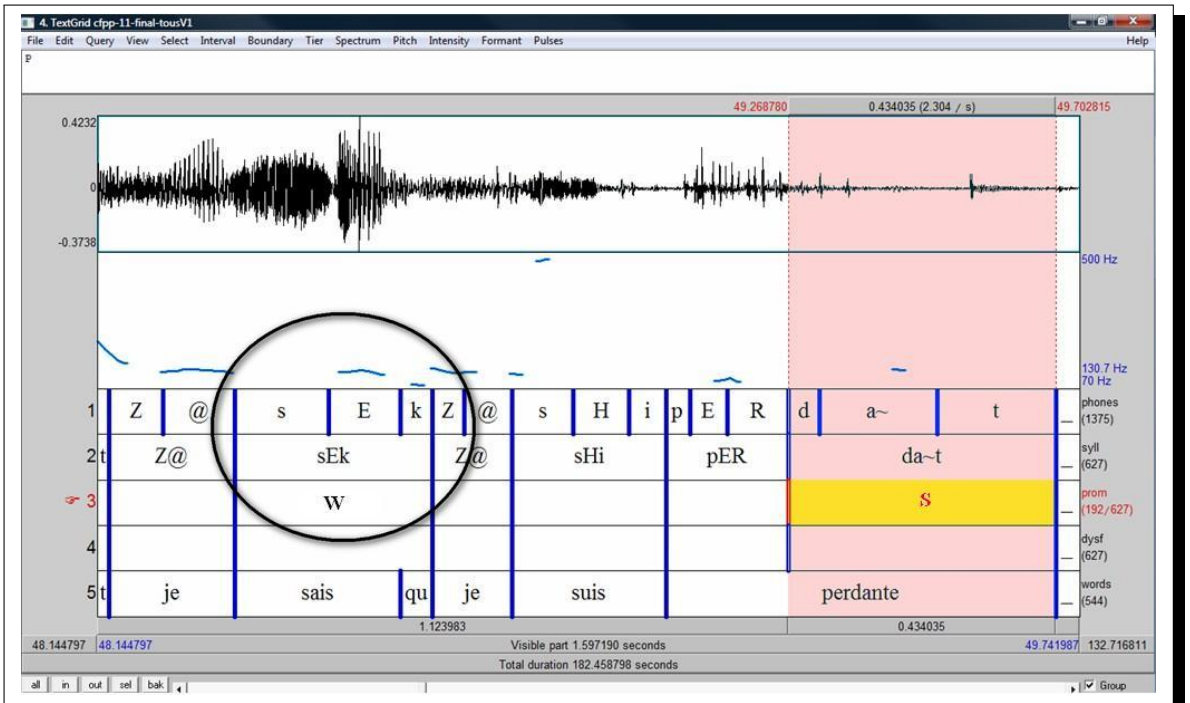


FIGURE 12 – D'autres paramètres que la hauteur rentrent en compte (exemple 2)

La perception des proéminences est *continue*, et non catégorielle. En d'autres termes, une syllabe ne peut se définir en proéminente ou non-proéminente. Il existe *différents degrés de proéminence* allant du non proéminent au très proéminent. Pour rendre compte de manière rudimentaire de ce processus continu, nous utilisons *3 niveaux* : **non proéminent** (marqueur « 0 » pour « nul »), **proéminent** (marqueur « W » pour « weak ») et **très proéminent** (marqueur « S » pour « strong »).

### 3.2 Disfluence et principe de non linéarité

On définira la disfluence comme un élément qui brise le déroulement syntagmatique dans la chaîne parlée, une sorte de *trébuchement vocal*. Il peut prendre différentes natures et correspond souvent à un allongement syllabique excessif associé au travail de formulation en cours. Il peut s'agir aussi d'une répétition de morphème et de bien d'autres choses encore. Toute syllabe perçue comme disfluente sera repérée par une balise spécifique, notée « H ».

### 3.3 Les périodes, unités prosodiques majeures

## 4 Consignes pour le codage des proéminences et des disfluences

### 4.1 Les conditions d'un bon codage

Nous donnons ci-dessous un ensemble de règles à respecter pour stabiliser autant que faire se peut le codage et limiter la variation inter-annotateur qui dériverait d'un manque de la clarté dans l'explicitation des consignes et dans le cadrage méthodologique.
Cette variation doit être limitée au maximum surtout si l'on veut travailler sur des annotations multiples et évaluer le consensus inter-annotateurs. En d'autres termes des règles explicites pour l'annotation sont essentielles pour les résultats fournis ultérieurement par le calcul du taux d'accord inter-annotateur : un mauvais taux d'accord provient souvent d'un manque de rigueur méthodologique.

**Les conditions à respecter** :

- **Utiliser un casque de qualité** pour le codage ;
- **Coder dans un environnement calme** ;
- Toujours coder en **respectant le protocole** à toujours avoir sous la main et à consulter au moindre doute ;
- Annoter chaque fichier **en suivant les instructions à la lettre et en ne dépassant pas un temps de codage de 30 minutes**. Il est conseillé de ne **pas coder plus de 30 minutes** au risque de « saturer » et de ne plus rien entendre. Il vous faudra, quand cela est nécessaire, morceller votre tâche de codage en étapes de 30 minutes. La qualité sonore d'un fichier, les stratégies vocales du ou des locuteurs sont autant de facteurs pouvant compliquer votre tâche de codage : il faut en moyenne 20 minutes pour annoter un fichier de 60 secondes, mais ce temps peut être doublé voire triplé selon les fichiers ;
- Il est **strictement interdit** de toucher aux barres de segmentation : en cas de déplacement ou effacement involontaire il suffit de le reporter dans la tire *COMMENTS*, ce conseil est aussi valable lorsqu'on repère une erreur de segmentation ;
- Reporter les erreurs et questions posées dans la tire *COMMENTS* ;
- **Enregistrer régulièrement** son travail sous un autre nom que le fichier original tant que le codage n'est pas terminé pour ne pas risquer de le perdre.

### 4.2 Les étapes de codage

Avant de procéder à tout codage, il est nécessaire d'**écouter quelques secondes du fichier** : le but est de s'imprégner des différentes caractéristiques de chaque fichier. Il est important de prendre conscience que différents locuteurs ne s'expriment pas de la même manière, aussi **il ne faut coder qu'en fonction d'un locuteur sans prendre en compte les caractéristiques d'un autre locuteur** (que ce soit dans le fichier ou hors du fichier). Nous possédons tous nos propres caractéristiques prosodiques qui sont à prendre en compte dans une tâche de codage : certaines personnes vont parler de façon assez monotone quand d'autres vont accentuer presque chaque mot qu'elles prononcent ; **il est donc indispensable d' écouter un peu chaque locuteur d'un fichier pour avoir un aperçu de son style vocal et calibrer intuitivement sa perception sur la voix du locuteur en cours.** Ainsi, pour un locuteur qui a un registre restreint, une variation de perception tonale peu importante peut néanmoins donner lieu à la perception d'une proéminence forte ; ça ne sera pas le cas pour un locuteur qui distribue ses variations mélodiques sur un registre large.
**Cette écoute préalable permettant la mise en place de stratégie d'annotation est indispensable, elle permet d'éviter le *sous-codage* ou le *sur-codage*.**

### 4.2.1 Principes généraux de codage

Le codeur procède en plusieurs étapes pour procéder à l'annotation prosodique du fichier :

- Le codeur traite le fichier à analyser de gauche à droite ;
- Chacune des étapes de codage se fait de façon séquentielle : *empan par empan*[3] et non sur tout le corpus. Sur un segment donné (un groupe de cinq secondes maximum, cf. *infra* 4.2.2), on traite :
  - d'abord les intervalles de confiance (proéminences fortes (« S »),
  - ensuite les proéminences et les disfluences (« W » et « H »).
- Vous disposez d'**une tire commentaire** :
  - un symbole pour les erreurs d'alignement ;
  - un autre pour les erreurs de transcription.

  Cette tire est une « point » tire, placez la barre de segmentation à l'endroit où vous avez rencontré une erreur et attribuez lui les marqueurs suivants suivis d'un commentaire le cas échéant :

  1. Erreur d'alignement
  2. Erreur de transcription syllabique
  3. Erreur de transcription orthographique
  4. Erreur dans la tire PROM ou DYS : présence d'un symbole alors qu'il ne devrait rien y avoir ou encore marqueur erroné (*** vs %)
  5. Autre avec commentaire adjoint

  Cette tire est *très importante*, elle vous permettra de *mettre en évidence les erreurs rencontrées*, mais aussi vos *questionnements* divers, vos *hésitations* sur une annotation : **il ne faut pas hésiter à utiliser cette tire pour consigner les retours d'expérience des codeurs**.

Figure 13 – Exemple de commentaires

**Nombre d'écoutes** : quand le doute s'installe, ne pas réécouter trop de fois le segment sonore au risque de ne plus rien entendre : pour chacune des deux étapes précédemment citées, identification des balises de confiance (proéminences fortes) puis proéminences moins fortes et disfluences, le nombre d'écoute est limité à *trois fois maximum*.

### 4.2.2 Etape 0 : Choisir sa fenêtre de codage

Une expérience de perception, on le sait, (qu'elle soit visuelle ou auditive) est bornée dans un intervalle donné pour pouvoir construire ses repères, sinon, on a le sentiment d'être face à un tout amorphe au sein duquel on ne peut rien percevoir du tout et où il est donc impossible d'isoler des objets spécifiques.

3. Ici l'empan correspond à l'extrait temporel en cours de codage, c'est-à-dire la fenêtre de traitement choisie (elle varie entre 3 et 5 secondes selon les contextes)

En conséquence, l'expérience conduite ici sera bornée sur un intervalle temporel [4] ou empan de codage de **3 à 5 secondes max**.
L'empan de codage correspond à un segment linguistique avec un début et une fin. La question pour le codeur est donc : comment fixer la fin de ce segment ?
La fin d'un segment correspond à une unité linguistique « achevée » : autant que possible, alignez l'intervalle temporel avec la fin d'un groupe de souffle (identifiable par une pause), sinon avec la fin d'une unité de sens (un « énoncé »). Une coupure ne doit pas générer un sentiment d'incomplétude qu'il s'agisse d'une incomplétude de mot ou de groupe (ex. j'ai acheté _), c'est-à-dire que l'empan de codage commence au début d'un groupe et finit à la fin d'un groupe quelconque.
Les pauses de groupe de souffle constituent de bons indices pour ce faire, mais il n'y en a pas toujours. Dans l'exemple ci-dessous, la première pause tombe au bout de 5 secondes et correspond à une unité sémantiquement achevée :

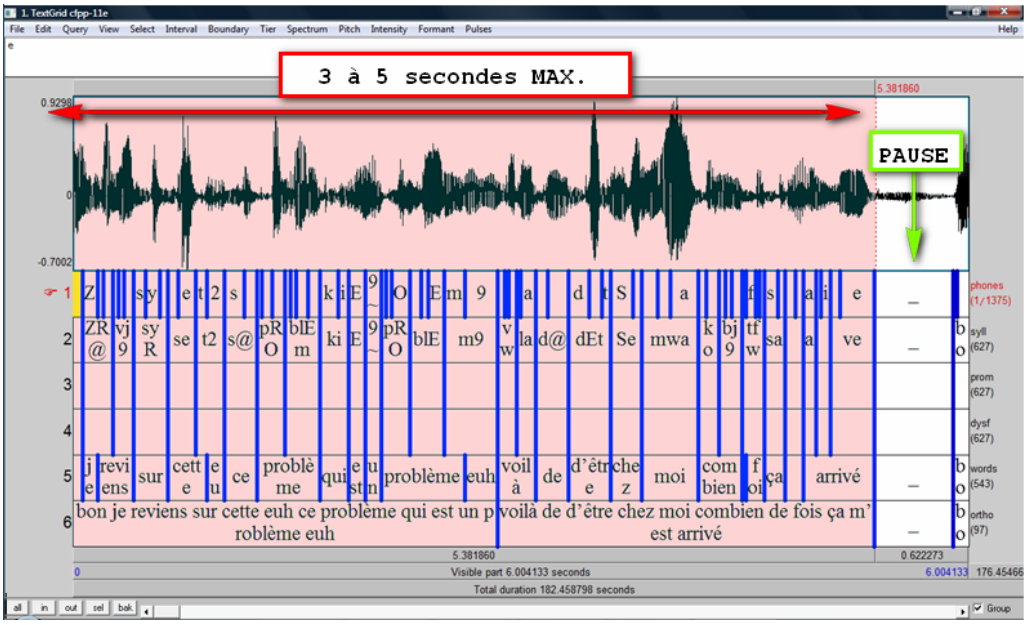


FIGURE 14 – Exemple de fenêtre de codage

Mais certaines unités inter-pausales [5] sont plus longues que 3/5 secondes, auquel cas on essaie de couper à l'intérieur de cette macro-unité, à la fin d'un « groupe » de sens. Dans l'exemple ci-dessous, on se sert de la disfluence (« tu tu ») pour la délimitation de la borne droite de la macro-unité :

4. L'intervalle temporel correspond à un segment sur lequel vous allez travailler avant de passer au segment suivant et ainsi de suite.
5. unités contenues entre des pauses

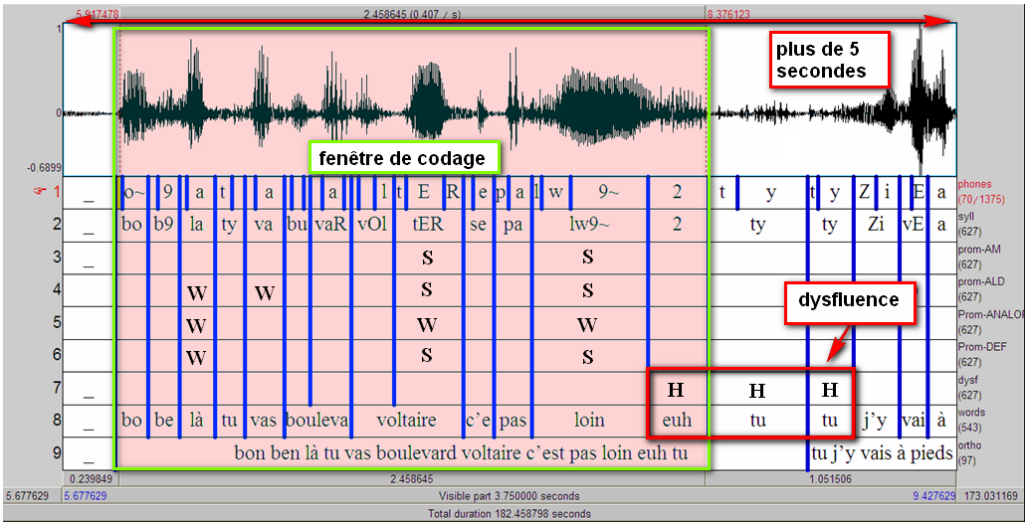


FIGURE 15 – Coupure au niveau de la disfluence en cas d'empan trop long

Cependant, toutes les pauses ne sont pas associées à la fin d'une macro-unité, certaines pauses peuvent couper un groupe de sens, notamment dans les contextes de disfluence. Il faut donc faire ***attention*** de ne pas couper après de telles pauses mais à la pause suivante qui démarque bien notre unité de sens. Ainsi, dans la séquence « bon autre chose on a nos # nos chaises qui ont besoin de refaire#, le premier symbole '#' indique une pause non démarcative liée à une disfluence, le second est associé à une pause démarcative, c'est donc ce dernier qui va nous servir pour fixer l'empan.

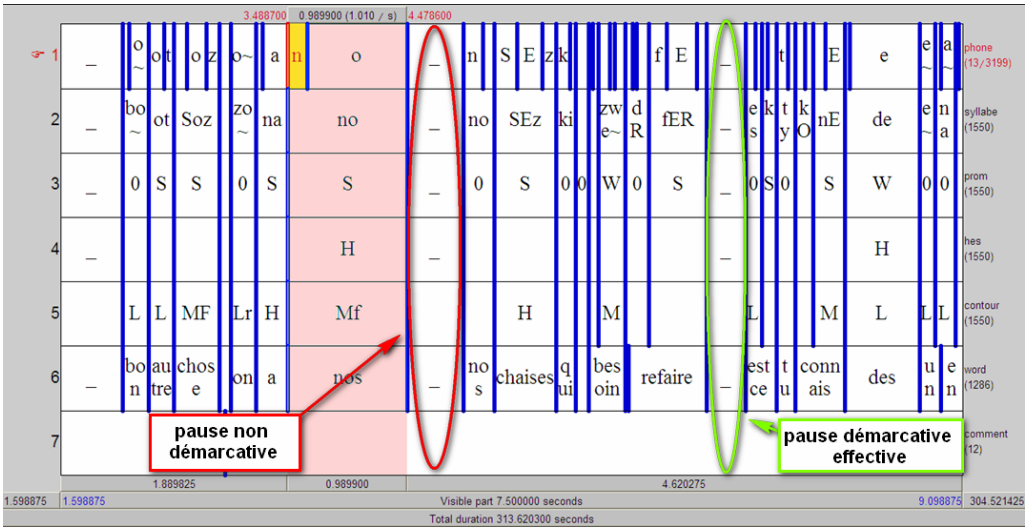


FIGURE 16 – Différence entre pause démarcative et pause non démarcative

Cependant, il est tout à fait possible de se retrouver confronté à un cas où il est impossible de segmenter comme précédemment expliqué. Dans un tel cas de force majeure, vous fixerez votre intervalle de confiance au-delà de la limite imposée de 5 secondes *sur un segment achevant le groupe* :

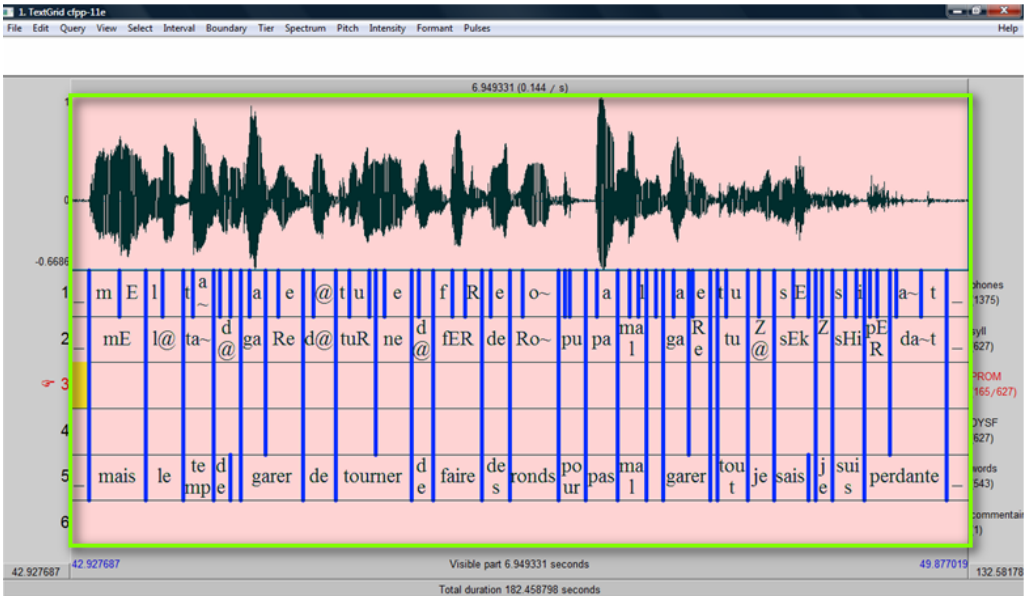


FIGURE 17 – Intervalle de confiance de plus de 5 secondes

Enfin, *la segmentation des tours de parole* peut aider à délimiter les macro-unités de cette première étape. On ne code pas en même temps deux locuteurs, mais on écoute les portions de signal d'un locuteur, puis celle d'un autre. Il faut donc prendre en compte cette alternance de tour dans l'intervalle de confiance choisi :

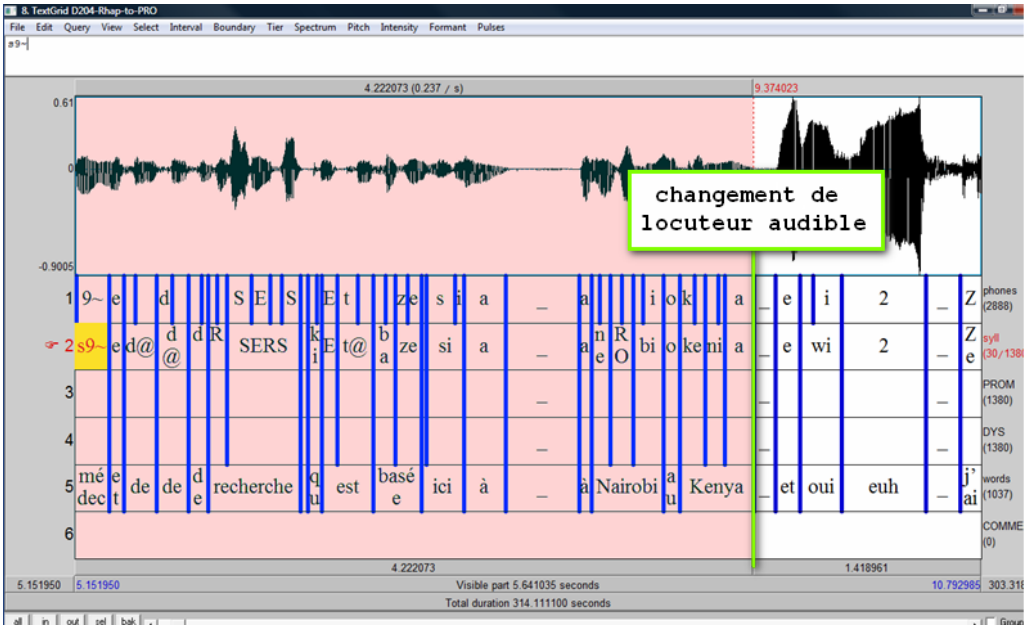


FIGURE 18 – Changement de locuteur

### 4.2.3 ETAPE 1 : CODAGE DES PROÉMINENCES FORTES : INTERVALLE DE CONFIANCE OU EMPAN DE CODAGE

**Les syllabes fortement proéminentes vont servir de point d'ancrage pour la suite : elles vont permettre de délimiter des unités de rang inférieur (des intervalles internes aux segments de 5 secondes), dans lesquelles on va coder d'autres événements.**

Une fois la macro-unité identifiée lors de l'étape 0, on la sélectionne [6] et on zoome dedans ( Ctrl + N )

6. Se reporter à la présentation de Praat si on ne sait pas comment sélectionner.

afin de ne pas déborder de la zone à traiter qui apparaît alors en rose.

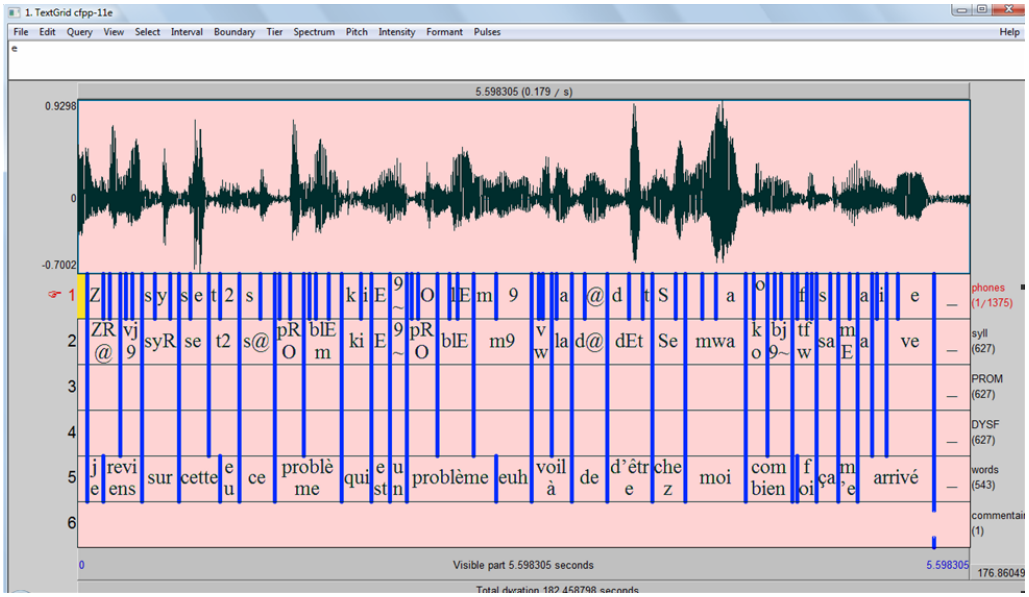


FIGURE 19 – Sélection de la zone

Maintenant, il faut jouer l'extrait sélectionné en cliquant sur la bande correspondante ou en utilisant TAB.
On note « S » les intervalles syllabiques perçus **fortement proéminents**, c'est-à-dire les syllabes qui se détachent de leur environnement lors des premières écoutes comme précédemment expliqué. **Le degré de la proéminence (W/S) doit être fixé relativement à ce qui suit et non relativement à ce qui précède.**
En cas de doute après le nombre maximal d'écoutes, il convient de laisser cette syllabe vide dans un premier temps puis de l'annoter faible « W » lors de l'étape suivante (étape 2). Si le codeur hésite entre proéminence forte et faible, il lui faut se rappeler que les proéminences fortes doivent se détacher suffisamment de leur environnement pour ne pas laisser la place au doute. Il est possible aussi d'annoter la syllabe avec un « ? », la décision pourra alors être prise lors de l'étape de finalisation du fichier.
Enfin, il est important de se rappeler qu'**une fin de groupe de souffle n'est pas nécessairement marquée par une proéminence forte** , dans certains cas (tels que les parenthèses) comme dans l'exemple ci-dessous, la proéminence peut être faible voire inexistante, vous la notez alors « W ».

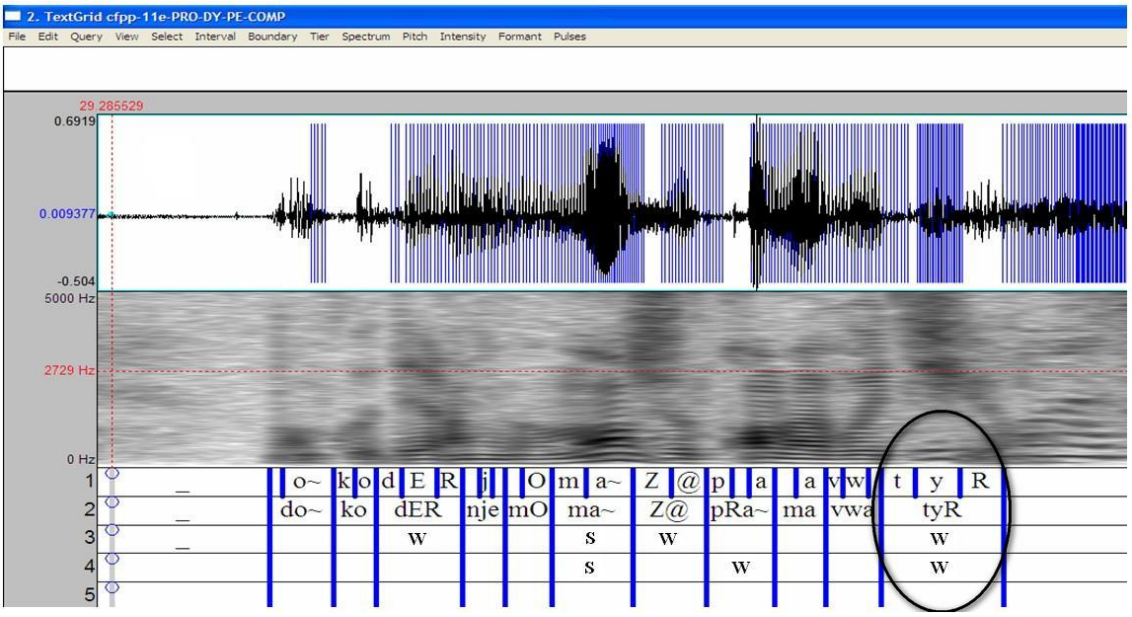


FIGURE 20 – Proéminence faible en fin de groupe de souffle

### 4.2.4 ETAPE 2 : CODAGE DES PROÉMINENCES FAIBLES (« W ») ET DES DISFLUENCES (« H »)

On sélectionne le premier intervalle du segment qui est délimité à sa droite par une proéminence forte.

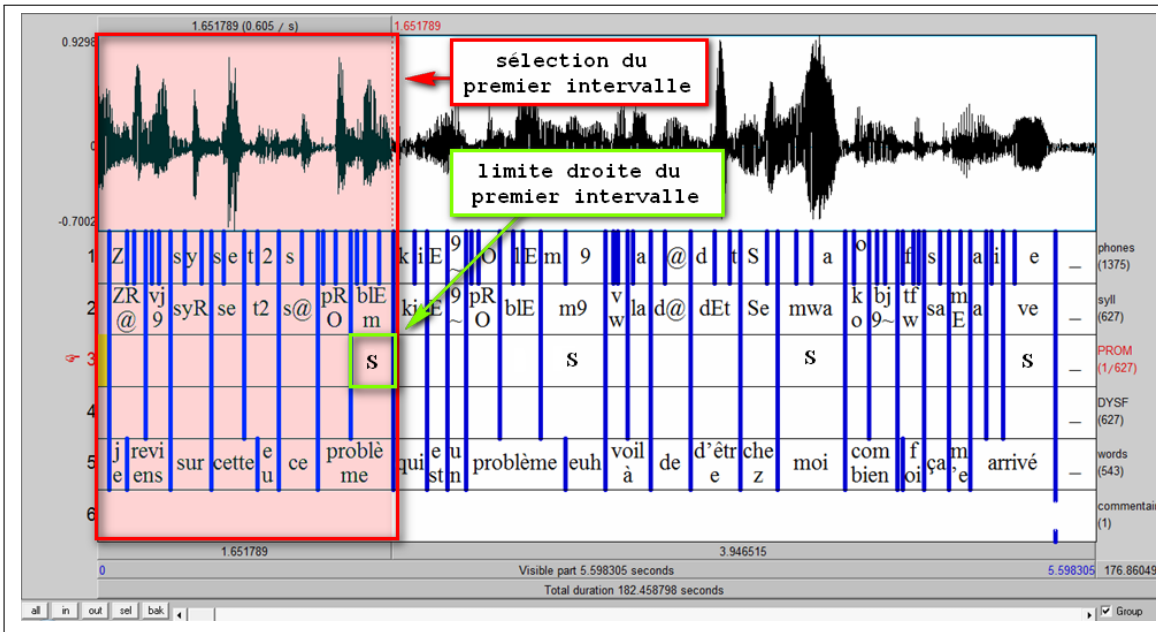


FIGURE 21 – Sélection du premier intervalle du segment délimité à sa droite par une proéminence forte

#### 4.2.4.1 CODAGE DES « W »

À l'intérieur de cet intervalle séparé par une proéminence forte, on note les syllabes perçues comme faiblement proéminentes avec le symbole « W » (en ne dépassant jamais le nombre maximal d'écoutes !). **Le degré de la proéminence (W/S) doit être fixé relativement à ce qui suit et non relativement à ce qui précède.** On ne code jamais de « S » lors de la seconde étape, mais seulement des « W ». On part du principe que c'est la première écoute qui est la plus importante (la perception des contrastes change selon la fenêtre temporelle choisie).

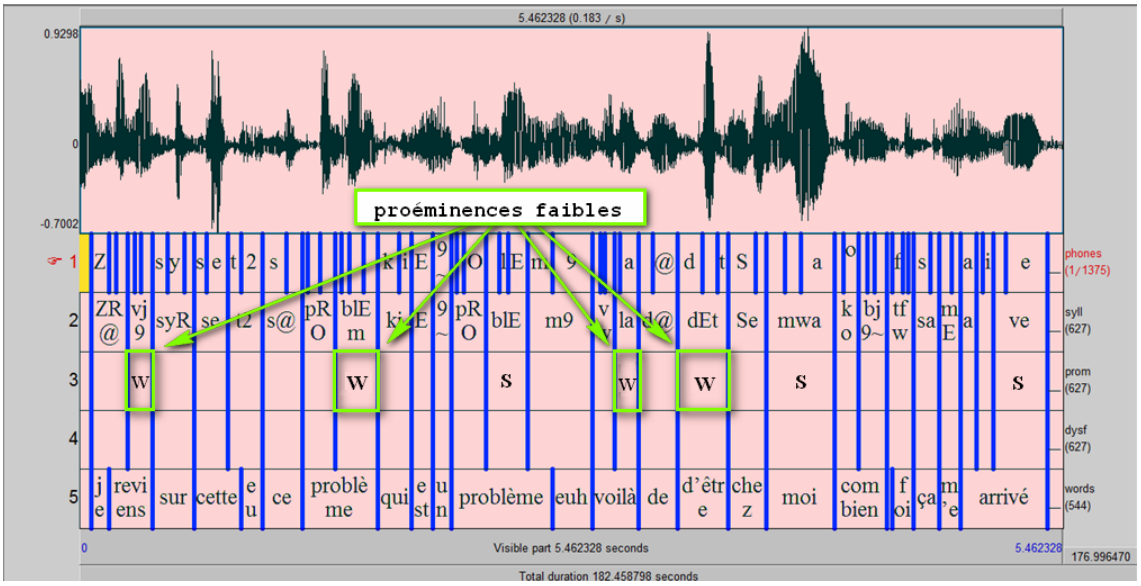


FIGURE 22 – Codage des proéminences faibles « W »

Ensuite, on laisse les autres cases vides et :

1. soit on annote les disfluences s'il y en a de ce passage
2. soit on continue sur l'empan suivant en sachant que pour l'annotation des disfluences, deux stratégies sont possibles :
   (a) elle est effectuée simultanément à l'annotation des proéminences : chaque intervalle marqué par une proéminence terminale « S », une fois les proéminences « W » annotées, fait l'objet de l'annotation des disfluences avant de passer à l'intervalle suivant.
   (b) elle est réalisée de manière autonome une fois la tire des proéminences remplie pour l'ensemble du fichier.

#### 4.2.4.2 CODAGE DES « H »

Une fois les proéminences faibles codées au sein de l'intervalle en cours, il faut coder les disfluences « H » dans la tire prévue à cet effet.
**Pour rappel, sont disfluentes les syllabes que le codeur considère comme relevant de l'hésitation (trébuchement vocal).** Comme pour les autres étapes, le nombre d'écoutes est *limité à 3.*

On note « H » les disfluences clairement identifiées, « ? » les segments pour lesquels un doute se fait sentir (ce type de contexte est essentiellement lié à la perception des allongements syllabiques et des sons /E/) ; on laisse vide les autres cases.

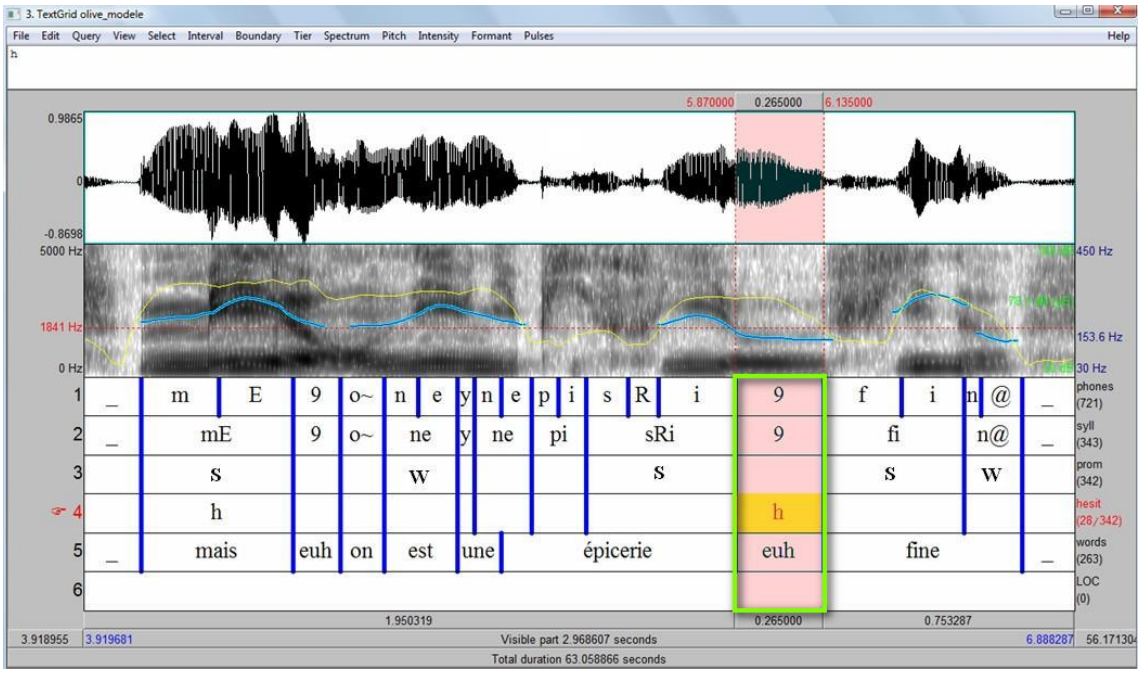


FIGURE 23 – Codage des disfluences « H »

## Stratégie/empan d'annotation pour les disfluences

Les cas de disfluences les plus classiques sont les suivants :

- Segments abandonnés, amorces de morphèmes : *c'est pas loin* ***tu*** *j'y vais*
- Segments repris/répétés : *c'est pas loin* ***tu tu*** *y vas*
- « Euh » d'hésitation
- Allongement syllabique excessif (ne correspondant pas à une démarcation structurale (*il y* ***a*** *: : : : :*)

**Ces phénomènes ne sont pas exclusifs, mais peuvent se combiner.** ***Comment les annoter ?***

S'il s'agit de disfluences simples (une seule syllabe disfluente), la syllabe disfluente est notée « H » dans la tire dédiée :

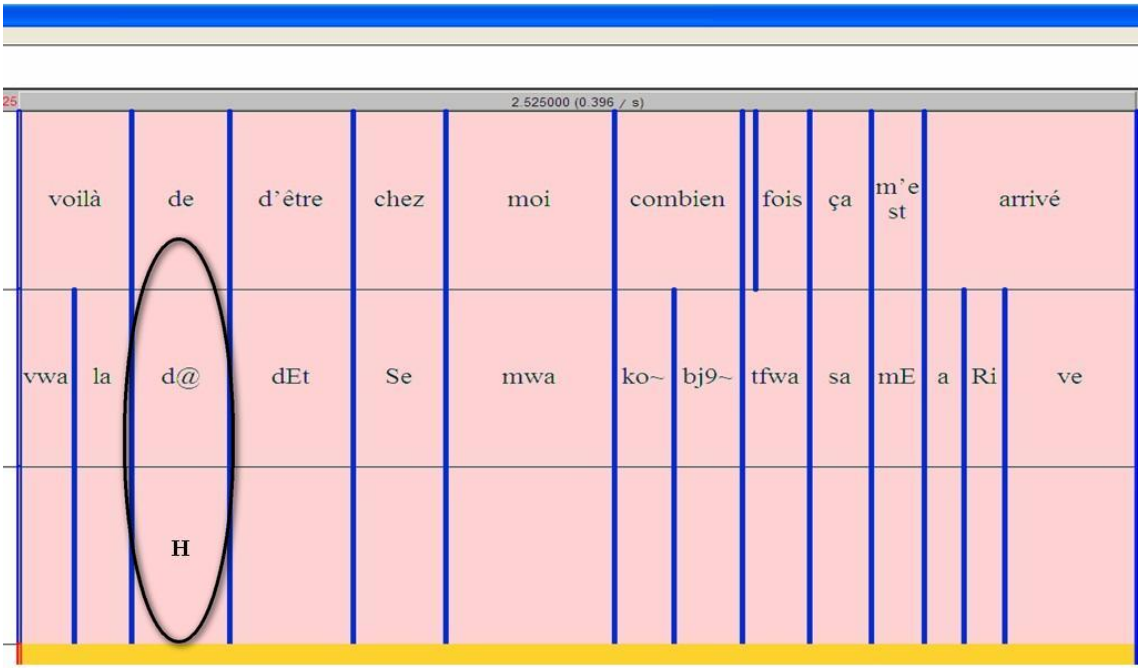


FIGURE 24 – « voilà de d'être chez moi »

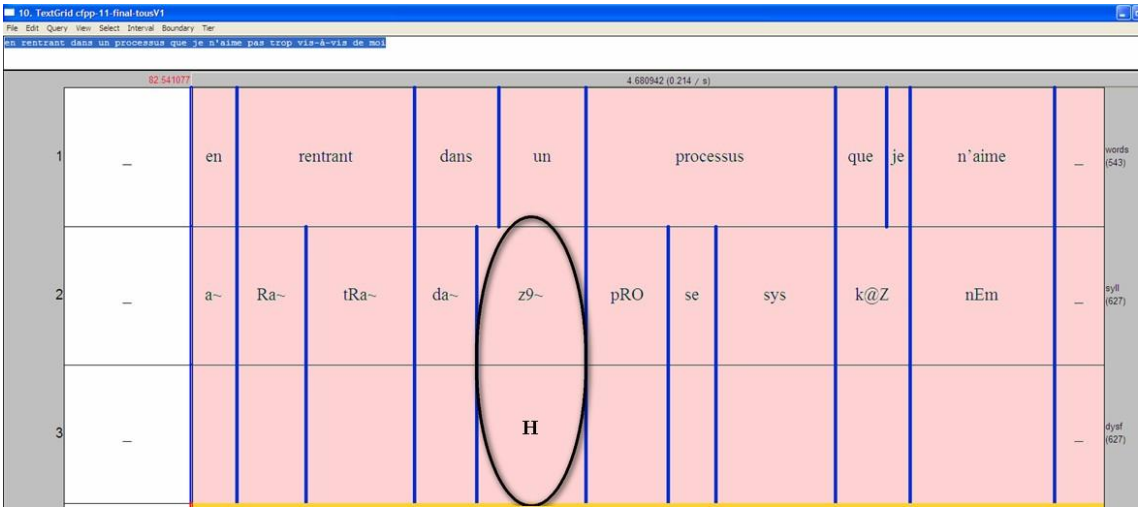


FIGURE 25 – « en rentrant dans un processus que je n'aime pas trop vis-à-vis de moi »

S'il s'agit de disfluences combinées (plusieurs intervalles qui se suivent), il faut alors coder toutes les syllabes comprises dans l'intervalle disfluent :

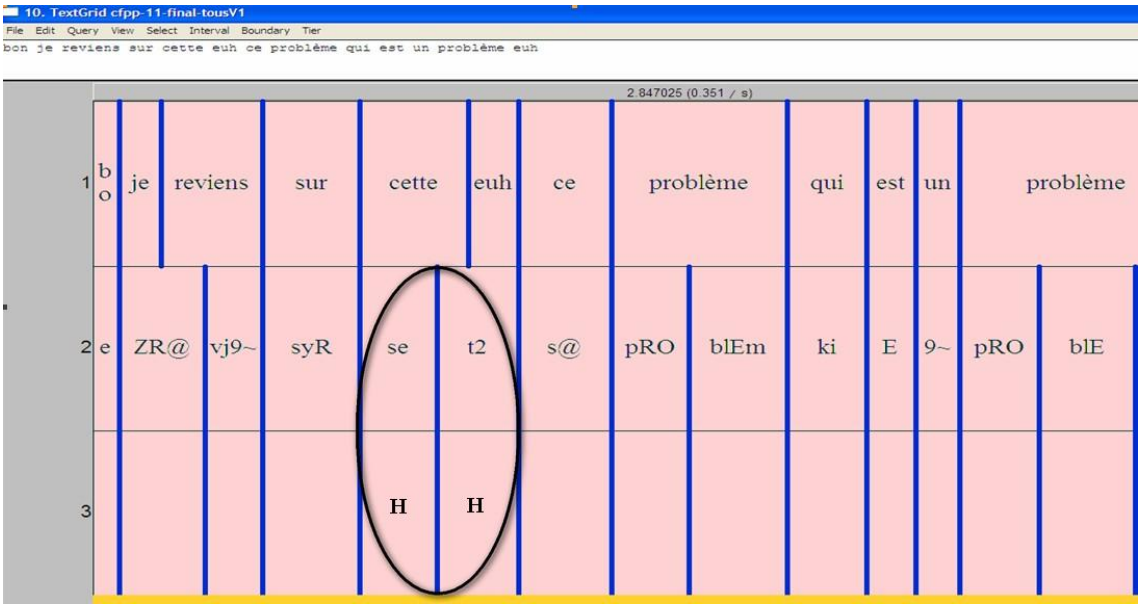


FIGURE 26 – « je reviens sur cette euh ce problème »

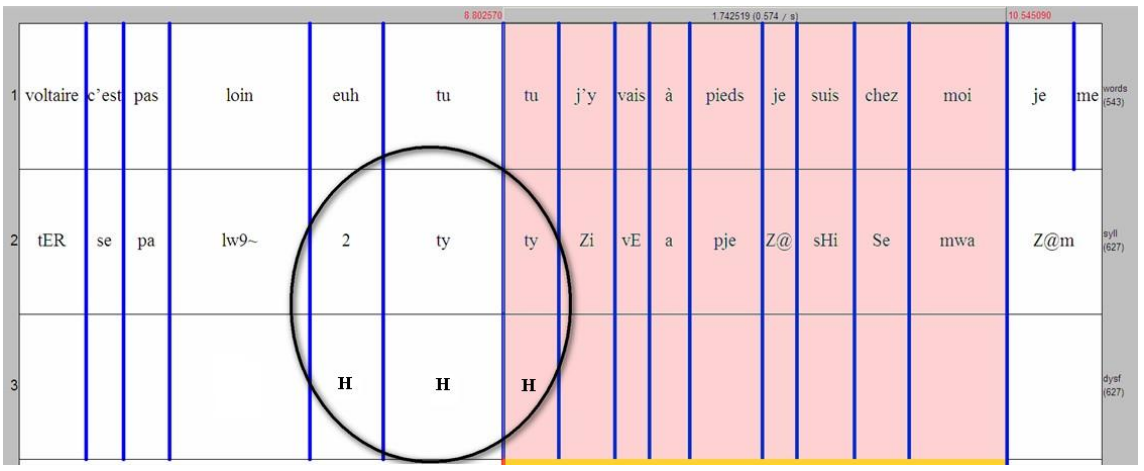


FIGURE 27 – « c'est pas loin euh tu tu j'y vais à pieds »

### 4.2.5 ETAPE 3 : EMPAN SUIVANT

Une fois le premier empan annoté avec les proéminences et les disfluences le cas échéant, on passe au suivant, et ainsi de suite jusqu'à la fin du fichier à coder.
Il est important de bien travailler *empan par empan* en respectant les intervalles de confiance délimités par les proéminences fortes, cependant, il faut aussi faire attention à ne pas couper n'importe où.

Il faut donc se rappeler qu'**on ne coupe pas au milieu d'un mot**. Ainsi, dans l'exemple ci-dessous, la deuxième fenêtre de codage interne démarre après « problème », la syllabe disfluente « euh » ne fait pas partie du second segment d'écoute. En effet, étant donné l'enchaînement phonologique à la frontière des mots, la syllabe /m9/ est associée au mot « problème » et non à l'unité qui suit :

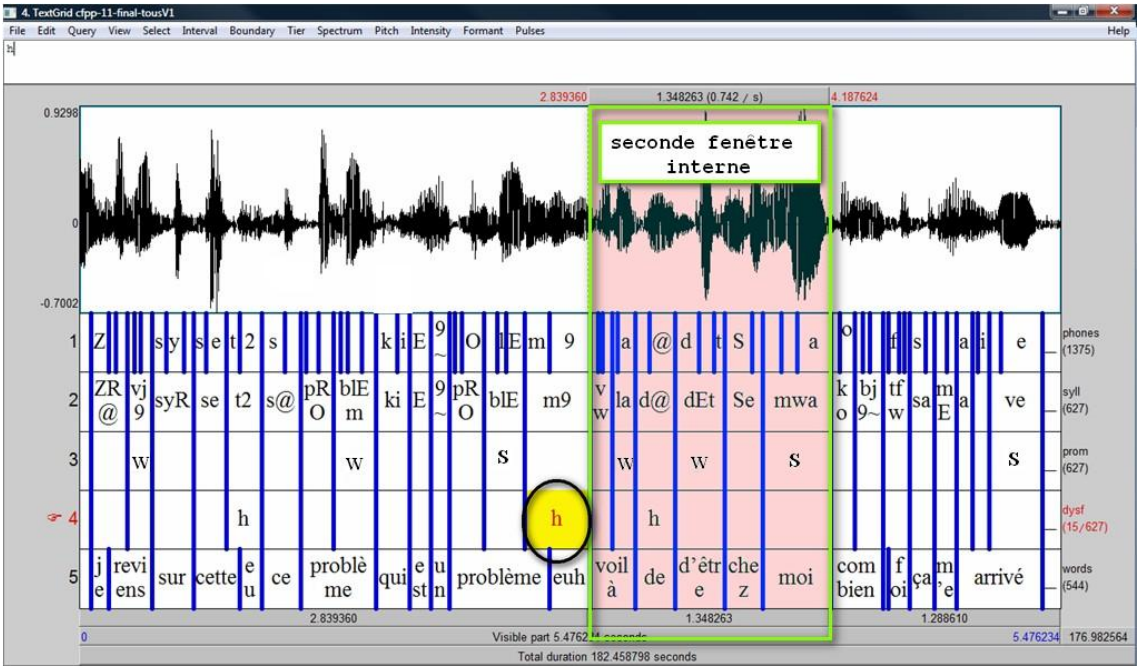


FIGURE 28 – Sélection d'une fenêtre interne après une disfluence

De même, une proéminence forte peut être au milieu d'un groupe de sens : il ne faut donc pas couper sur cette proéminence mais bien élargir la fenêtre interne jusqu'à la proéminence forte suivante :

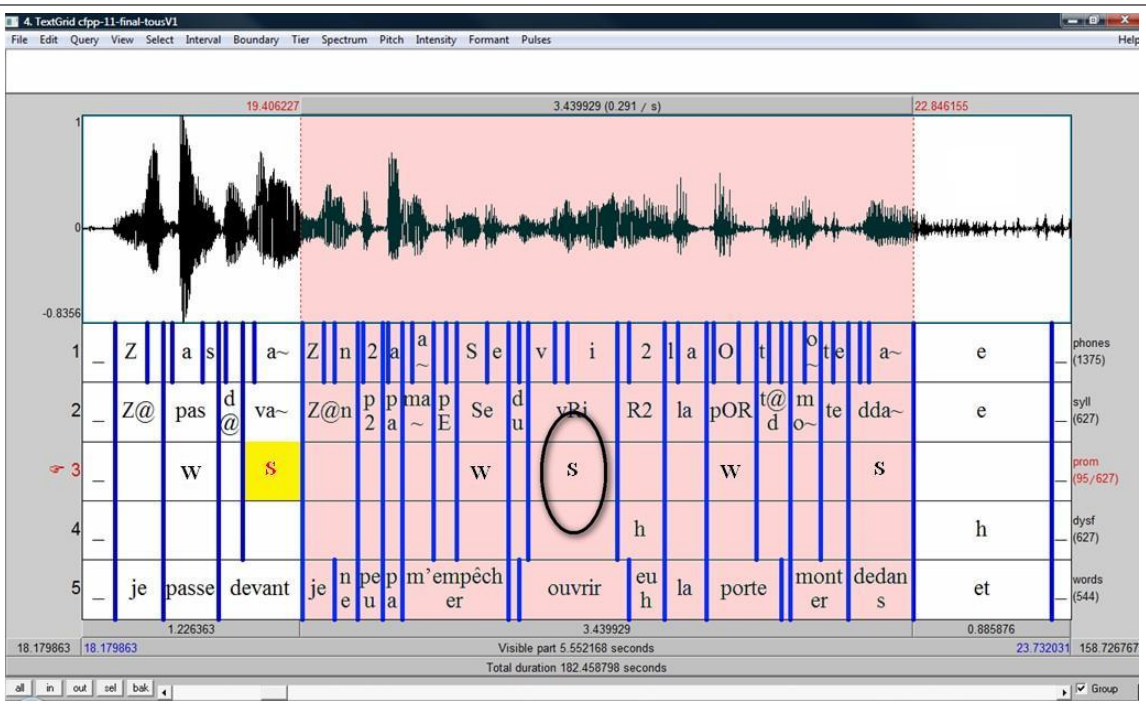


FIGURE 29 – Sélection d'une fenêtre interne avec une proéminence forte au sein d'un groupe de sens

Pour changer d'empan : zoom out (CTRL + O) puis déplacement vers la droite jusqu'à la pause suivante.

### 4.2.6 Etape 4 : Finalisation du codage

Une fois tout le fichier traité, le codeur, après avoir cliqué sur la tire à vérifier), doit activer dans le menu « Edit » la fonction « Find » (Ctrl-F) pour retrouver le marqueur « ? », réécouter les segments en question s'il y en a et prendre la décision qui lui paraît la plus pertinente au bout du compte, mais toujours en se référant au guide. Il ne doit rester aucun marqueur « ? » à la fin du travail !

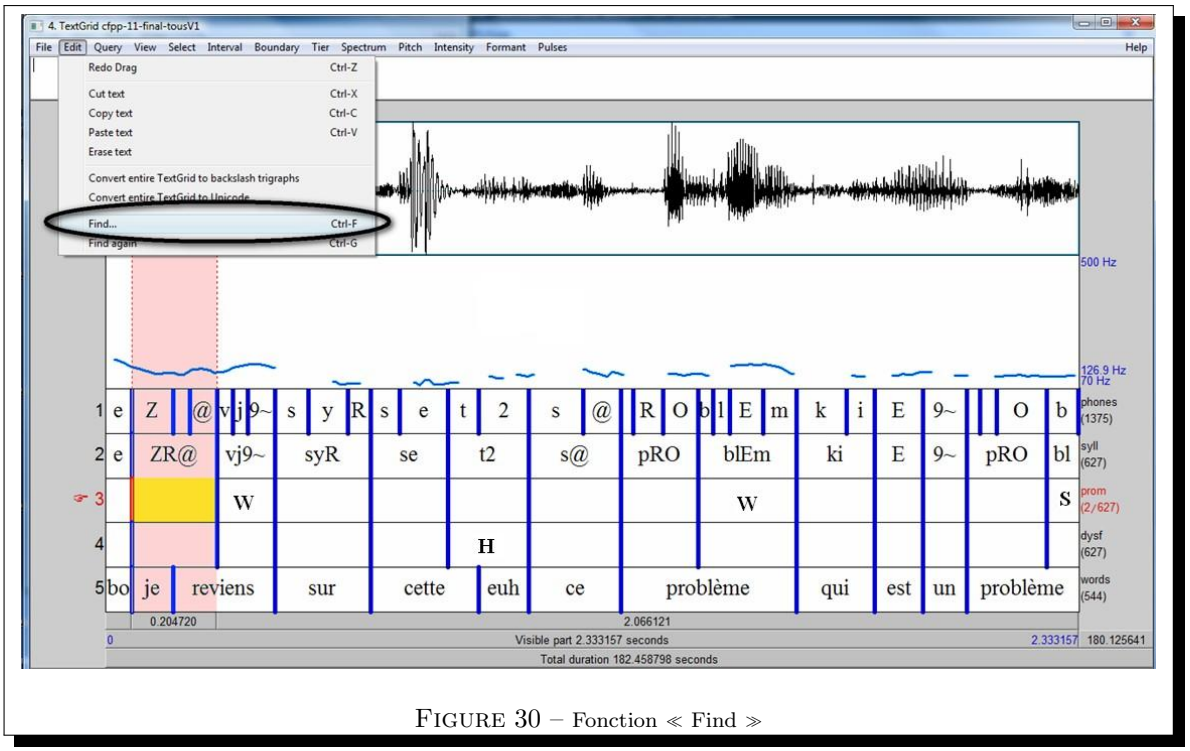


FIGURE 30 – Fonction « Find »

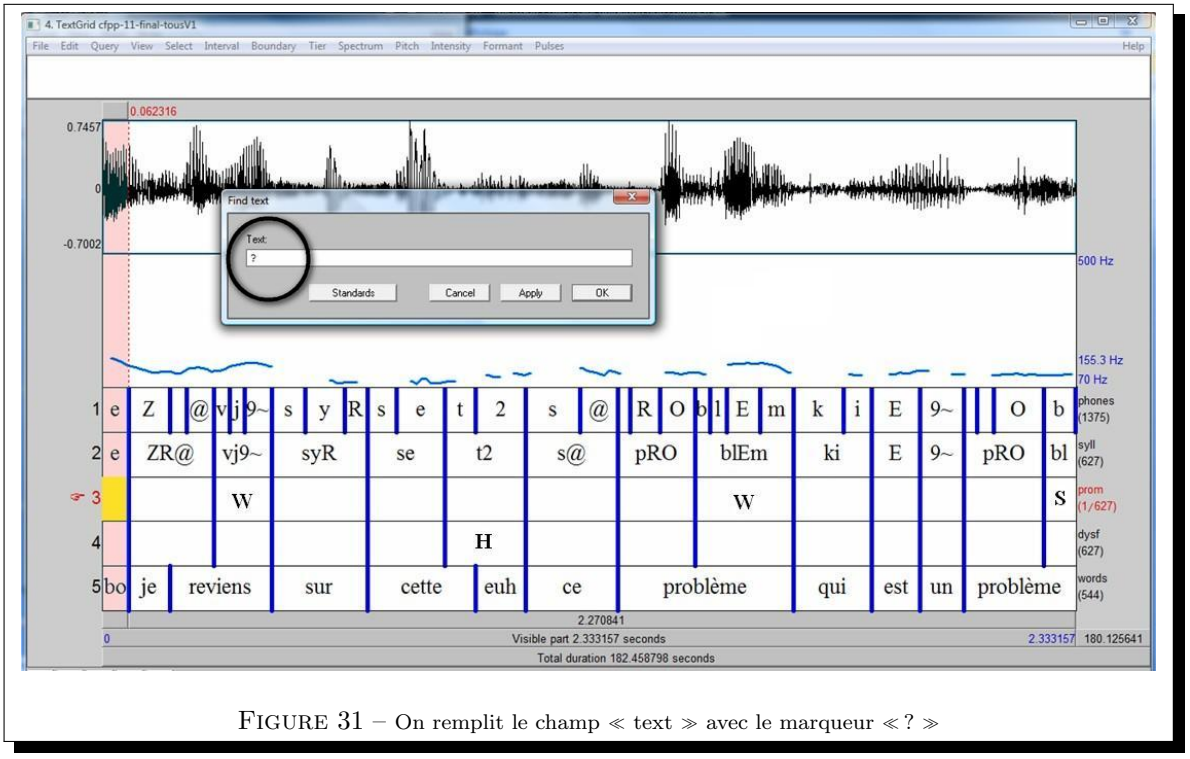


FIGURE 31 – On remplit le champ « text » avec le marqueur « ? »

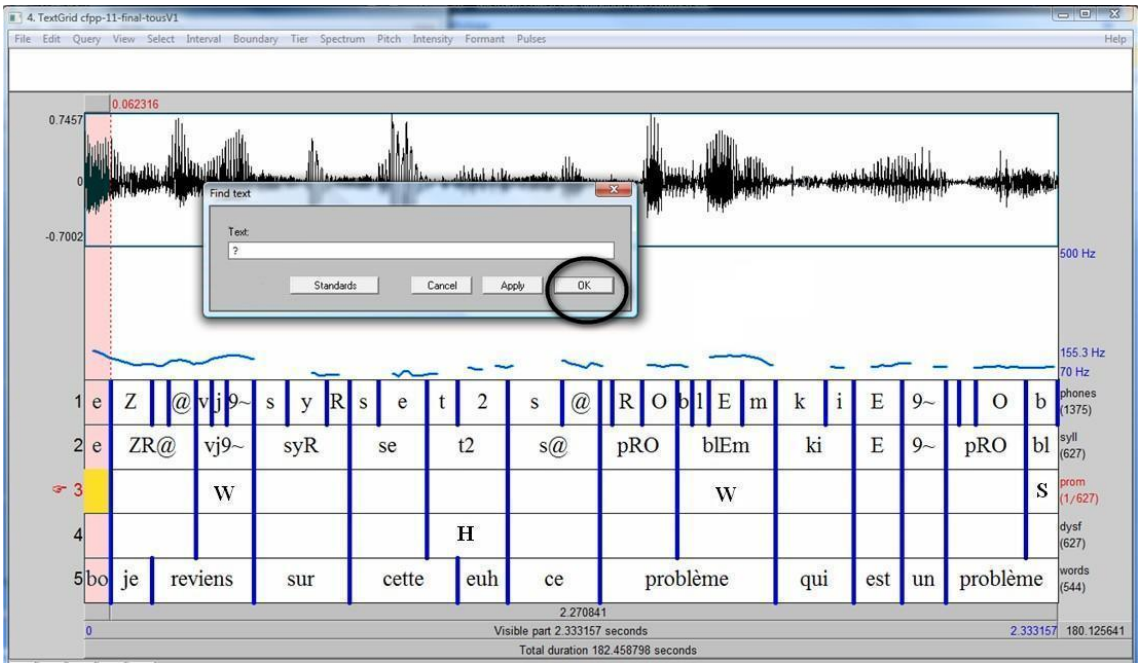


FIGURE 32 – On clique sur « OK »

Une fois sur la syllabe annotée « ? », on réécoute le segment jusqu'à 3 fois et on tranche.
Pour continuer, il suffit d'utiliser le raccourcis « Ctrl-G » (fonction find again), cela évite de refaire apparaître la fenêtre. On continue jusqu'à ce que le fichier ne contienne plus aucun « ? ».

N'oubliez pas de faire des sauvegardes intermédiaires (en tapant CRTL+S dans le Textgrid ouvert), et d'enregistrer votre travail sous un nom différent avant de quitter Praat.
Si vous fermez le Textgrid avant de sauver, sélectionnez le Textgrid dans la fenêtre Praat Objects et cliquez sur EDIT >WRITE >WRITE TO SHORT TEXTFILE :

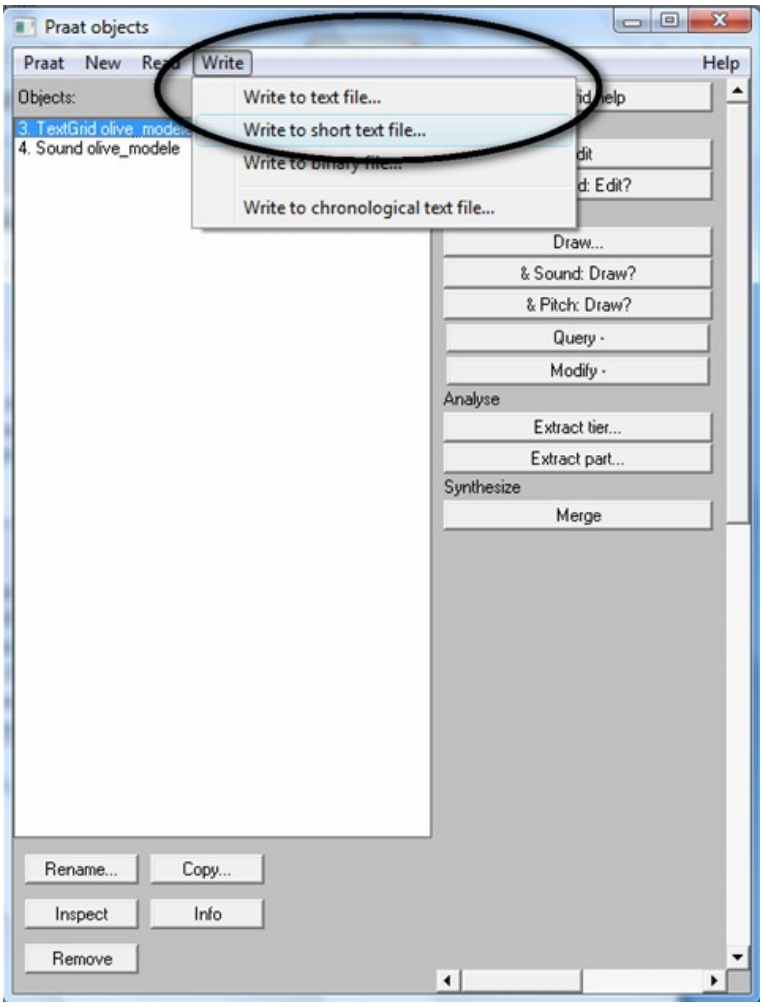


FIGURE 33 – Write to short textfile...

# 5 SURCODAGE VS PROÉMINENCES CONTIGUËS EFFECTIVES

Un problème qui peut survenir est le surcodage, c'est-à-dire l'annotation excessive de proéminences contiguës, faisant disparaître du même coup le concept même de proéminence syllabique comme une figure qui se détache sur un fond sonore. Dans cette section, nous allons voir comment éviter ce problème. Pour mémoire : on identifie les proéminences fortes (« S »), ensuite on se sert du marquage de ces proéminences fortes comme balises pour identifier de nouveaux intervalles de travail dans lesquels ont va coder des proéminences faibles « W ».

En fait, quand on identifie les proéminences fortes au début, il ne faut identifier que celles qui sont vraiment fortes par contraste relativement à celles qui précèdent (ou par rapport à celles qui suivent à l'initiale d'un nouveau groupe à traiter). Si on code trop de proéminences fortes lors de la première passe, on risque d'avoir trop de sous-fenêtres de codage pour l'étape suivante, ce qui va entraîner un surcodage. À ce propos, **il n'est pas obligatoire de trouver des « W » dans les segments balisés par des « S »** (autrement dit, il ne faut pas à trouver des objets perceptifs quand il n'y en a pas surtout quand les groupes sont courts.)

**Il ne faut vraiment annoter que les syllabes qui se détachent perceptivement de façon significative.** Si on procède de la sorte, il est *très rare* que l'on trouve plus de deux syllabes proéminentes qui se suivent. Il existe cependant quelques contextes classiques de proéminences contigües.

## 5.1 Contextes terminaux : syllabe pénultième + syllabe finale

La proéminence terminale peut être anticipée perceptivement sur l'avant-dernière syllabe que les syllabes appartiennent ou non au même mot :

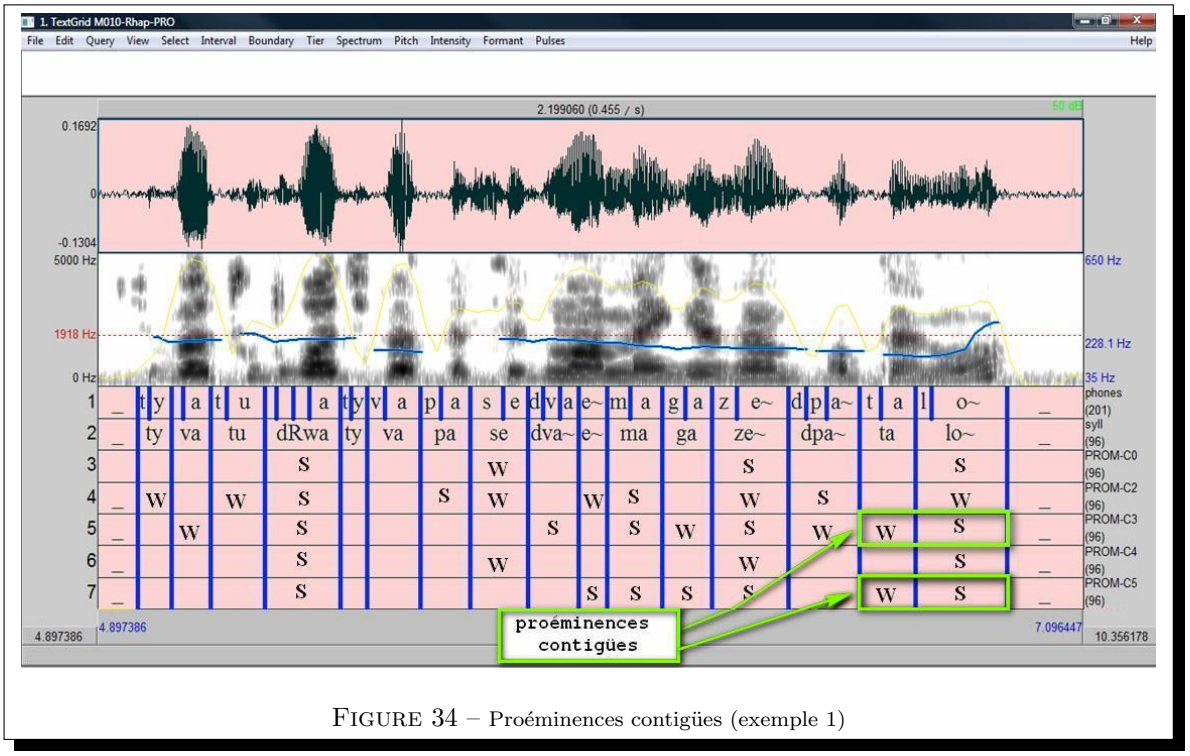


Figure 34 – Proéminences contigües (exemple 1)

Les exemples de cette section sont des annotations de cinq codeurs (C0, C2, C3, C4 et C5) : *C0 et C4 constitueront les annotations de référence.*

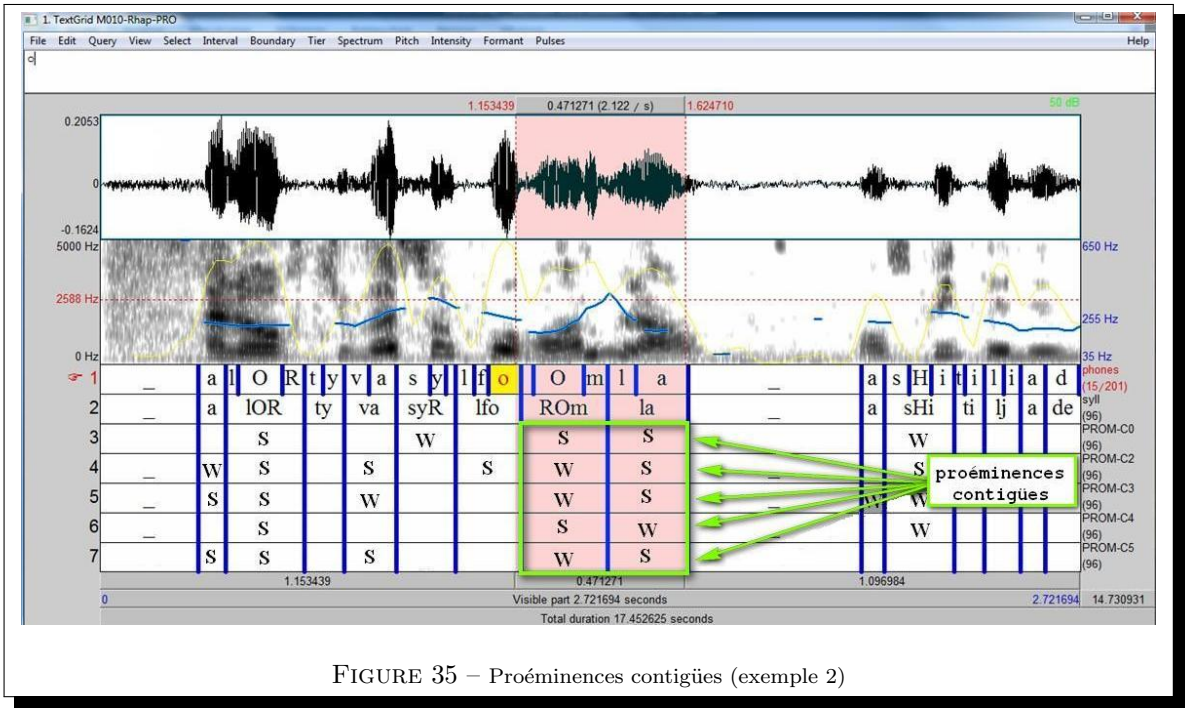


Figure 35 – Proéminences contigües (exemple 2)

## 5.2 Contiguité d'un accent final et initial

Il s'agit là de contextes où une syllabe initiale de groupe est perçue comme proéminente. Dans l'exemple ci-dessous une proéminence se place sur la fin du groupe « j'y vais à pieds » et à l'initiale du groupe suivant « moi ma voiture elle est garée dans la rue » :

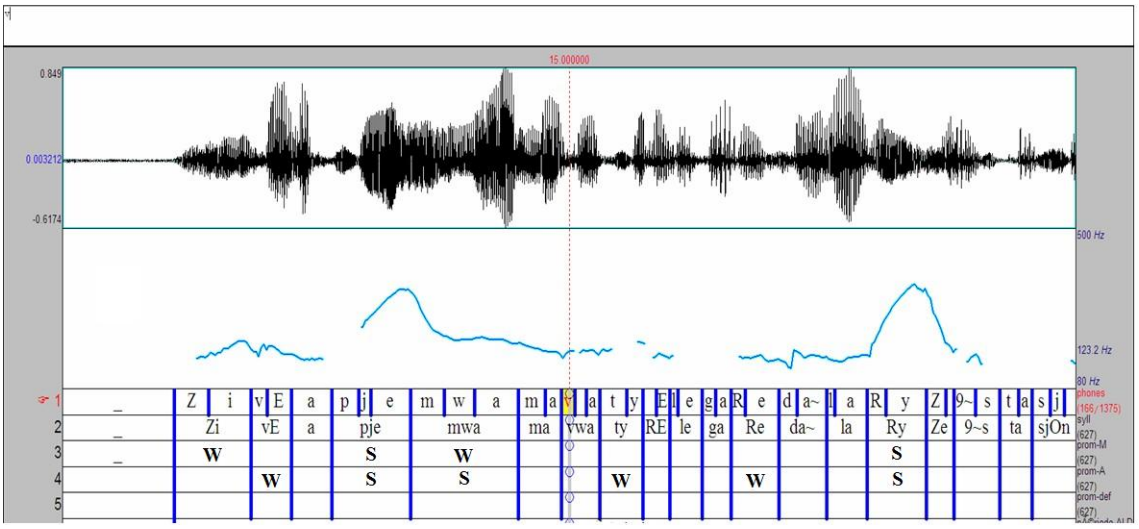


Figure 36 – Proéminences contigües (exemple 3)

## 5.3 Contextes d'insistance ou de scansion rythmique

Ci-dessous *neuf + fois*

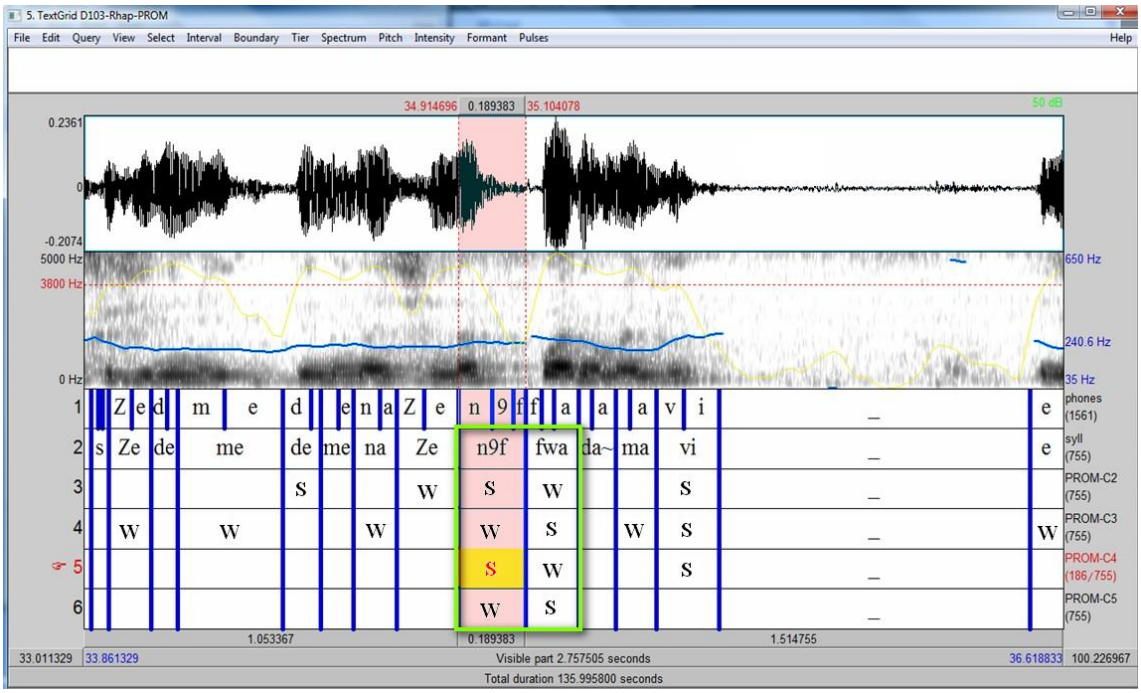


Figure 37 – Proéminences contigües (exemple 4)

## 5.4 La part de l'acoustique

Il est possible d'avoir une impression, sans pour autant avoir de certitude, que la proéminence s'étale sur deux syllabes contigües. Les causes peuvent être de nature acoustique or, vous ne pouvez pas savoir quelle est la part de l'acoustique dans cette histoire puisque vous ne devez pas avoir accès aux courbes de F0 et d'intensité (pour les codeurs novices).
C'est pour ces raisons que le nombre d'écoutes est limité à trois fois maximum. En effet, plus on écoute, plus on annote ! En d'autres termes, on annote lors de la première écoute, puis on réécoute une fois voire deux afin de lever les quelques doutes.

# 6 Remarques et précisions pour le codage

## 6.1 Les proéminences

### 6.1.1 Que faire lorsqu'on hésite entre « W » et « S » ?

Ce genre d'hésitation ne peut survenir que dans l'étape d'annotation des proéminences fortes qui permettent de déterminer les intervalles de confiance (étape 1). Ainsi, si le codeur hésite entre proéminence forte et faible pour une syllabe, c'est vraisemblablement que cette syllabe ne se détache pas assez fortement de son environnement, il convient donc de laisser cette syllabe vide dans un premier temps puis de l'annoter faible « W » lors de l'étape suivante (étape 2).

### 6.1.2 Que faire lorsqu'on hésite à annoter une syllabe « W » ou non ?

La procédure est la même que pour les disfluences :
On réécoute le segment jusqu'à 3 fois maximum, si l'hésitation persiste, on annote « ? ». Une fois le fichier entièrement codé, on effectuera une recherche dans la tire correspondante (ctrl+f) des marqueurs « ? » et on tranchera à ce moment-là.

### 6.1.3 Comment fixer le degré d'une proéminence par rapport à son entourage ?

Le degré de la proéminence (W/S) doit être fixé **relativement à ce qui *suit* et non relativement à ce qui *précède*.**
L'ilôt de confiance est donc le contexte droit (i.e. groupe prosodique en cours), *le gauche n'est utilisé que quand il n'y a pas de contexte droit.*
Soit l'exemple ci-dessous : pour étiqueter correctement « moi », on compare « moi » à « rue » et non à « pieds »

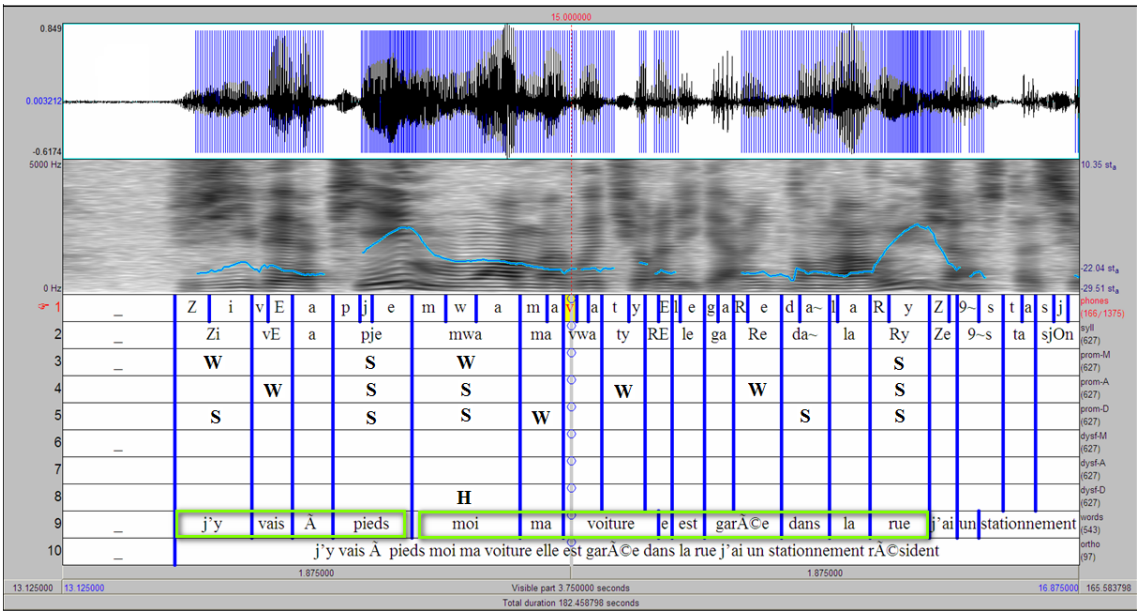


Figure 38 – (j'y vais à pieds) (moi ma voiture elle est garée dans la rue)

Autrement dit, quand on a affaire à une proéminence forte, il faut aller voir le premier contexte droit avec une proéminence forte (qui n'est pas forcément contiguë syllabiquement), ce qui suppose de faire abstraction de la perception des hauteurs intermédiaires.

### 6.1.4 Faut-il tenir compte des allongements syllabiques perçus ou des variations mélodiques pour déterminer la proéminence ?

Lorsqu'on compare deux syllabes proéminentes pour déterminer leur degré de proéminence (W/S), on peut se poser la question : la première est-elle aussi saillante que la seconde par la hauteur et/ou par la durée ?
Si on prend l'exemple ci-dessous, « moi » est codé « S » par la durée, mais si on ne tient compte, de façon réductrice, que de la F0, l'étiquetage sera contestable (on aura « W » tout au plus).

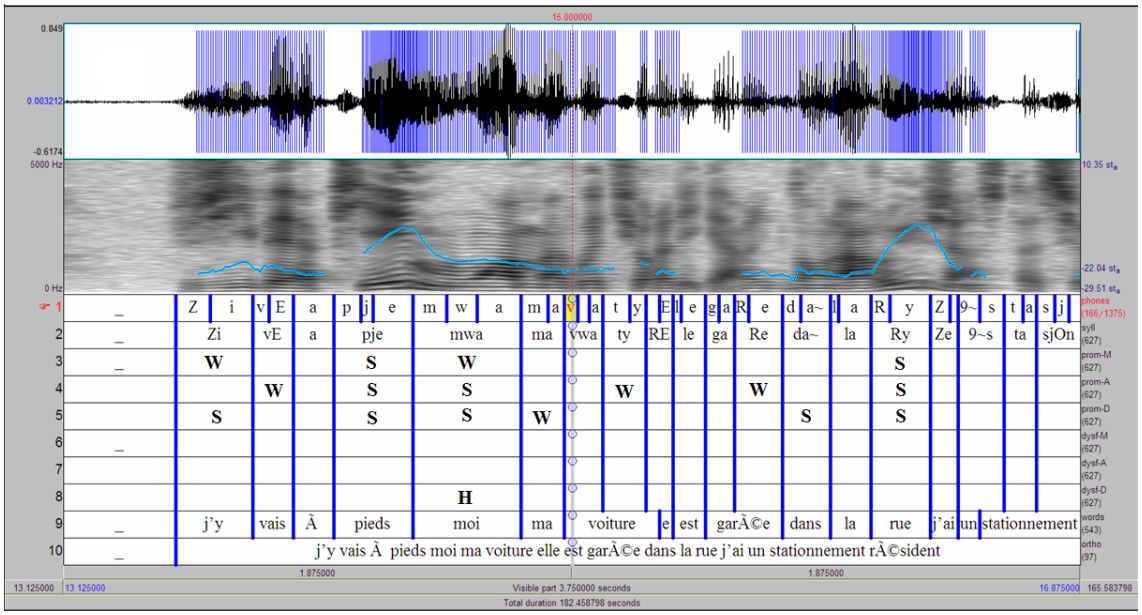


FIGURE 39 – « moi » est-il « W » ou « S » ?

En conclusion, la perception des allongements temporels est aussi fondamentale que la perception des variations mélodiques pour l'annotation des proéminences, d'autant que, pour certains contextes et/ou certains locuteurs, ce sont les variations temporelles qui sont significatives là où les variations mélodiques restent ténues.

### 6.1.5 Comment traiter le « e » extramétrique ou post-tonique ?

Pour les syllabes finales suivies d'un « E » extramétrique ou post-tonique, associé au « e » terminal de la syllabe (schwa) et/ou à un « euh » d'hésitation, deux stratégies sont possibles pour le marquage des proéminences mais il faut se tenir à l'une ou l'autre de ces stratégies tout le long de la phase d'annotation :

1. annoter la syllabe pleine si elle est perçue proéminente 'W' ou 'S', dans ce cas, ne pas annoter la syllabe subséquente contenant le « E » extramétrique proéminente
2. annoter la syllabe subséquente contenant le « E » extramétrique comme proéminente, dans ce cas, ne pas annoter la syllabe pleine précédente comme proéminente

Ainsi, si une proéminence forte est perçue sur la syllabe de « problème » dans « *qui est un problème euh* » alors l'annotation est soit :

| pro | blE | m9 |
|---|---|---|
| | S | |

Soit :

| pro | blE | m9 |
|---|---|---|
| | | S |

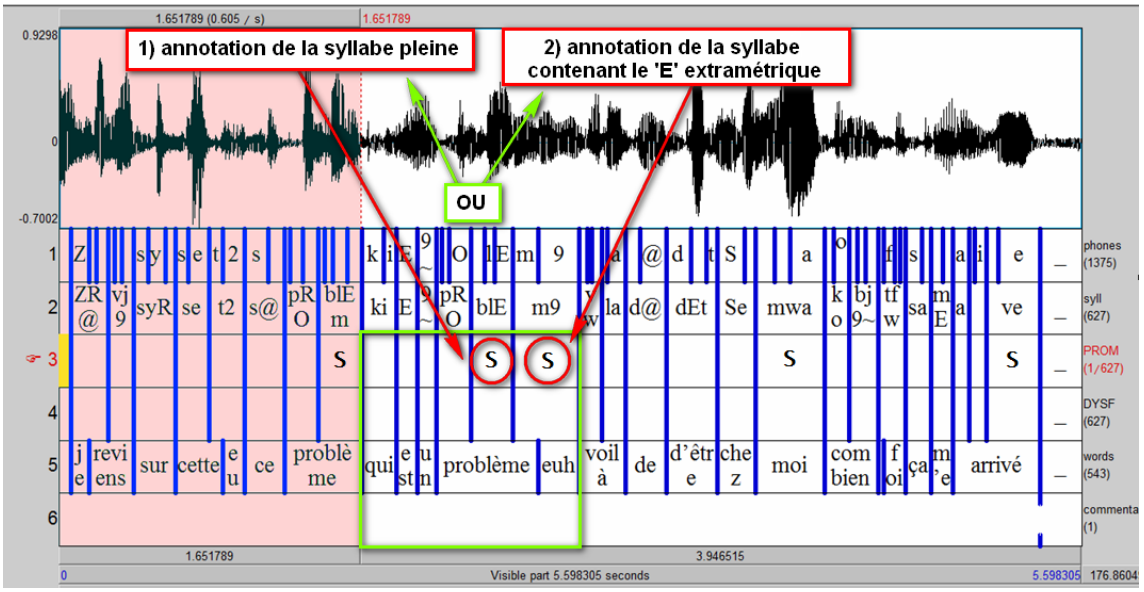


FIGURE 40 – « E » extramétrique dans « qui est un problème euh »

De même pour « limites » dans « il faut quand même me mettre des limites », l'annotation est soit :

| | | |
|---|---|---|
| li | mi | t@ |
| | S | |

Soit :

| | | |
|---|---|---|
| li | mi | t@ |
| | | S |

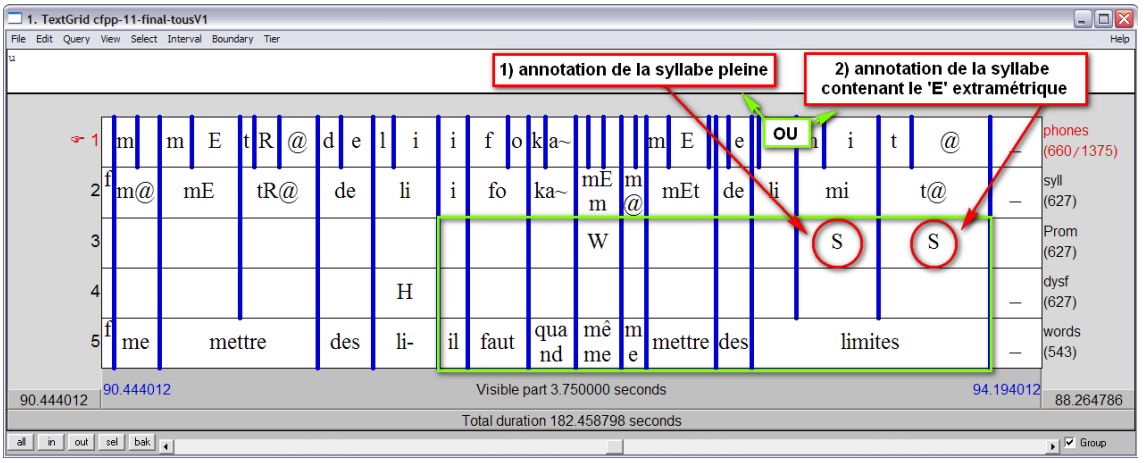


FIGURE 41 – « E » extramétrique dans « il faut quand même me mettre des limites »

A noter : le schwa post-tonique peut être disfluent ou annoté « 2 », dans ces cas voir la section 6.2.3.

### 6.1.6 COMMENT ANNOTER L'ATTAQUE CONSONANTIQUE EN INITIALE DE MOT ?

Si la parole du locuteur est resentie globalement comme une parole hyper-articulée lors de l'écoute initiale (cf. *supra*, 4.2), alors il ne s'agit pas de proéminence ; par contre s'il s'agit d'un phénomène d'hyper-articulation locale alors il est associé à une proéminence perceptive.

## 6.2 Les disfluences

### 6.2.1 Que faire lorsqu'on hésite à annoter une syllabe « H » ou non ?

On réécoute le segment jusqu'à 3 fois maximum, si l'hésitation persiste, on annote « ? ». Une fois le fichier entièrement codé, on effectuera une recherche dans la tire correspondante (ctrl+f) des marqueurs « ? » et on tranchera à ce moment-là.

### 6.2.2 Une séquence inachevée est-elle disfluente ?

Les séquences inachevées ne sont pas marquées disfluentessi elles sont suivies d'un autre locuteur et si, dans ce contexte d'interruption de parole, elles ne sont pas associées à un indice prosodique précis (saut de registre dans le grave, allongement).

### 6.2.3 Le schwa post-tonique ou extra-métrique peut-il être annoté disfluent ?

Dans certains cas, c'est facile, le schwa final est codé « 2 » et accompagné d'un « euh » sur la tire orthographique, dans ces cas, on annote la syllabe « H ».
Pour d'autres contextes, c'est plus difficile car la tire orthographique n'indique rien et la tire phonétique « @ ». Dans ces contextes, si vous avez l'impression que la prononciation du « e » final est liée à une recherche de dénomination lexicale ou de désignation et pas à un marqueur phonostylistique du locuteur, c'est-à-dire que l'enchaînement consonantique est possible alors il convient de noter disfluent « H ».

### 6.2.4 Les marqueurs de discours sont-ils tous disfluents ?

- Les éléments du type « bon », « ben », « quoi » sont marqués comme des hésitations **si et seulement si** ils sont acoompagnés d'un allongement marqué. Ainsi, dans la séquence « *eh ben euh* », seule la syllabe « euh » sera étiquetée « H ».
- Si aucun marqueur prosodique n'est associé aux constructions du type « il me semble », « je crois » , « enfin », « j'veux dire » alors ces constructions ne sont pas annotées « H ».
- Les marqueurs phatiques liés à l'interlocution (« hein », « vous voyez », etc.) *ne sont pas annotés disfluents.*

### 6.2.5 Les reformulations sont-elles des disfluences ?

S'il n'y a aucun marqueur prosodique associé à la reformulation, il ne faut pas annoter disfluent puisqu'il s'agit alors juste d'une spécification (ajout d'information). ***Ne pas se fier à la lecture de la transcription (au visuel) mais à l'écoute !*** Des exemples :

- *(La réponse) (question-réponse) (vous allez l'avoir)* : le segment 1 *(La réponse)* n'est pas une hésitation
- *(le point de vue du spectateur) (le point de vue du spectateur)* : le premier segment est un rhème descendant final de période, le second est une reprise thématique montante en début de période
- *(votre expérience par exemple si vous pratiquez la musique) (votre expérience en tant que pratiquant la musique)* : il n'y a aucune hésitation ici mais bien une reformulation

### 6.2.6 Les fausses hésitations sont-elles des disfluences ?

Les disfluences prosodiques perçues (allongements syllabiques importants et ou pauses non silencieuses par exemple) devront être étiquetées comme telles dans tous les genres discursifs, y compris les genres comme le genre MEDIA qui « miment » l'improvisation de la parole ordinaire. Autrement dit, à ce niveau de l'annotation, on ne fait pas de différences entre disfluences volontaires et non volontaires, entre fausses hésitations motivées par des stratégies rhétoriques et vraies hésitations associées par exemple à des problèmes de planification lexicale et/ou de réparation syntaxique.
S'il vous semble que c'est une façon de mettre en valeur ce qui va suivre, vous pouvez ajouter « S » sur le mot outil qui précède le lexème mis en valeur par cette stratégie ; vous pouvez joindre un commentaire du type « HS ? » ou « HW ? » dans la tire réservée aux commentaires.

## 6.3 Les proéminences et/ou les dyfluences

### 6.3.1 Une syllabe peut-elle être proéminente ET disfluente ?

Oui, une syllabe peut être proéminente *et* disfluente.

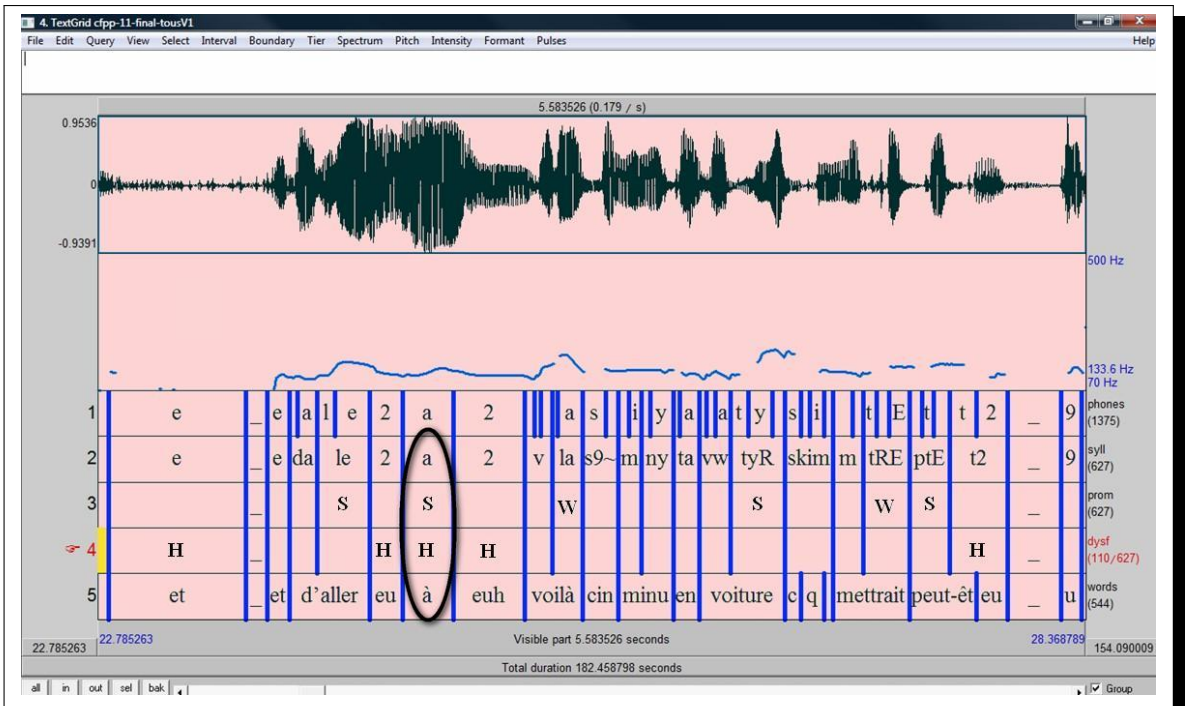


FIGURE 42 – Syllabe proéminente et disfluente (exemple 1)

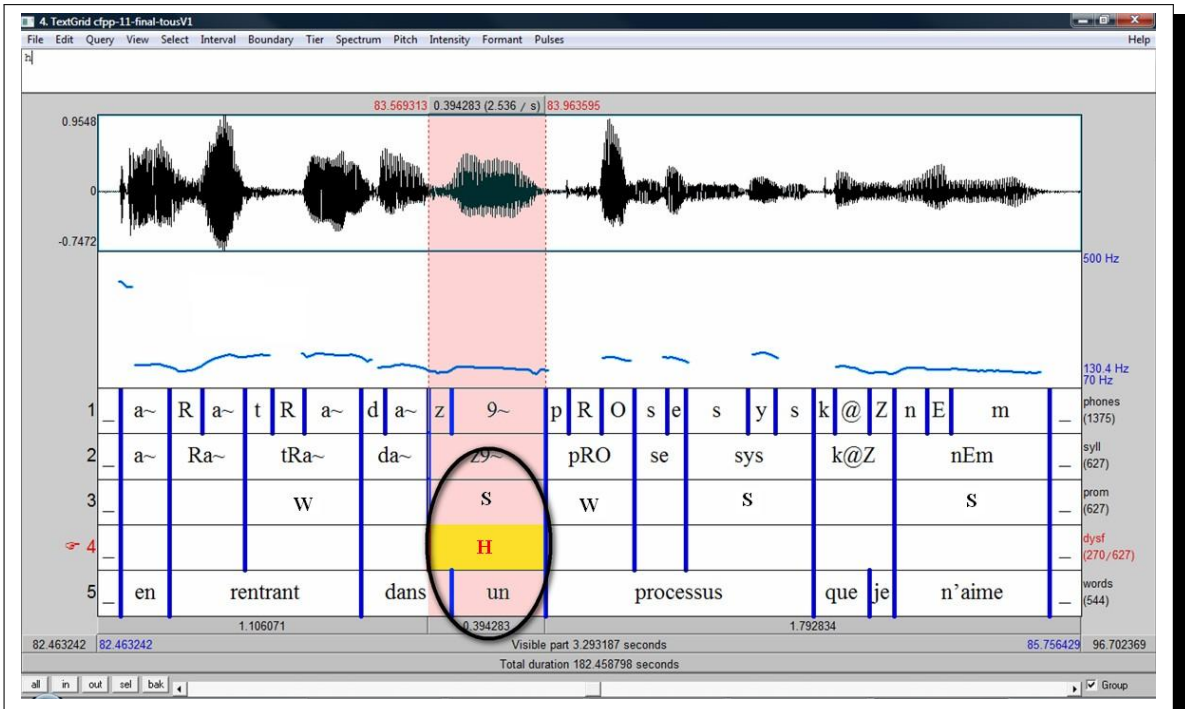


FIGURE 43 – Syllabe proéminente et disfluente (exemple 2)

### 6.3.2 QUE FAIRE LORSQU'ON HÉSITE ENTRE PROÉMINENCE ET DISFLUENCE POUR UNE SYLLABE ALLONGÉE ?

il n'existe pas de solution toute faite, le mieux est donc de rester le plus intuitif possible.

## 6.4 Divers

### 6.4.1 Comment coder les fichiers multi-locuteurs ?

Pour les fichiers multilocuteurs, si aucune balise n'est associée aux différents tours de parole (tout dépend la façon dont les fichiers de transcription ont été préparés) alors vous vous fiez au son.
Pour la perception des proéminences et des disfluences pour un locuteur donné, on prendra comme référence les stratégies vocales (tessiture, rythme, etc.) du locuteur en question uniquement (en aucun cas on ne prend comme référence les stratégies de l'interlocuteur ou des interlocuteurs). En d'autres termes, on traite un fichier locuteur par locuteur, on ne code pas en écoutant deux locuteurs en même temps.

# 7 Suivi d'annotations (pour les experts uniquement)

Il ne faut pas perdre de vue, dans cette tâche de suivi d'annotation, qu'il faut faire autant que possible avec les indices prosodiques auxquels nous avons accès perceptivement mais aussi en **visualisant le signal**, ce qui n'est pas le cas pour les codeurs novices.
Le suivi de codage doit donc se faire en visualisant les informations acoustiques mais ***ATTENTION*** **sur les bases d'une courbe de F0 corrigée (sous Praat ou sous Analor)**.

## 7.1 Accord inter-annotateur

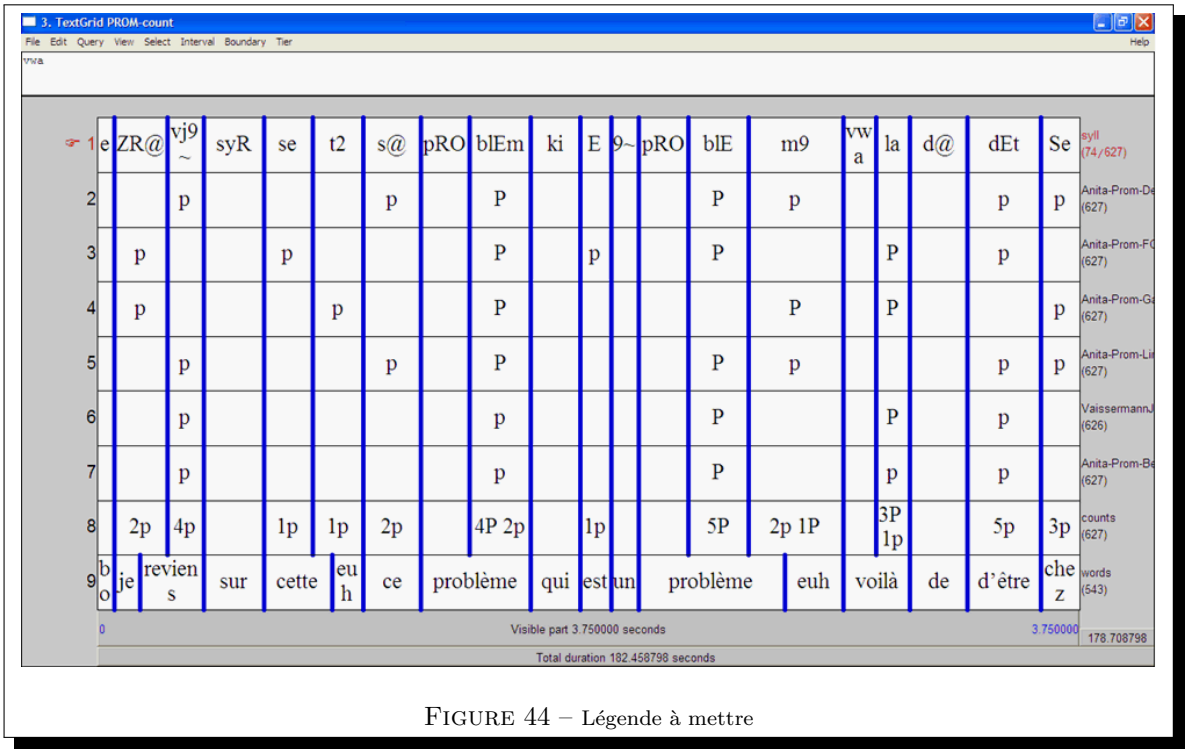


Figure 44 – Légende à mettre

## 7.2 Conseils pour la création de l'annotation de référence

Il est possible de dériver une annotation de référence en se basant sur des annotations réalisées par des novices. Dans ce cas, il convient d'apporter quelques modifications en se basant sur les indices prosodiques perceptibles mais aussi sur le signal auquel n'avaient pas accès les codeurs novices.

Il faut donc opérer des choix, faire des modifications, etc. mais en suivant quelques règles :

- **Penser aux requêtes que l'on va opérer sur le corpus :** dans quelle mesure le codage proposé va pouvoir rapporter de l'information pour les requêtes posées ? L'important est de donner les moyens au linguiste qui voudra faire une interprétation fonctionnelle des données de pouvoir la faire ; il faut donc faire des choix en conséquence quitte à ce que ces choix ne soient pas toujours bien traditionnels.

- **Respecter un principe d'économie dans le suivi d'annotation :** il faut faire la part des choses entre les erreurs graves et celles qui ne le sont pas dans l'étape de suivi : les erreurs d'annotations entraînent-elles des effets qu'il n'y aurait pas eus ? Ces erreurs se répercutent-elles sur les résultats obtenus via les requêtes ? Si non, il n'est pas nécessaire de tout recoder !
- **Utiliser la tire « commentaire »** notamment pour les recodages des zones consensuelles si vous en faites.

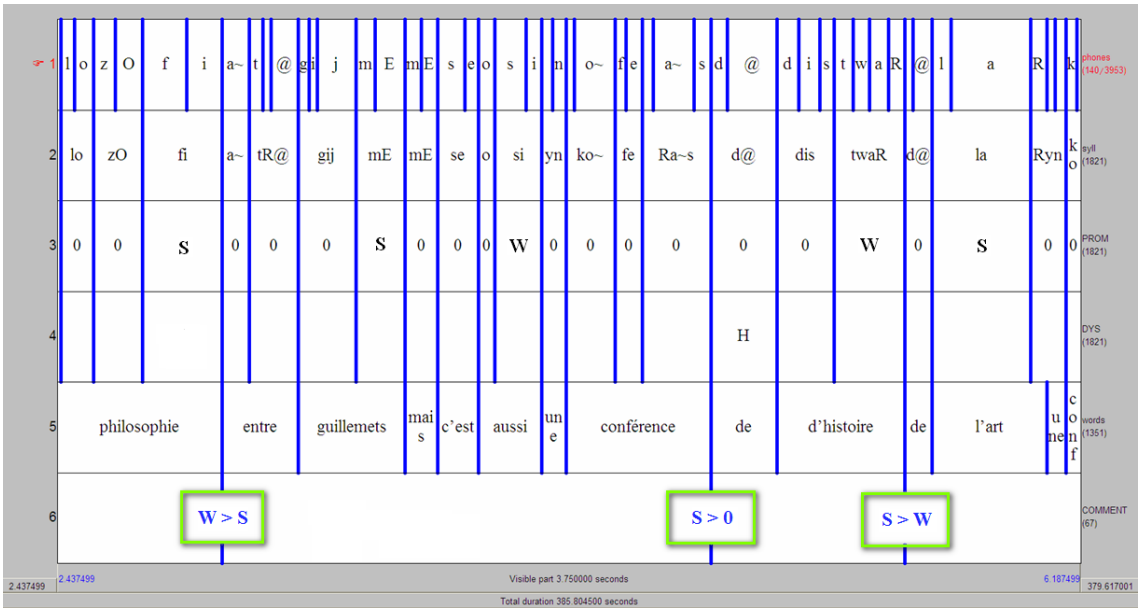


FIGURE 45 – Commentaires de recodage

De gauche à droite :
– la dernière syllabe du mot « philosophie », notée consensuellement « W », a été recodée « S »
– la dernière syllabe de « conférence », notée consensuellement « W », a été recodée « 0 »
– la dernière syllabe du mot « histoire » a été recodée « W

## 7.3 REMARQUES POUVANT AIDER AU SUIVI D'ANNOTATION

- Une théorie (Martinet 1960) démontre que plus une unité lexicale est rare statistiquement, plus elle a de chances d'être informative donc saillante, proéminente. Ainsi, des mots rares, des noms propres ou encore des mots étrangers auront plus de chance d'être saillants. Ce qui veut dire concrètement qu'en cas de non consensus chez les codeurs naïfs en ce qui concerne une proéminence sur ce type de mots, il vous faudra rétenir la proéminence.
- autre remarque ?

# 8 LA SEGMENTATION EN PÉRIODES INTONATIVES

## 8.1 À VENIR

## 8.2 À VENIR

# Glossary

## A

**attaque consonantique** Dans l'attaque consonantique, la durée et l'intensité de la consonne sont plus marquées que celles de la voyelle.

## D

**disfluence** « achoppement dans la linéarité de l'énoncé, marque de discours en cours d'élaboration » (Dister 2007).

**dysfluence** Elément qui brise le déroulement syntagmatique dans la chaîne parlée, sorte de trébuchement vocal.

## E

**encodage** Manière de représenter les différents caractères dans un système informatique.

## P

**prosodie** Enveloppe temporelle et mélodique de la voix.

**proéminence** Syllabe qui se détache de son environnement comme une figure sur un fond.

## T

**tire** Niveau d'annotation.

# Table des figures


1 Sélection de Text reading preferences . . . 4
2 Sélection de Text writing preferences . . . 4
3 Sélection de Read from file... . . . 5
4 Edition des fichiers . . . 5
5 On décoche ≪ show spectrogram ≫ . . . 6
6 Sélection . . . 6
7 On zoome avec ≪ sel ≫ . . . 7
8 Les différentes tires . . . 7
9 Correspondance API - SAMPA pour les consonnes du français . . . 8
10 Correspondance API - SAMPA pour les voyelles du français . . . 8
11 D'autres paramètres que la hauteur rentrent en compte (exemple 1) . . . 12
12 D'autres paramètres que la hauteur rentrent en compte (exemple 2) . . . 12
13 Exemple de commentaires . . . 14
14 Exemple de fenêtre de codage . . . 15
15 Coupure au niveau de la disfluence en cas d'empan trop long . . . 16
16 Différence entre pause démarcative et pause non démarcative . . . 16
17 Intervalle de confiance de plus de 5 secondes . . . 17
18 Changement de locuteur . . . 17
19 Sélection de la zone . . . 18
20 Proéminence faible en fin de groupe de souffle . . . 19
21 Sélection du premier intervalle du segment délimité à sa droite par une proéminence forte . . . 19
22 Codage des proéminences faibles ≪ W ≫ . . . 20
23 Codage des disfluences ≪ H ≫ . . . 21
24 ≪ voilà de d'être chez moi ≫ . . . 22
25 ≪ en rentrant dans un processus que je n'aime pas trop vis-à-vis de moi ≫ . . . 22
26 ≪ je reviens sur cette euh ce problème ≫ . . . 23
27 ≪ c'est pas loin euh tu tu j'y vais à pieds ≫ . . . 23
28 Sélection d'une fenêtre interne après une disfluence . . . 24
29 Sélection d'une fenêtre interne avec une proéminence forte au sein d'un groupe de sens . . . 24
30 Fonction ≪ Find ≫ . . . 25
31 On remplit le champ ≪ text ≫ avec le marqueur ≪ ? ≫ . . . 25
32 On clique sur ≪ OK ≫ . . . 26
33 Write to short textfile... . . . 27
34 Proéminences contigües (exemple 1) . . . 28
35 Proéminences contigües (exemple 2) . . . 28
36 Proéminences contigües (exemple 3) . . . 29
37 Proéminences contigües (exemple 4) . . . 29
38 (j'y vais à pieds) (moi ma voiture elle est garée dans la rue) . . . 30
39 ≪ moi ≫ est-il ≪ W ≫ ou ≪ S ≫ ? . . . 31
40 ≪ E ≫ extramétrique dans ≪ qui est un problème euh ≫ . . . 32
41 ≪ E ≫ extramétrique dans ≪ il faut quand même me mettre des limites ≫ . . . 32
42 Syllabe proéminente et disfluente (exemple 1) . . . 34
43 Syllabe proéminente et disfluente (exemple 2) . . . 34
44 Légende à mettre . . . 35
45 Commentaires de recodage . . . 36