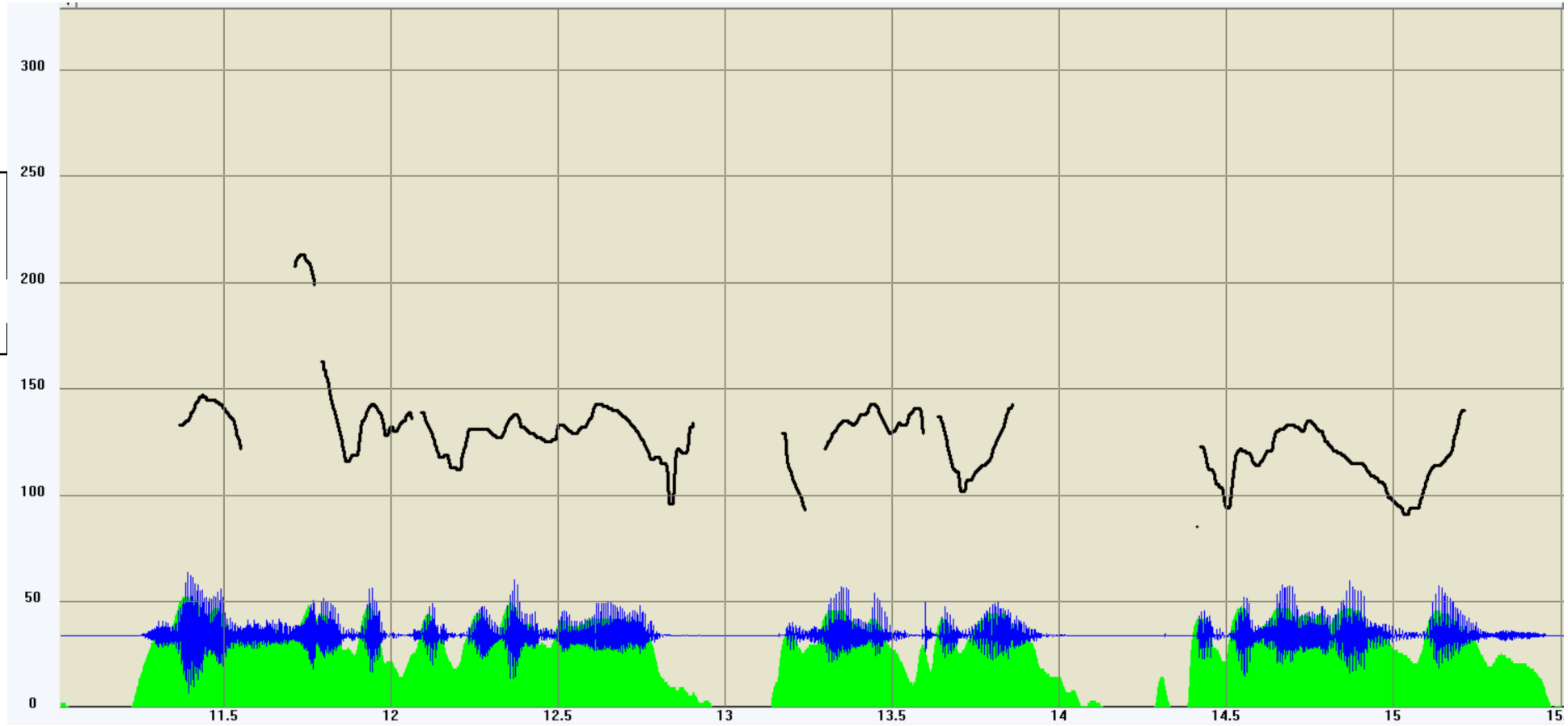
Pitch Instruments Inc.

Mesure de la fréquence fondamentale avec WinPitch

Philippe Martin

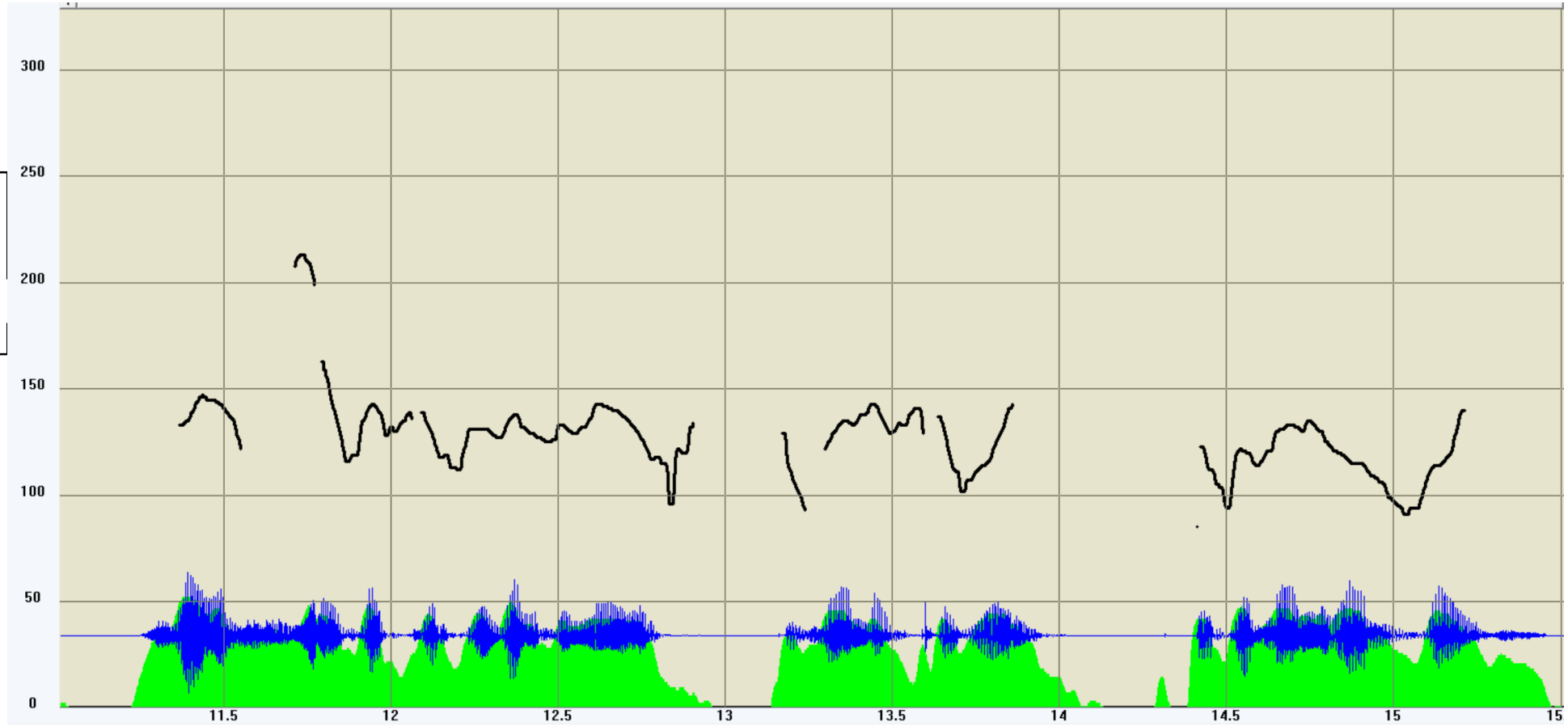
Object: use of the various imbedded function for F0 tracking and annotation

Fundamental frequency (aka melodic) curve on the analysis window



Object: use of the various imbedded function for F0 tracking and annotation

The liability of the F0 curve depends on a lot of factors.



Les ennemis de Fo

Micros à bande passante insuffisante
(plus de fondamentale)

Écho
(parois de la salle d'enregistrement trop réflexives, fenêtres…)

Niveau d'enregistrement trop faible
(micro trop éloigné du locuteur)

Sources sonores multiples
(bruit de moteur, chevauchement de parole…)

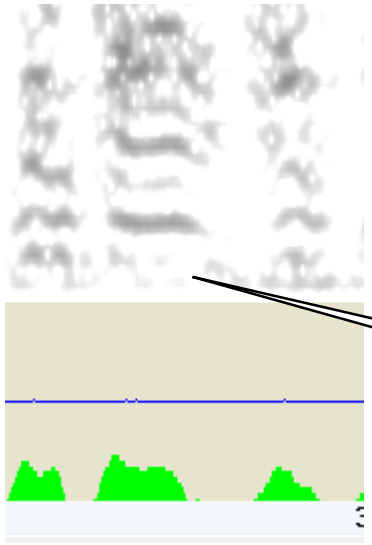
Contrôle automatique de volume (AVC)
(dénature les transitions voisés non voisés)

Codage mp3, wma… à faible débit
(déformation parfois inaudible du signal)

Format d'enregistrement > 22050 Hz
(temps de calcul excessif)

Les ennemis de Fo

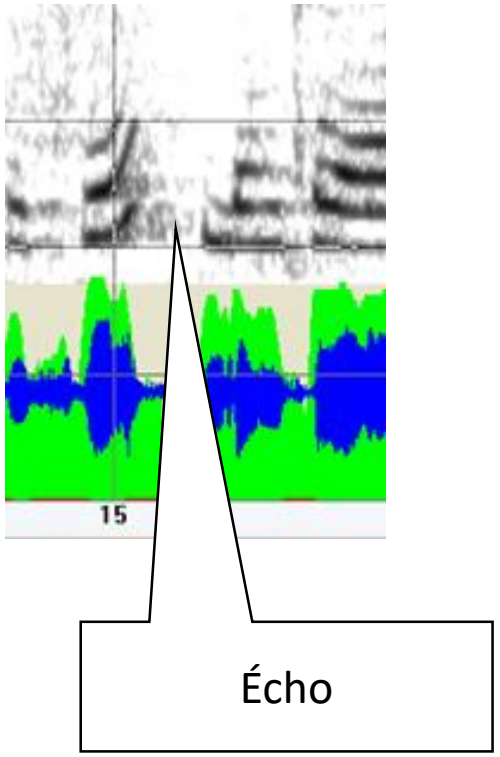
Micros à bande passante insuffisante

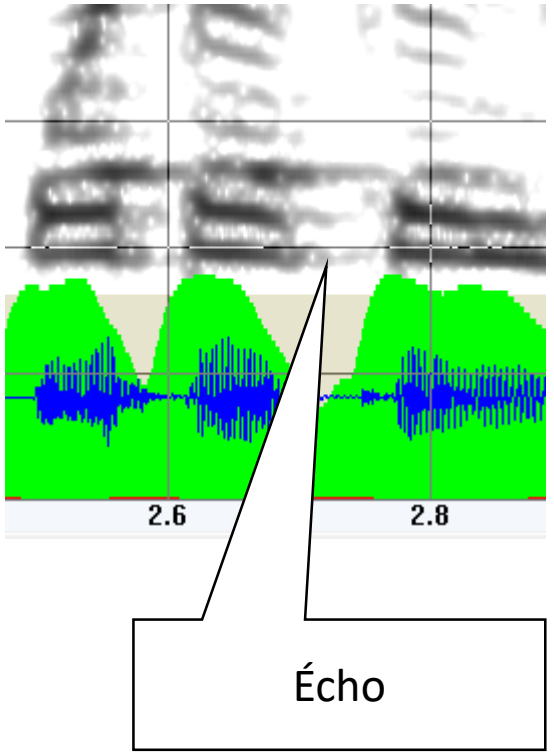


Fondamentale très faible ou absente

Les ennemis de Fo

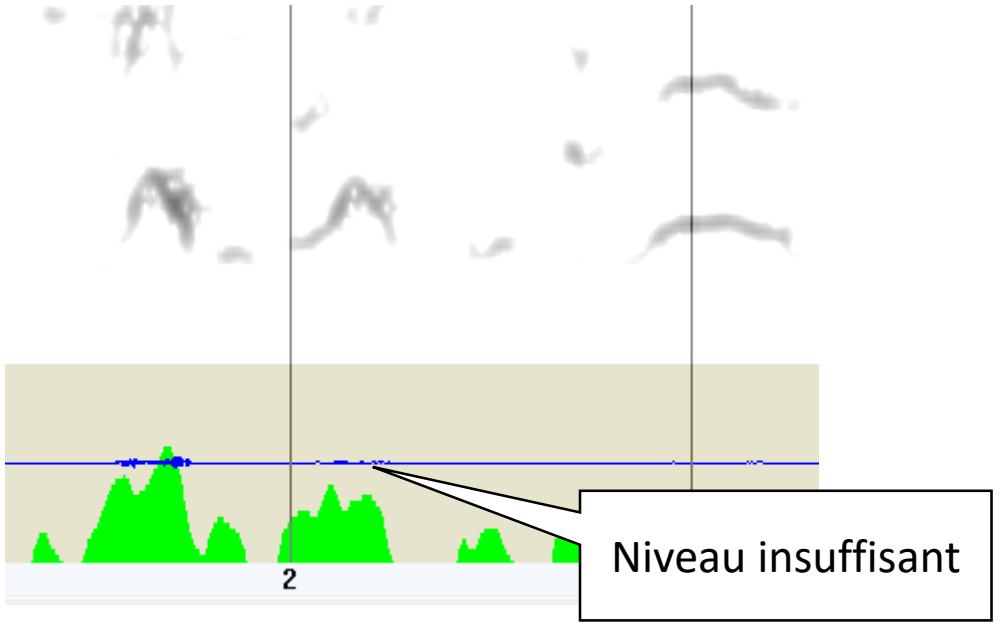
Écho





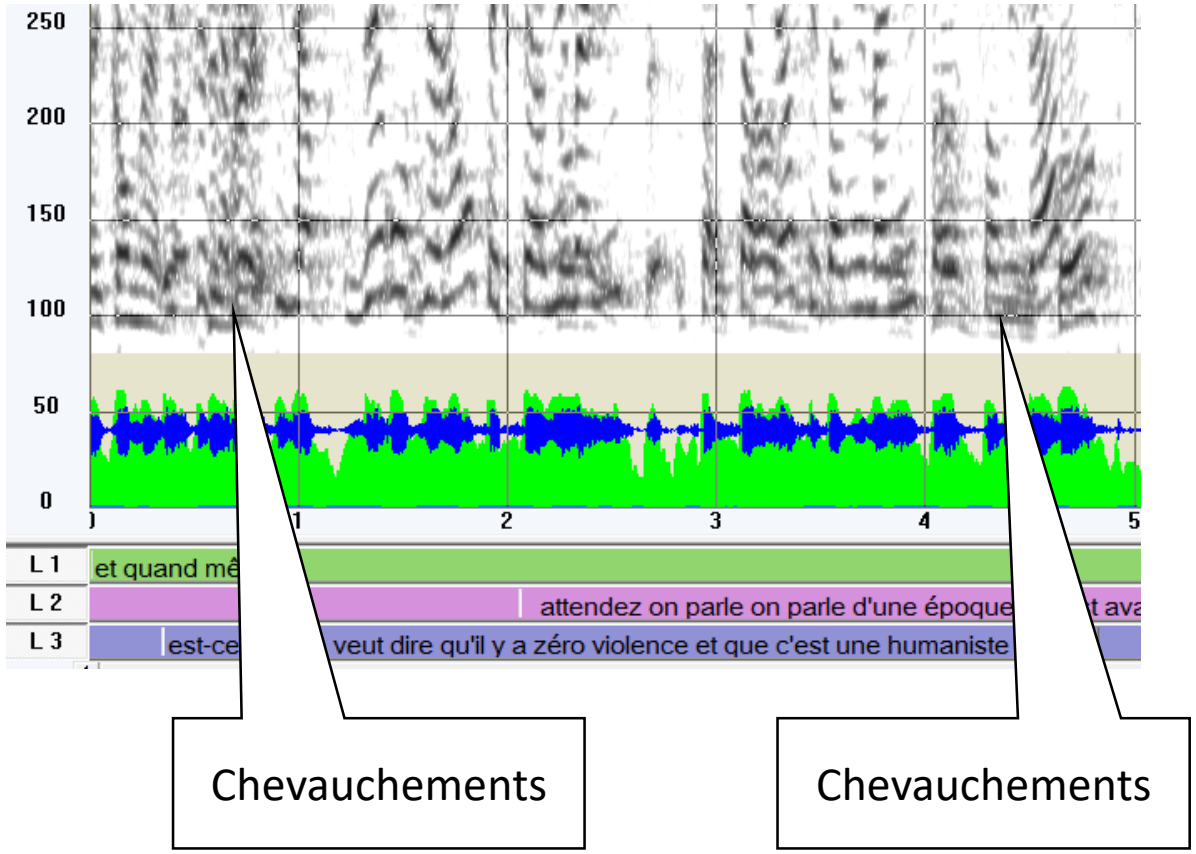
Les ennemis de Fo

Niveau d’enregistrement trop faible



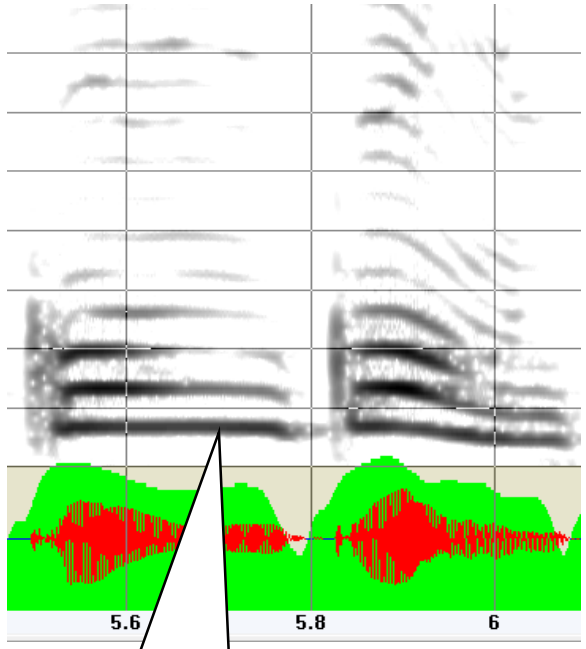
Les ennemis de Fo

Sources sonores multiples

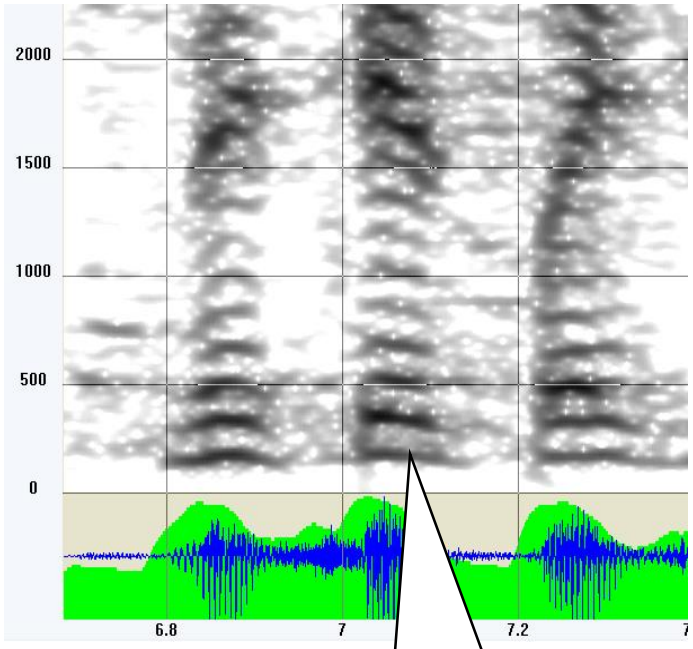


Les ennemis de Fo

Codage mp3, wma… à faible débit



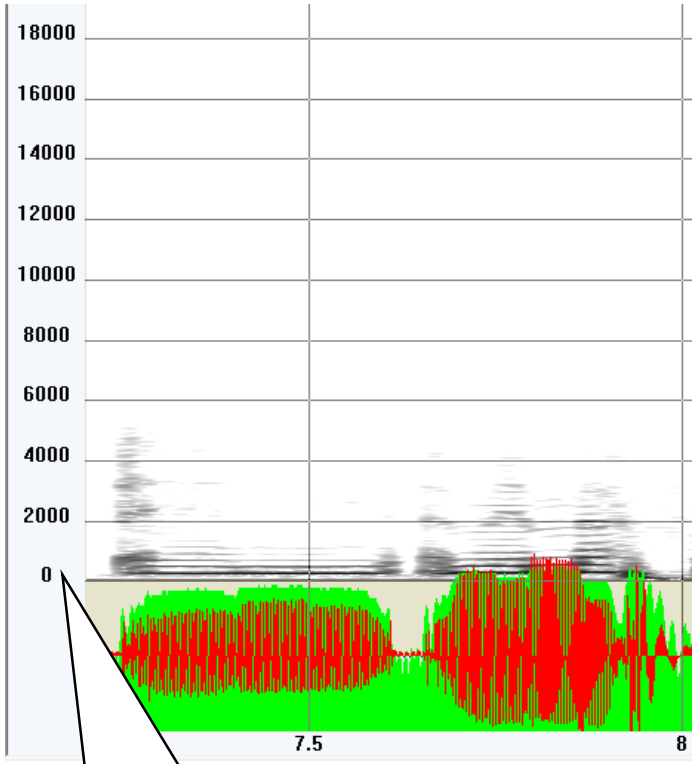
Format wav



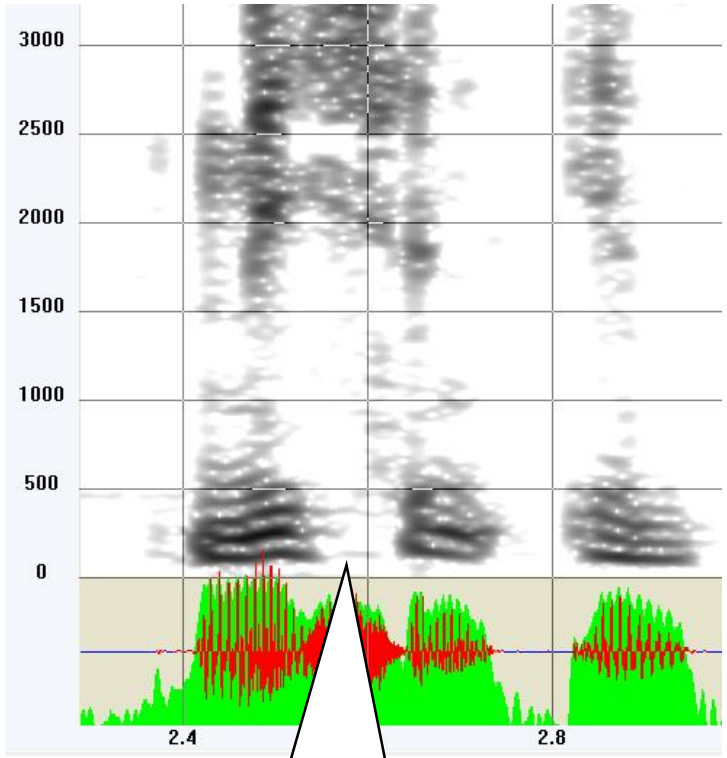
Codage mp3

Les ennemis de Fo

Format d'enregistrement > 22050 Hz



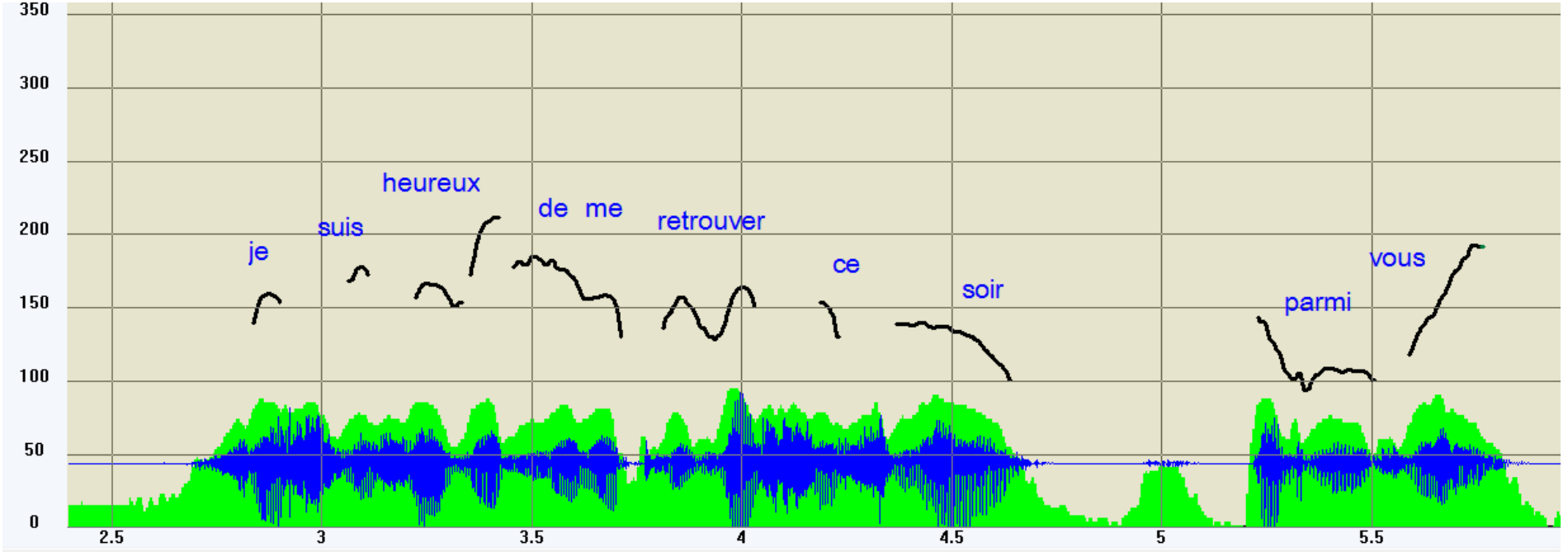
Format 44100 Hz



Format 44100 Hz
et codage mp3

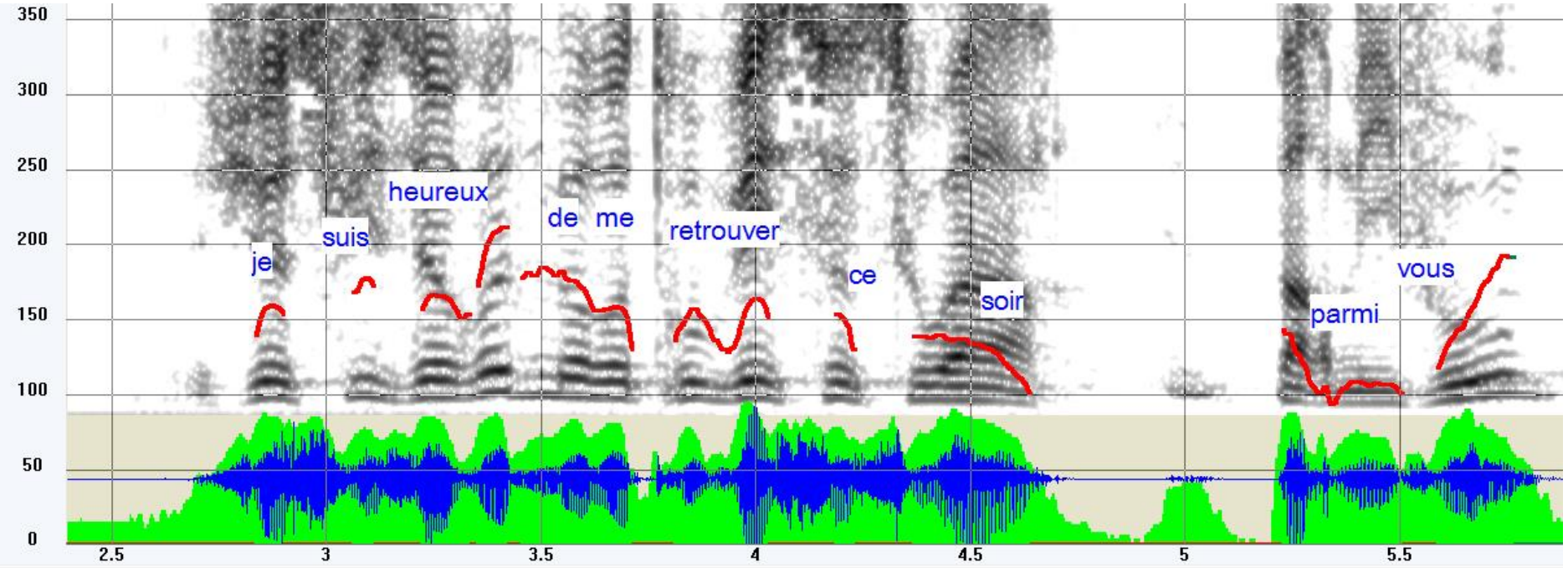
Nettoyage des courbes mélodiques

...quand la musique est bonne...



Peut on se fier à cette courbe mélodique ?

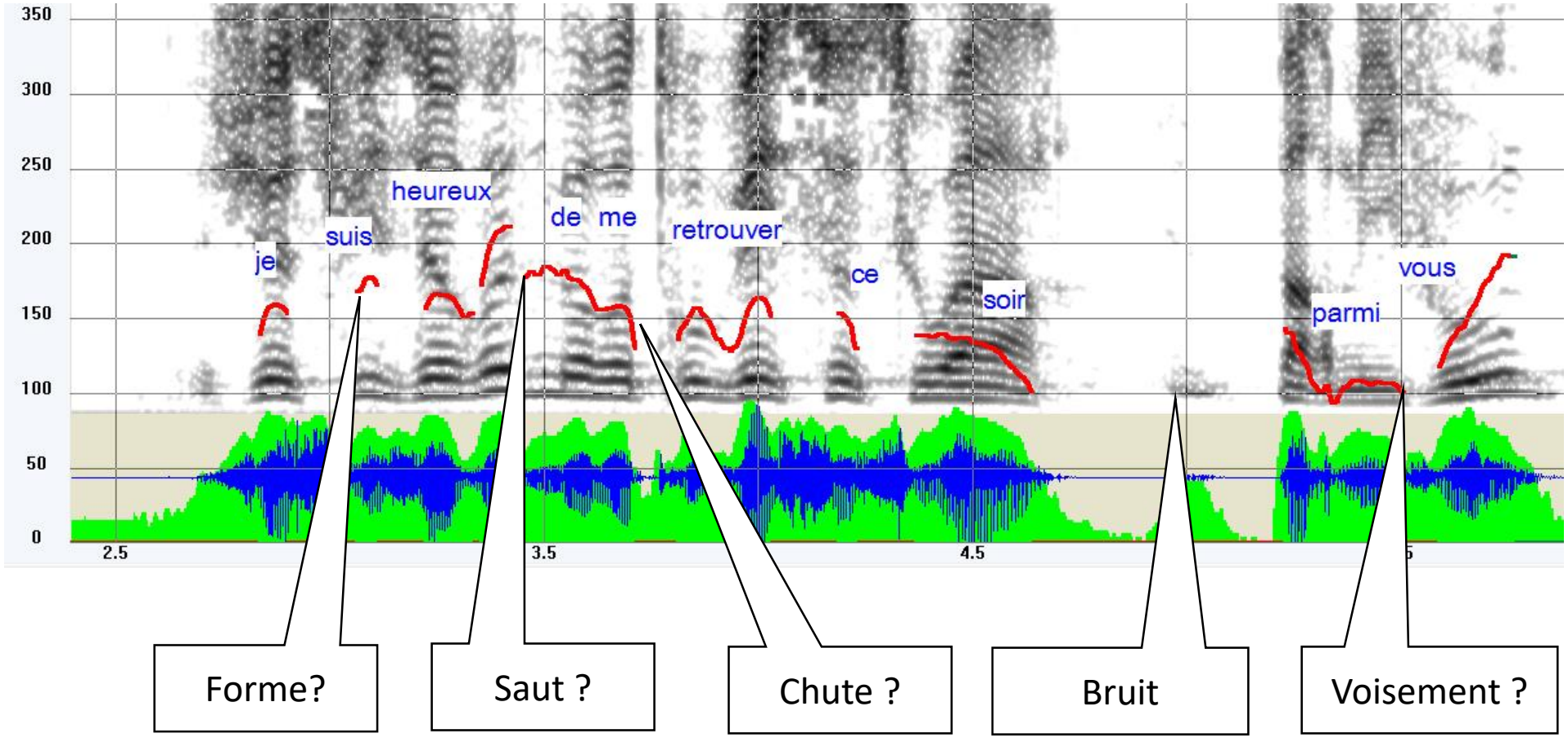
Vérification visuelle au spectrogramme



Spectrogramme « bande étroite »-> harmoniques

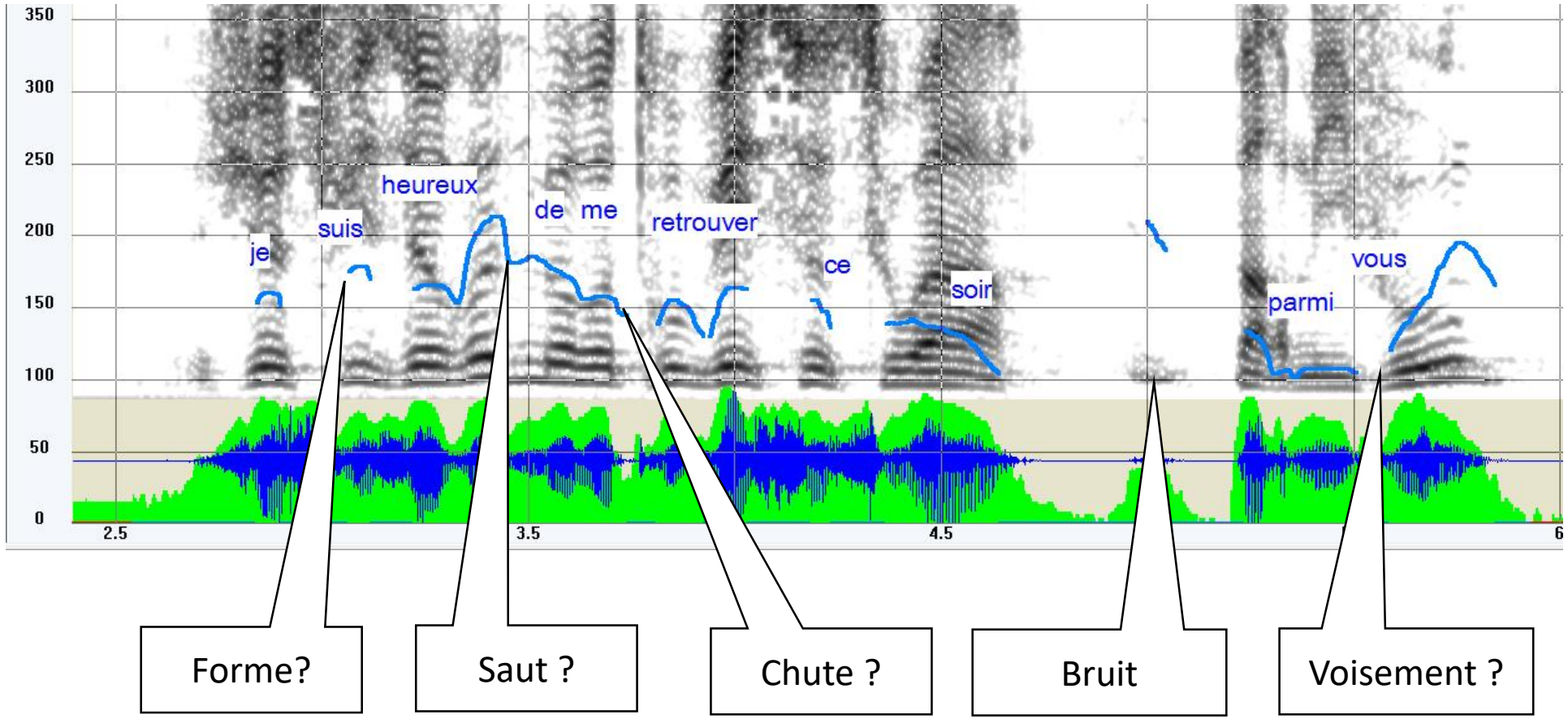
Comparaison de divers algorithmes

Le peigne spectral



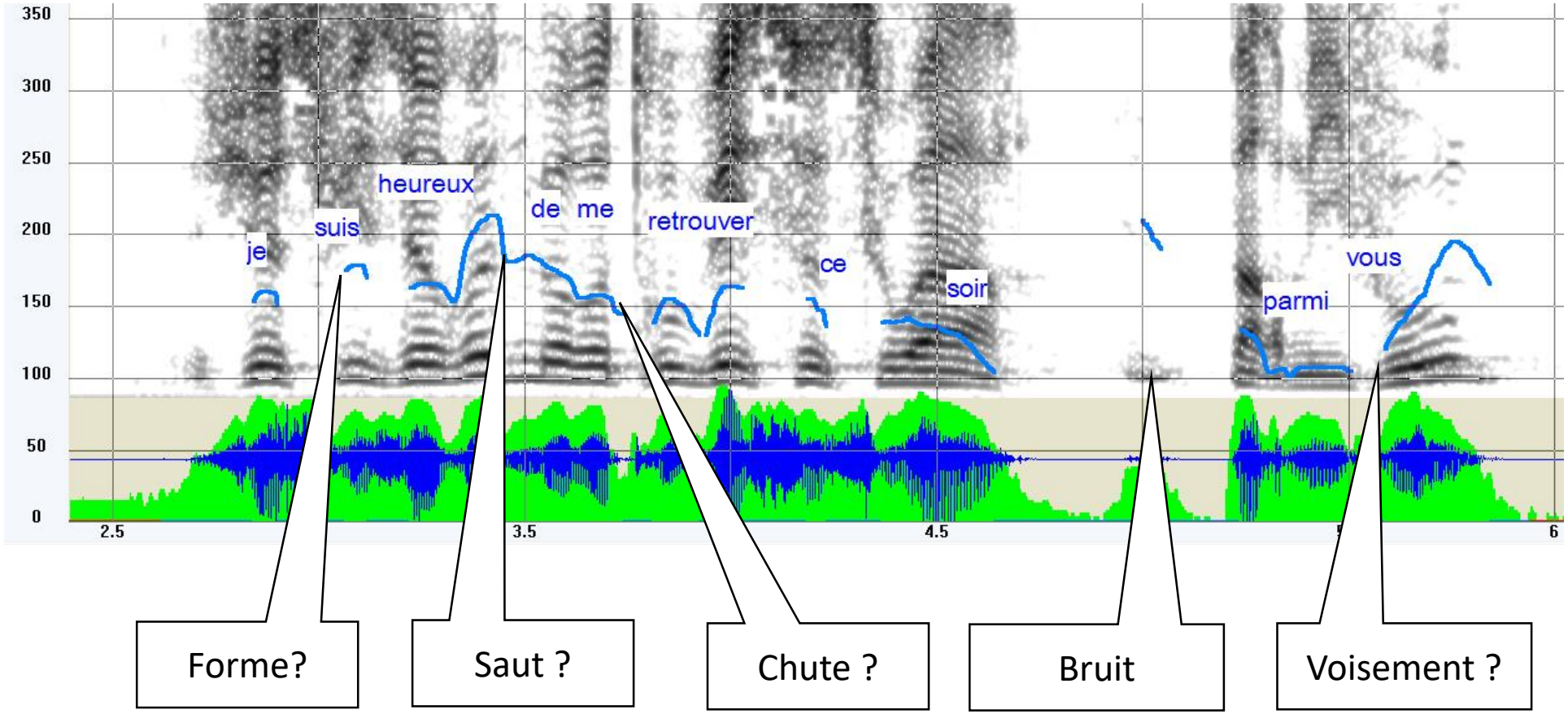
Comparaison de divers algorithmes

Autocorrélation classique



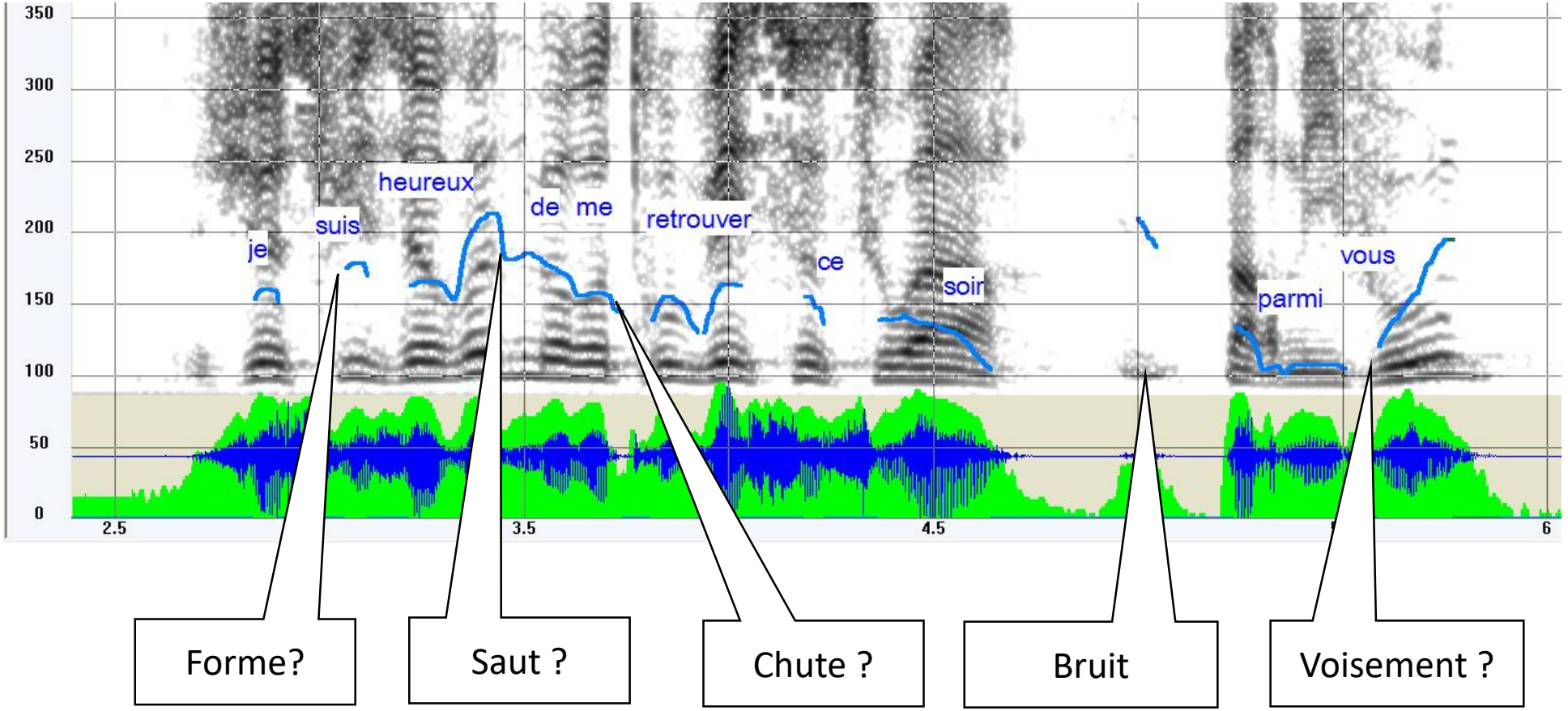
Comparaison de divers algorithmes

Autocorrélation Praat (Hann)



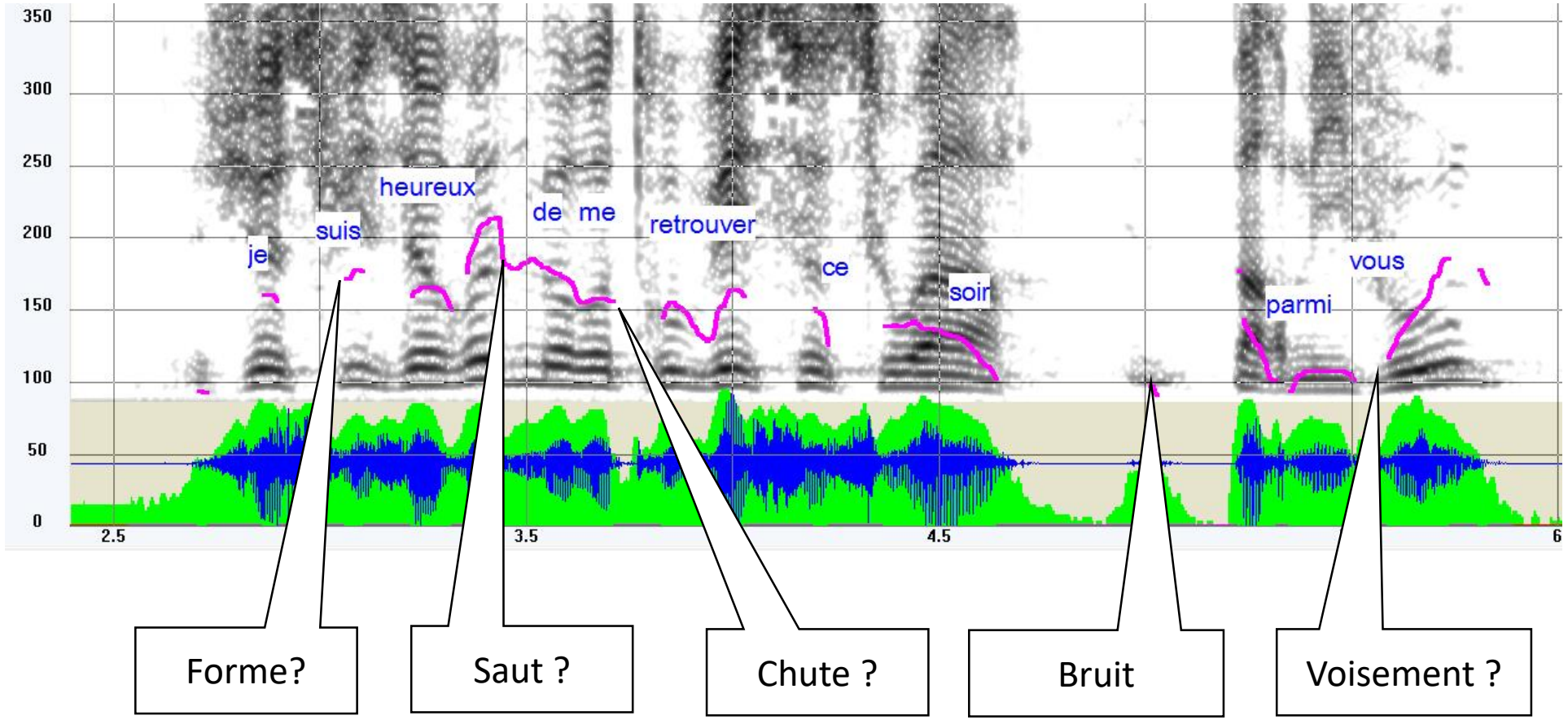
Comparaison de divers algorithmes

Autocorrélation Yin



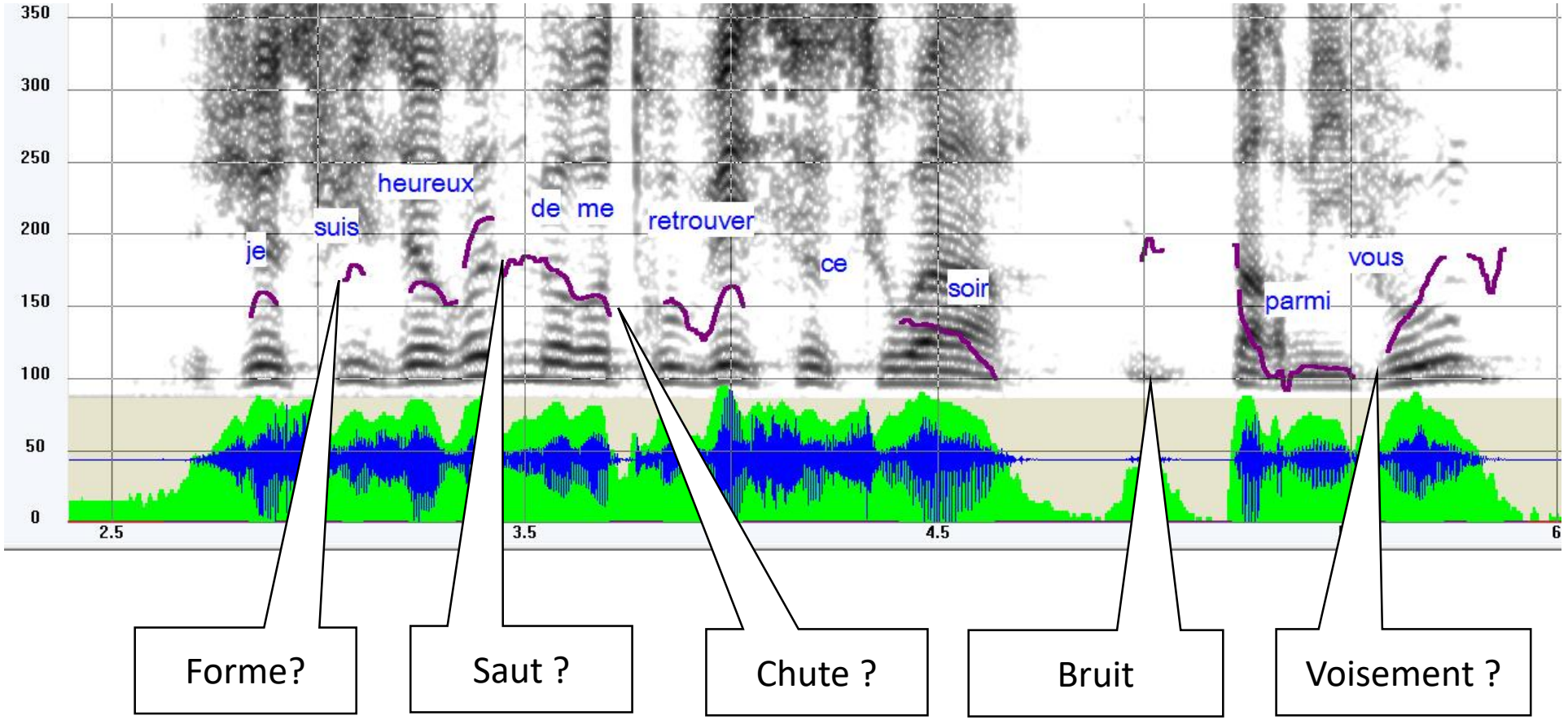
Comparaison de divers algorithmes

AMDF



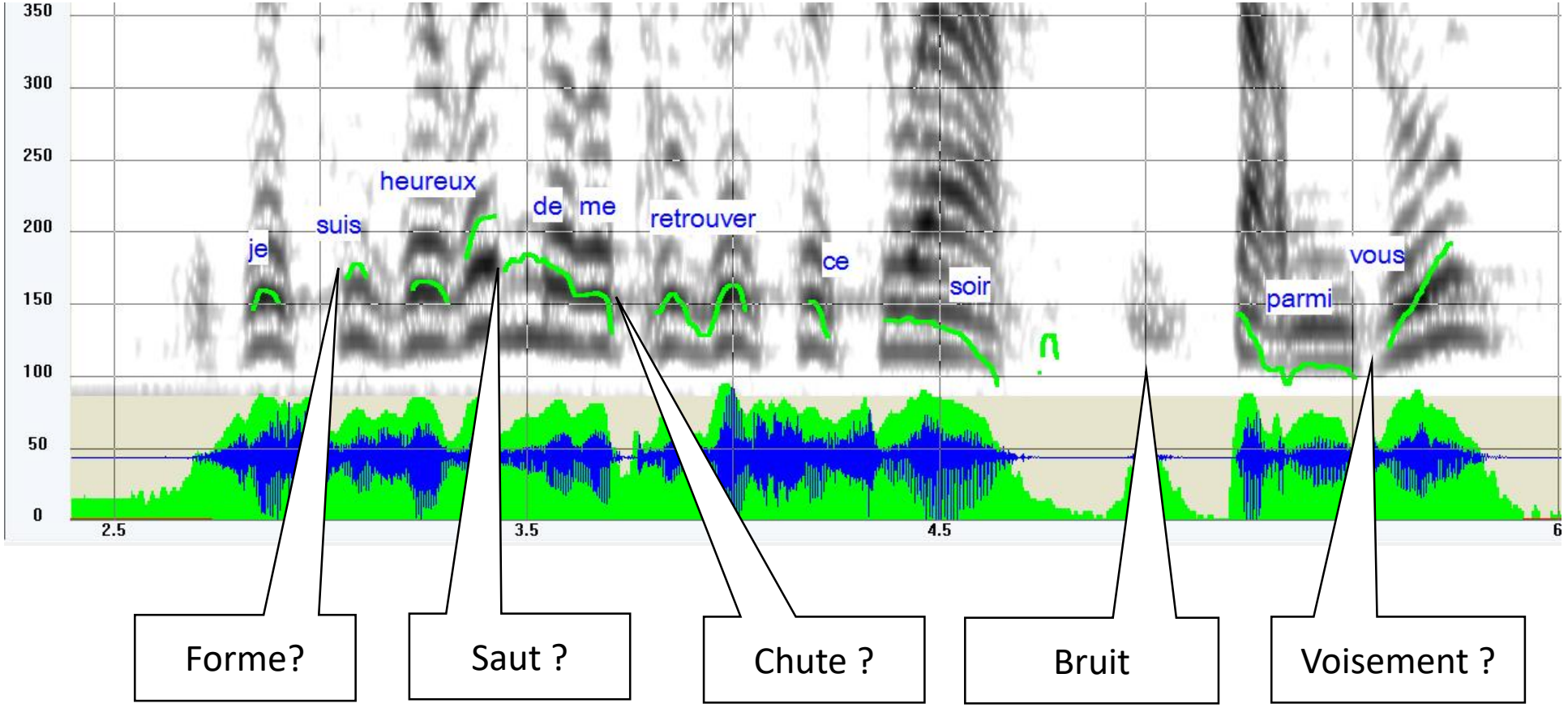
Comparaison de divers algorithmes

La brosse spectrale



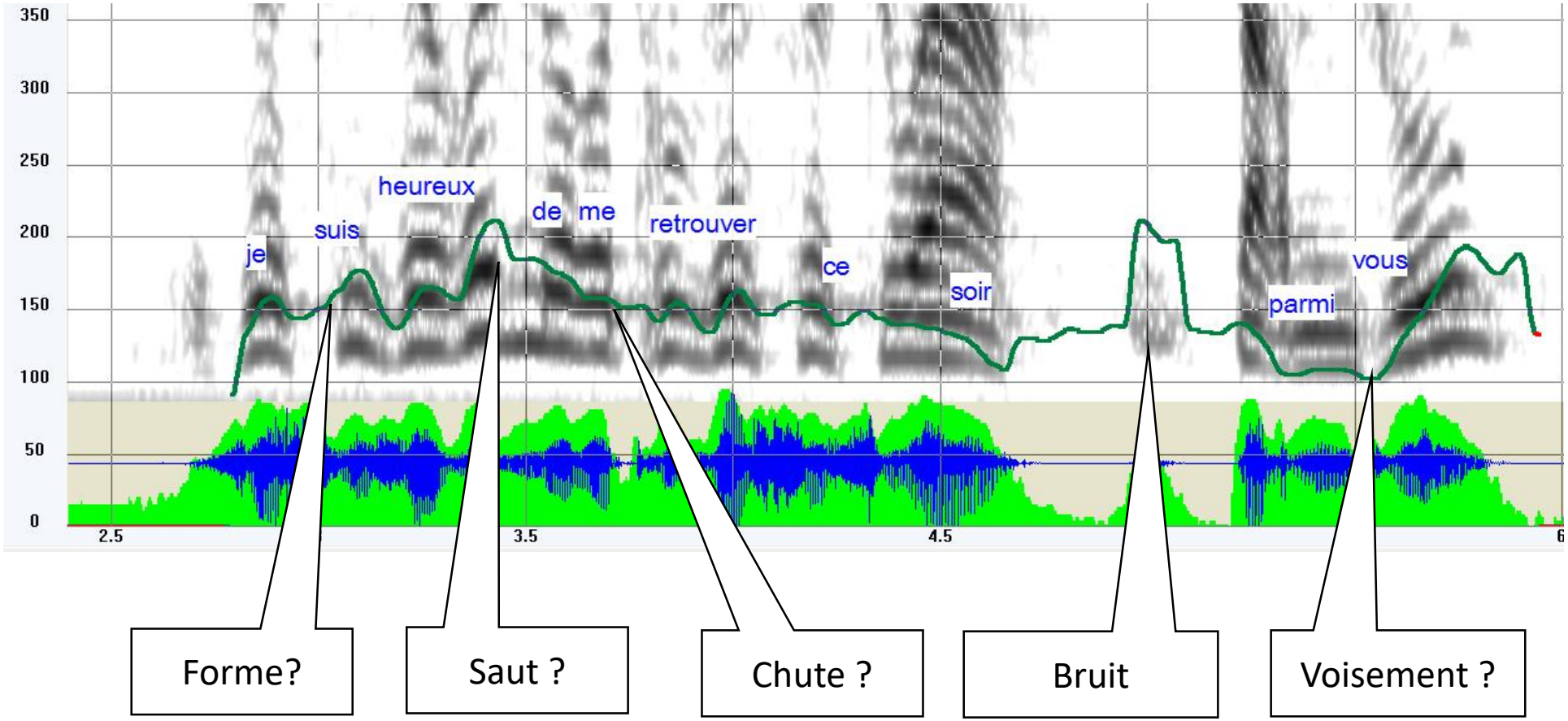
Comparaison de divers algorithmes

Sélection d'harmoniques



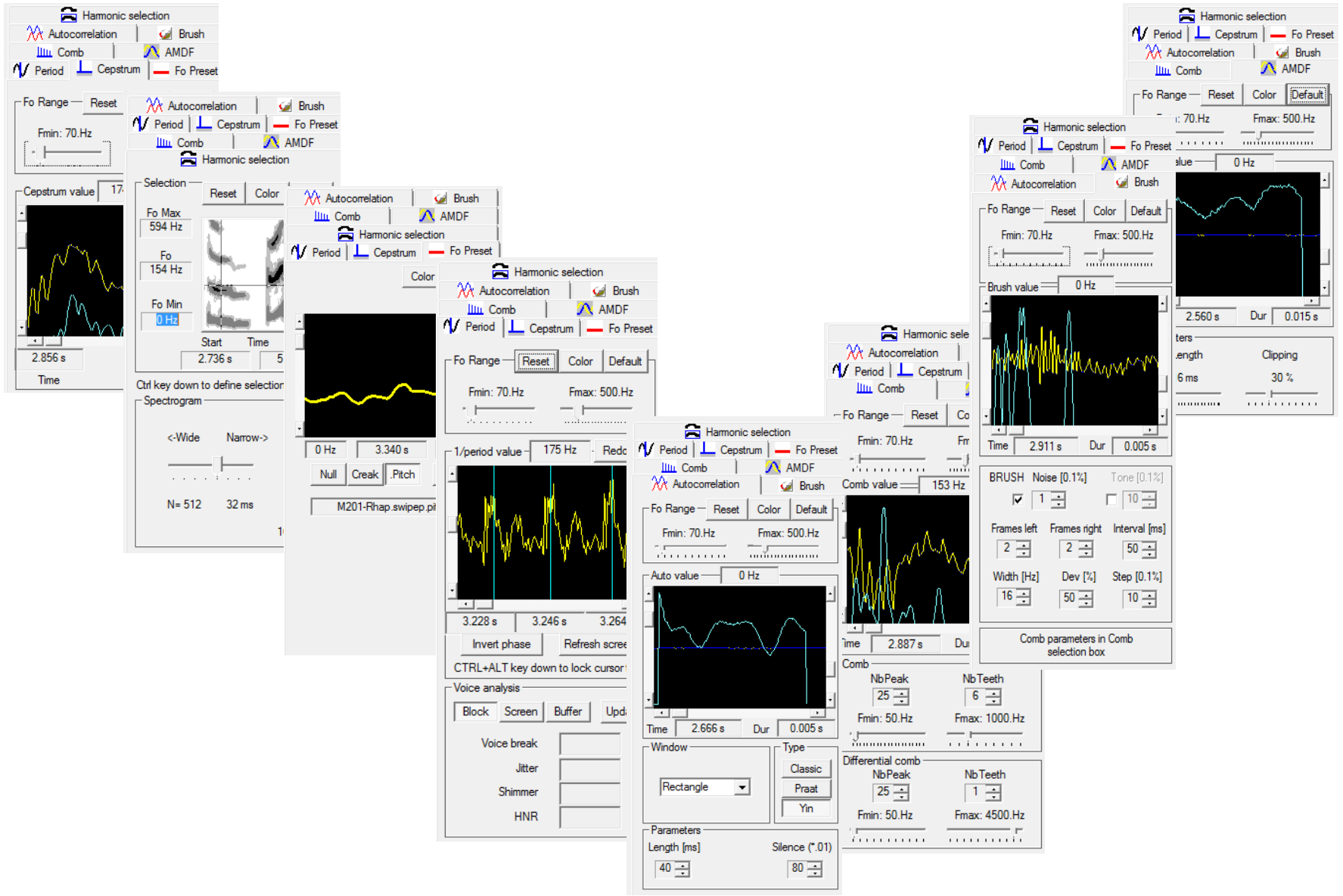
Comparaison de divers algorithmes

Swipep (Ircam)

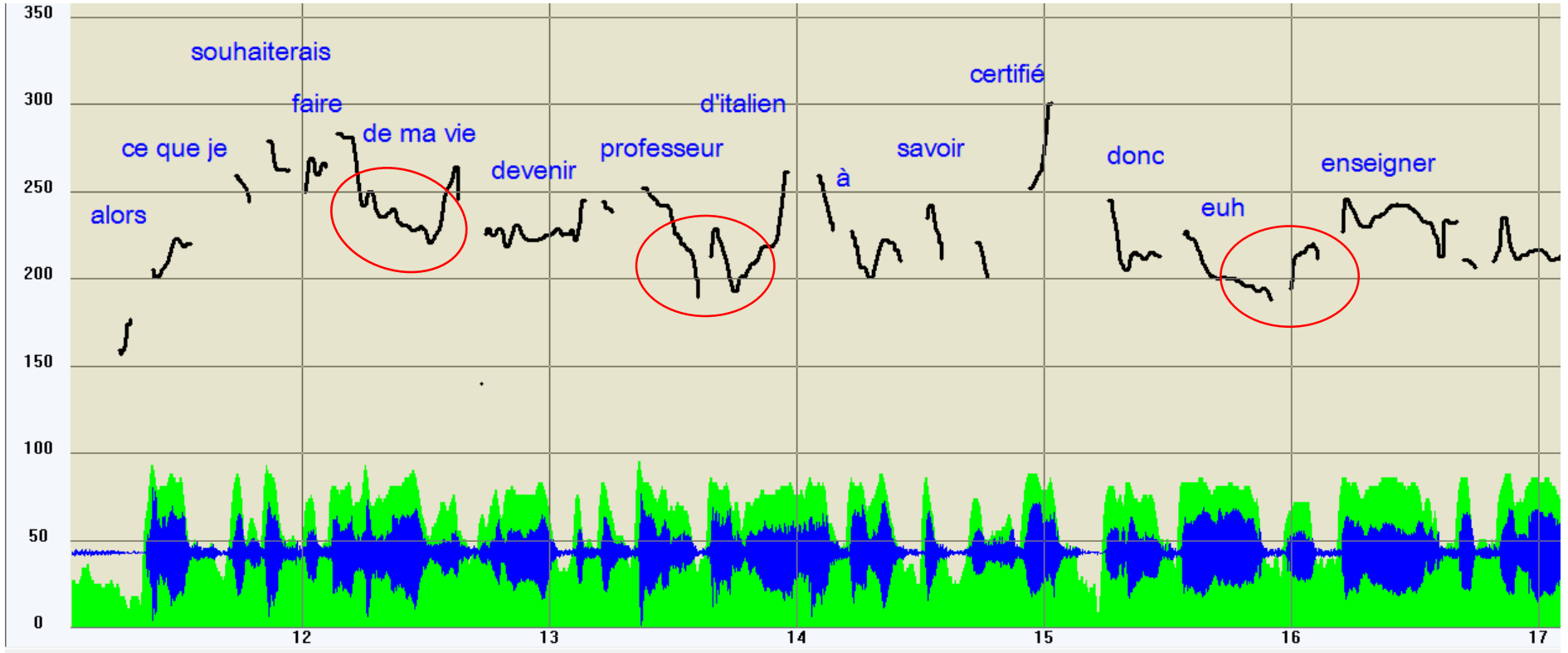


Mais c’était un exemple
d’enregistrement de très bonne qualité !

Algorithmes disponibles...

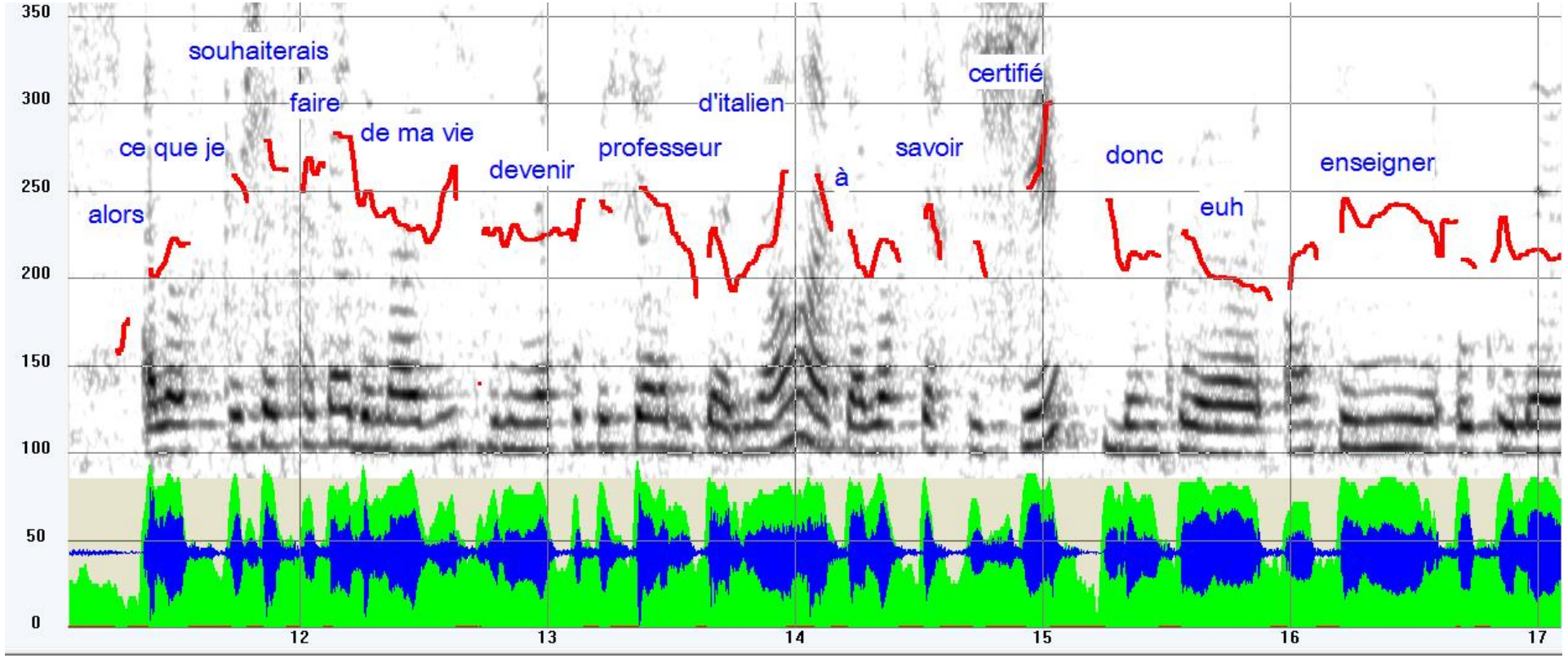


Nettoyage des courbes mélodiques



Peut on se fier à cette courbe mélodique ?

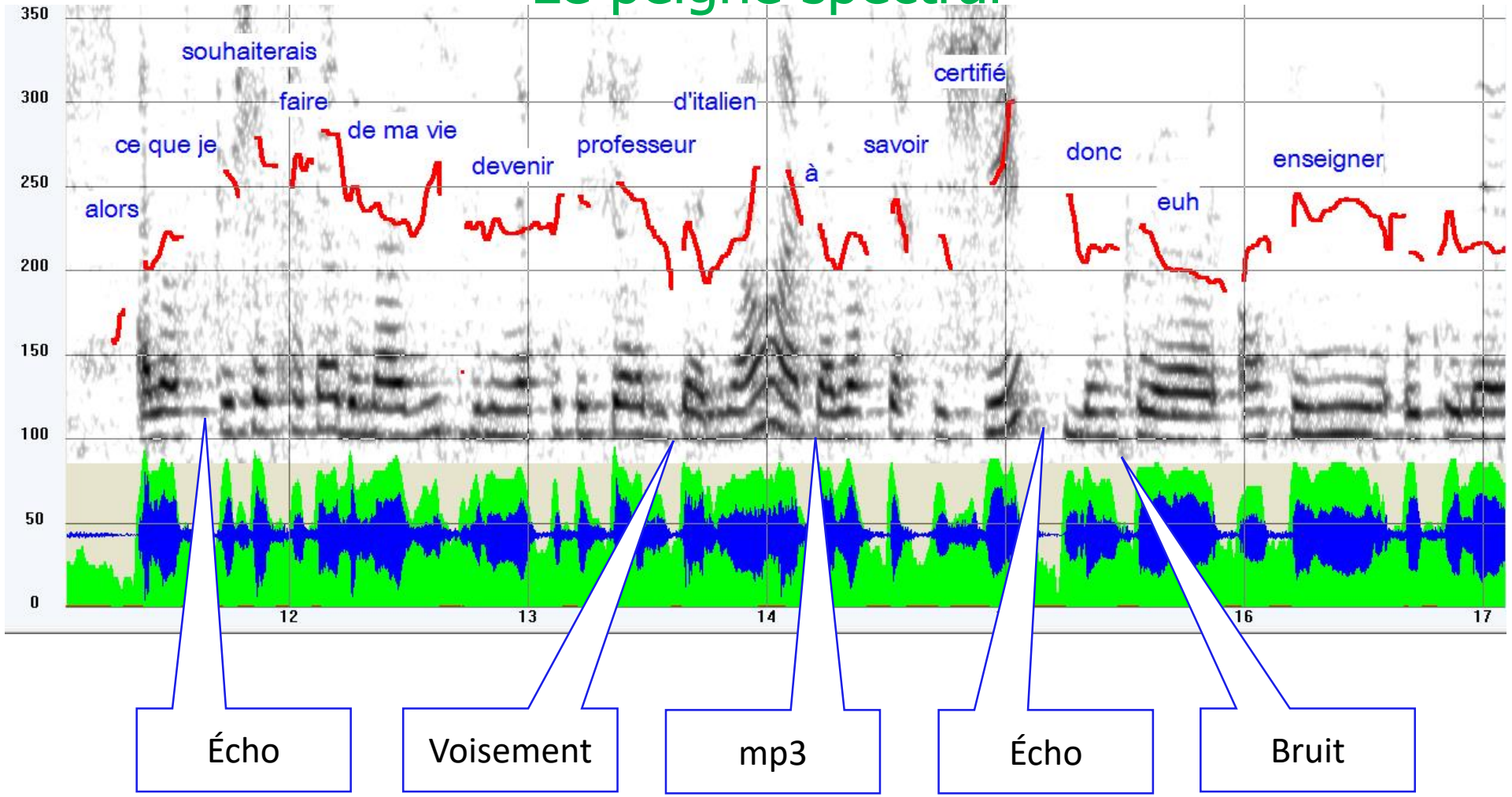
Vérification visuelle au spectrogramme



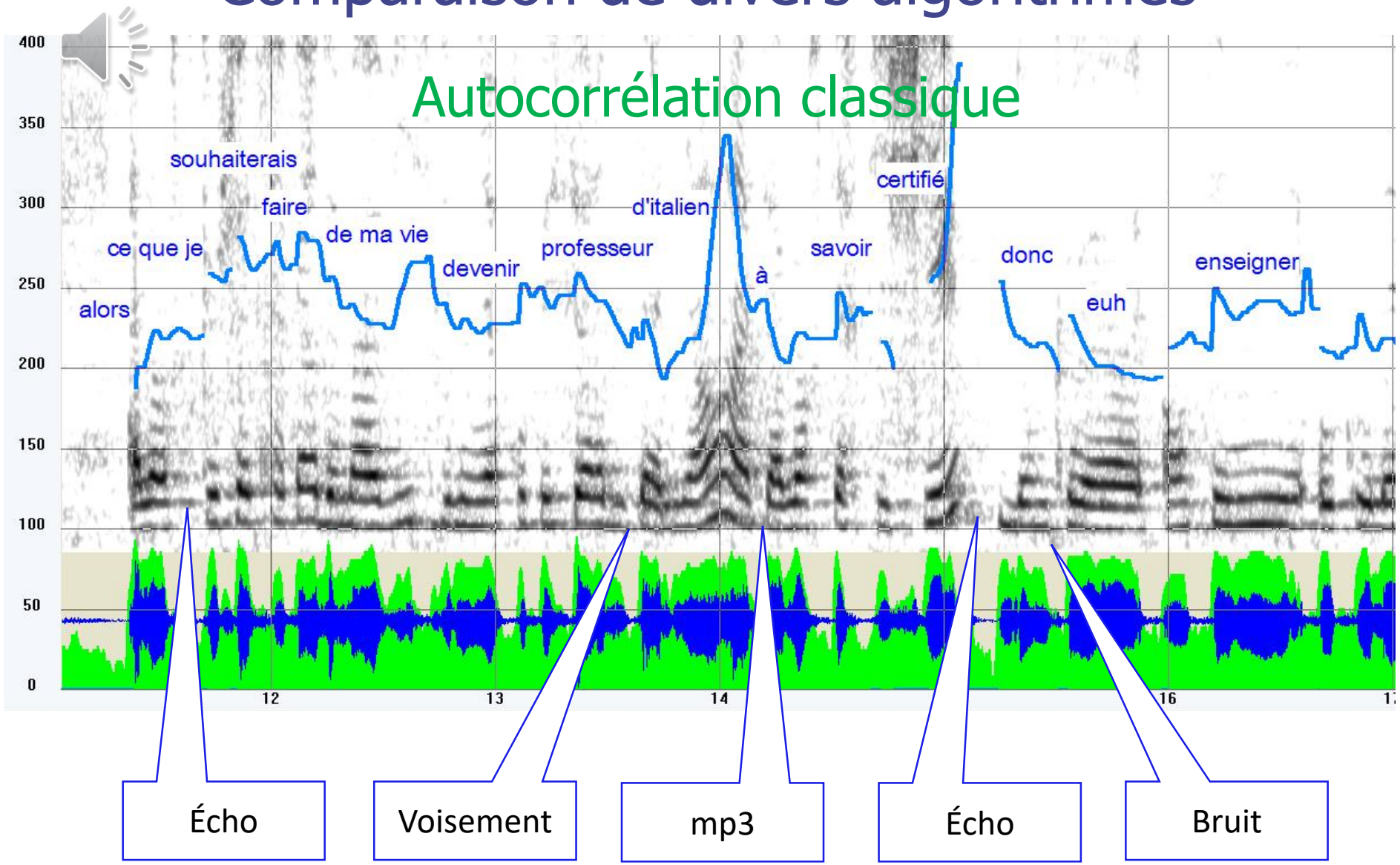
Spectrogramme « bande étroite »-> harmoniques

Comparaison de divers algorithmes

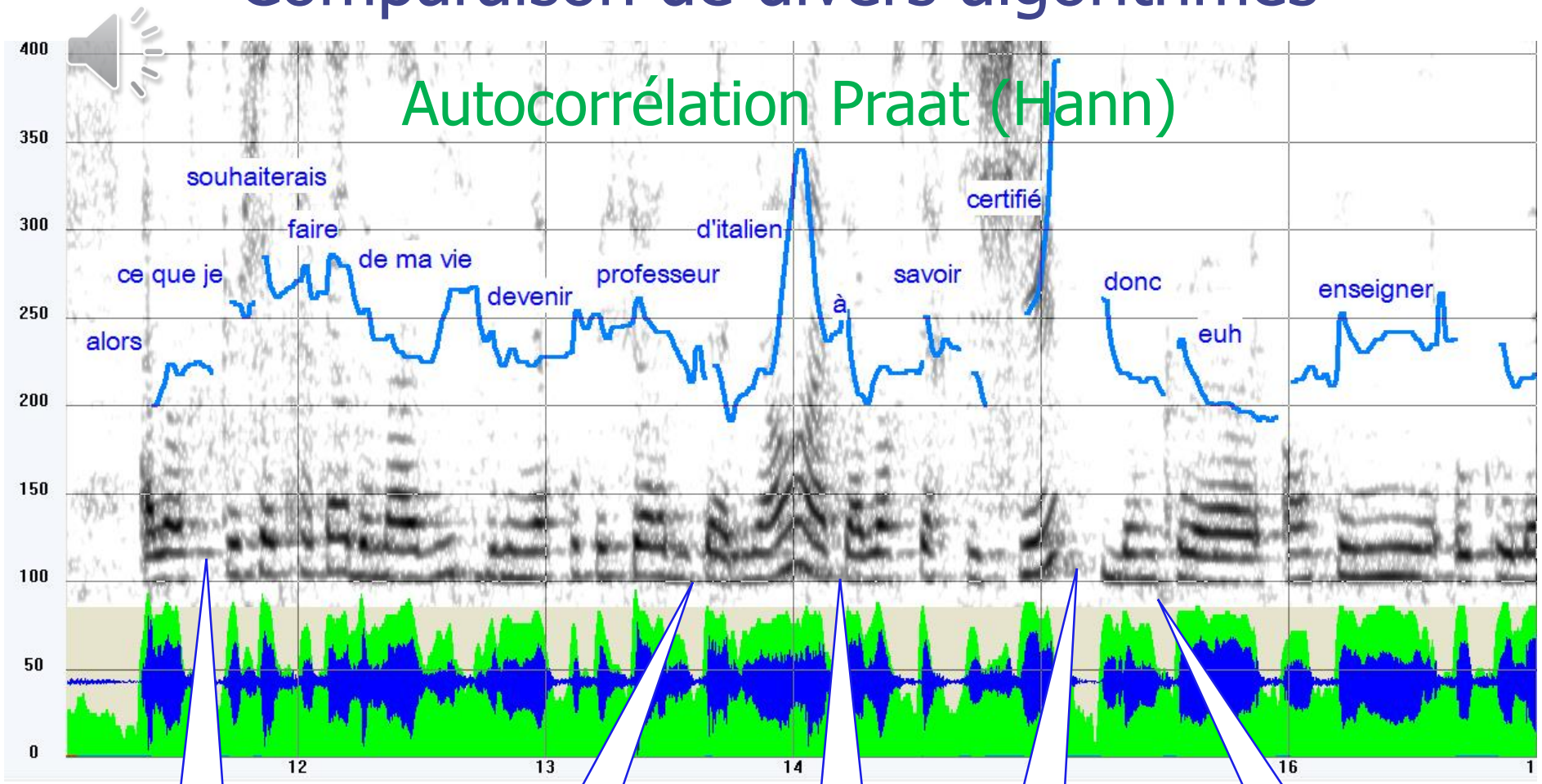
Le peigne spectral



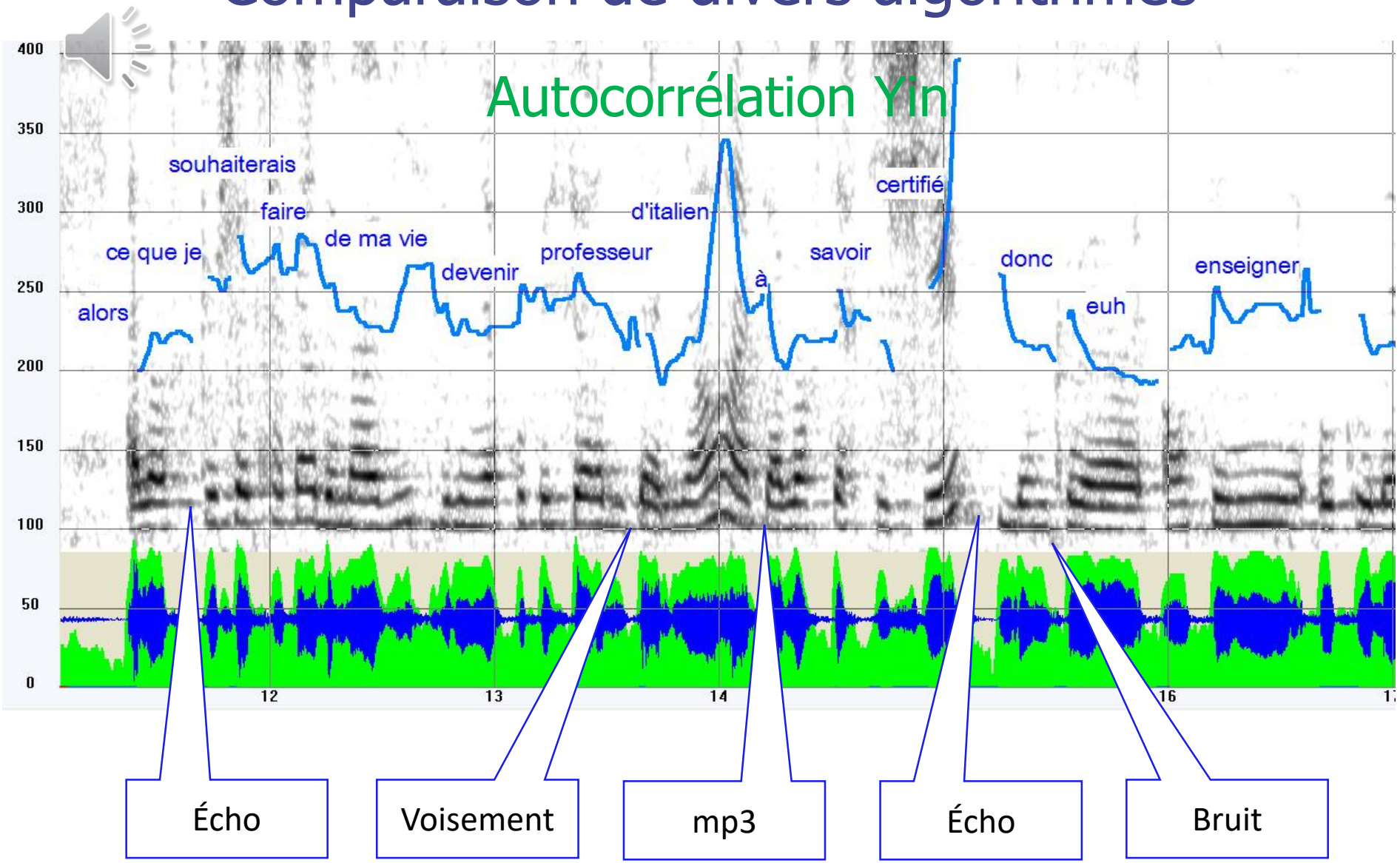
Comparaison de divers algorithmes



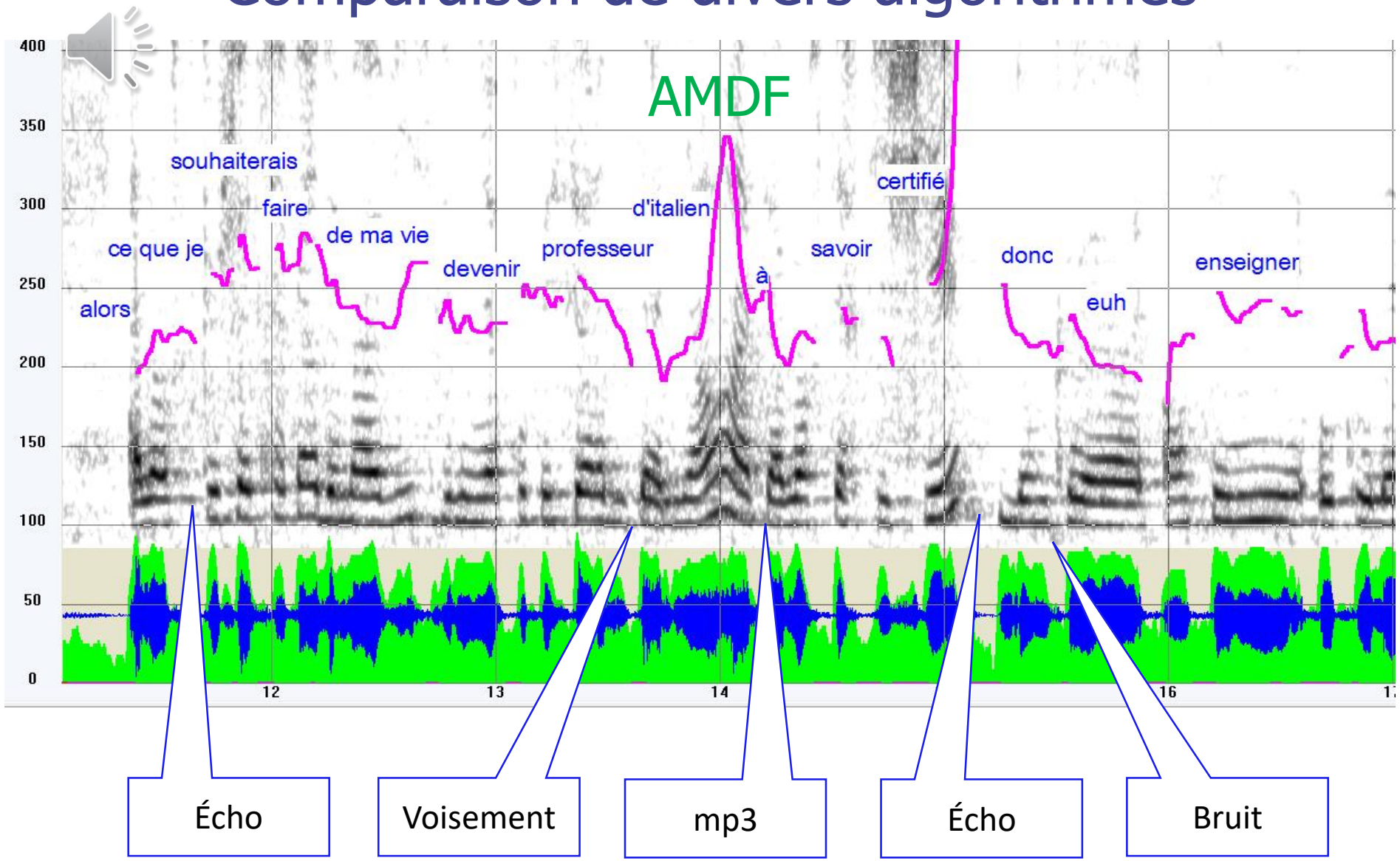
Comparaison de divers algorithmes



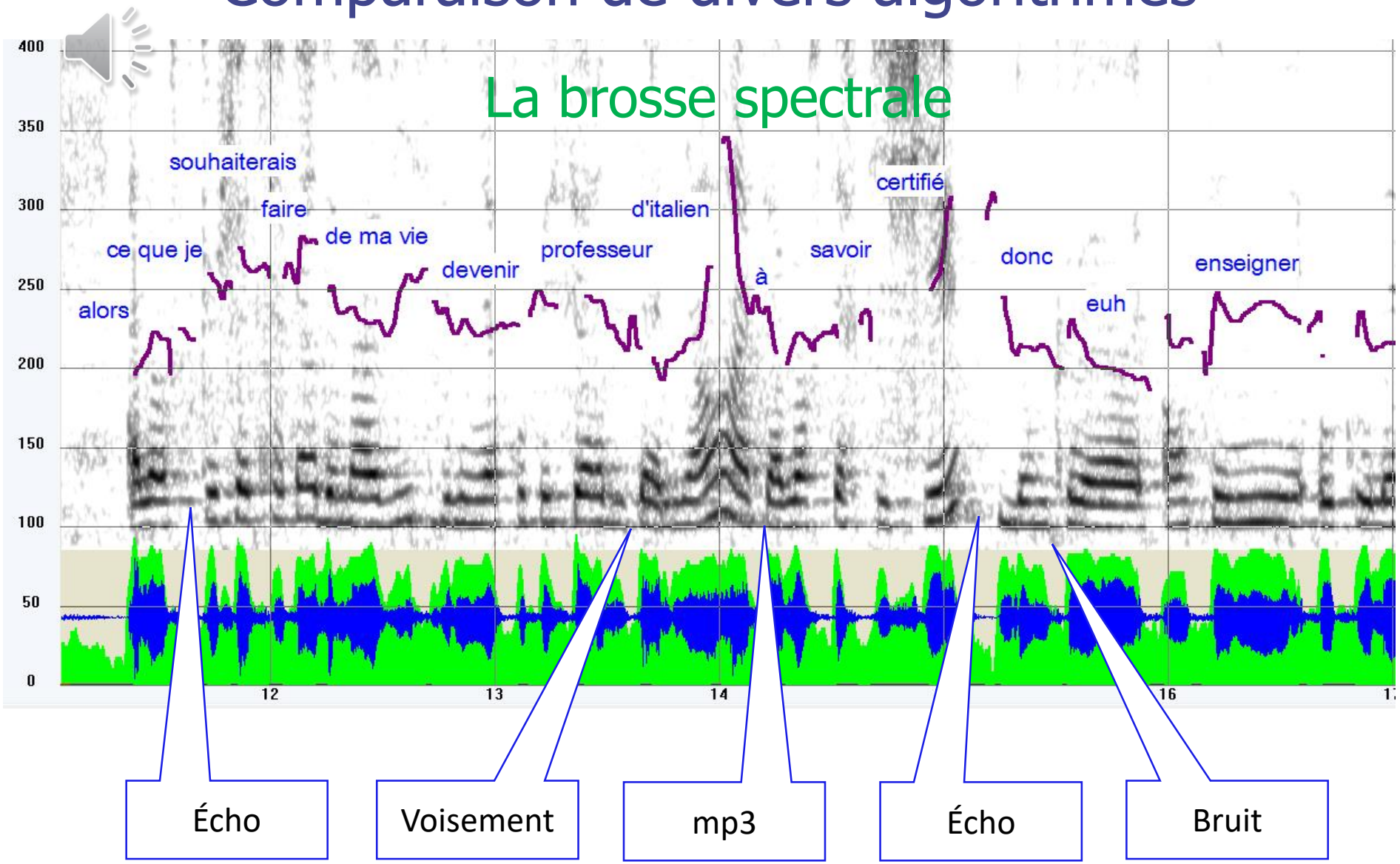
Comparaison de divers algorithmes



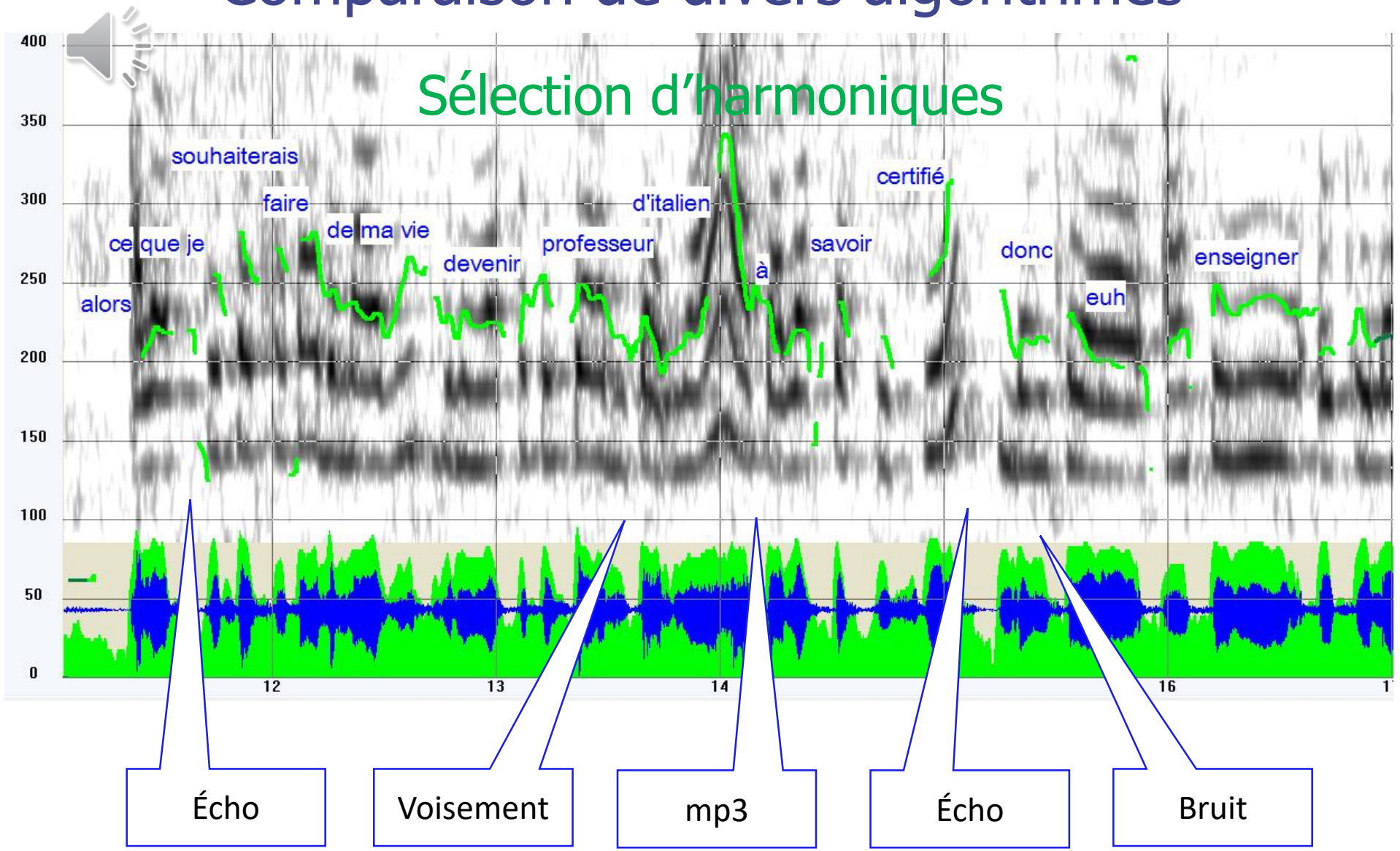
Comparaison de divers algorithmes



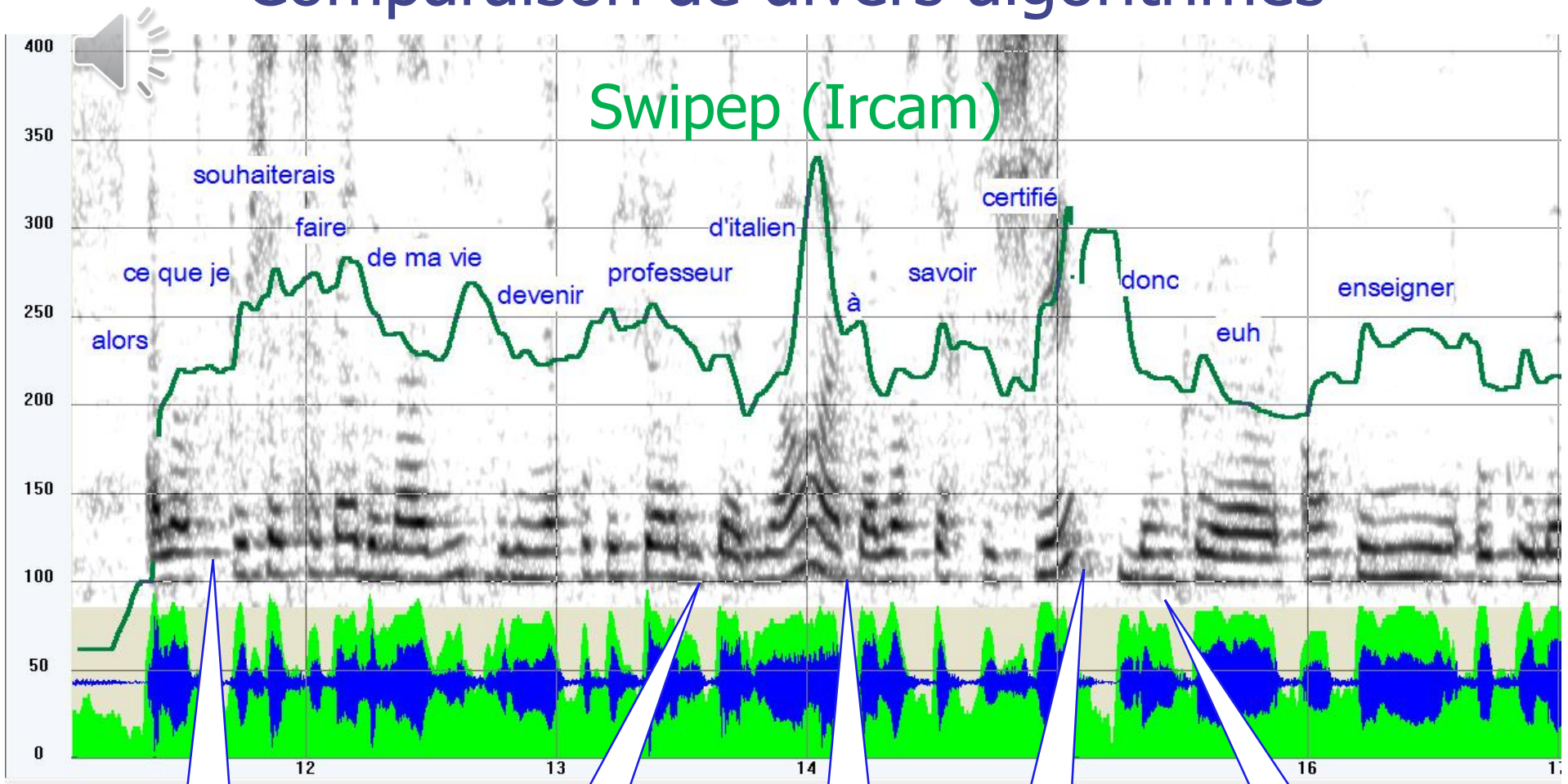
Comparaison de divers algorithmes



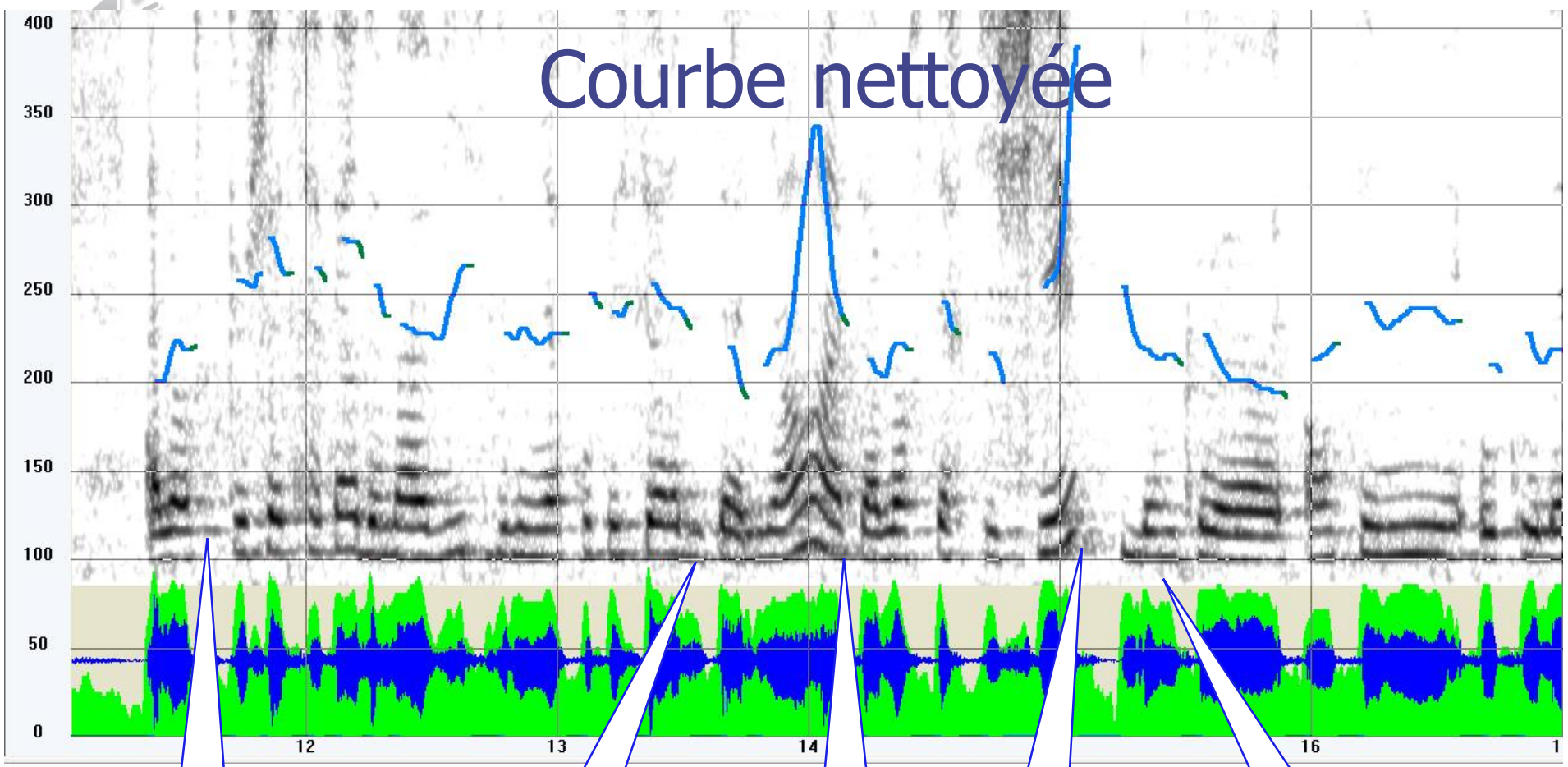
Comparaison de divers algorithmes



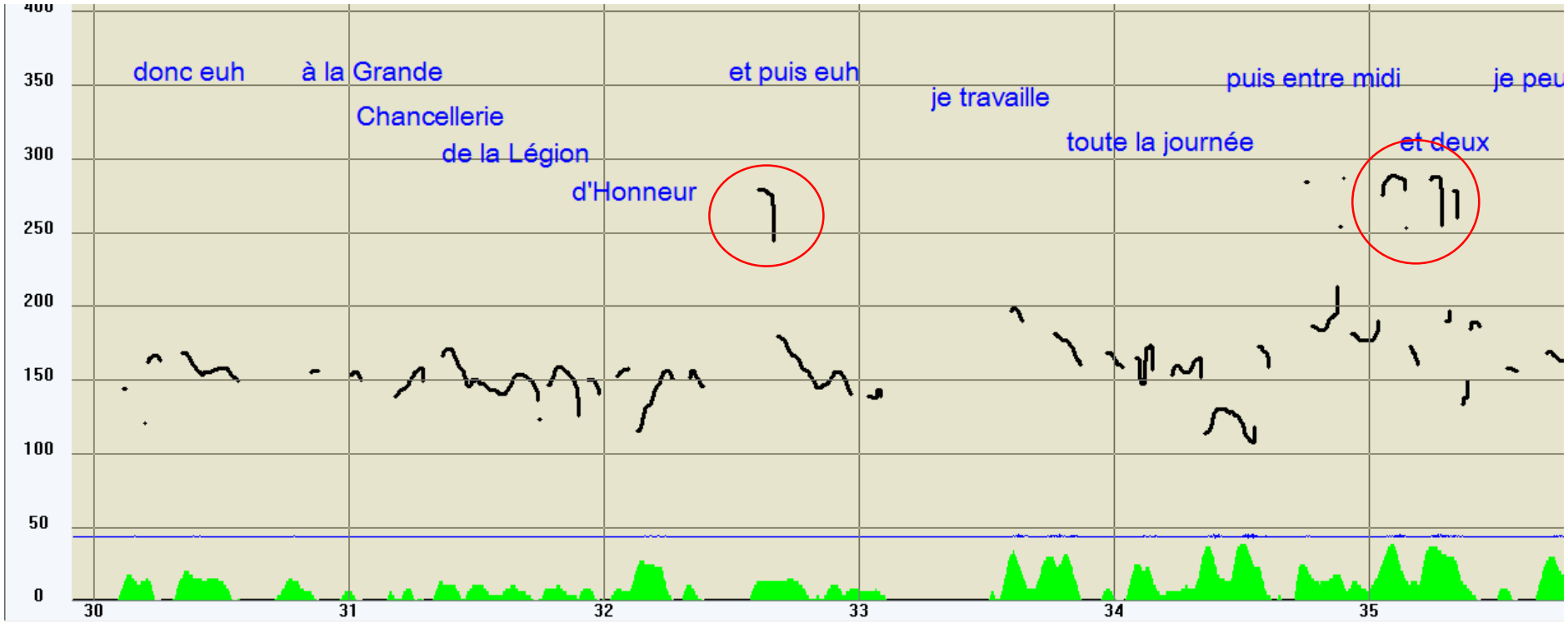
Comparaison de divers algorithmes



Courbe nettoyée

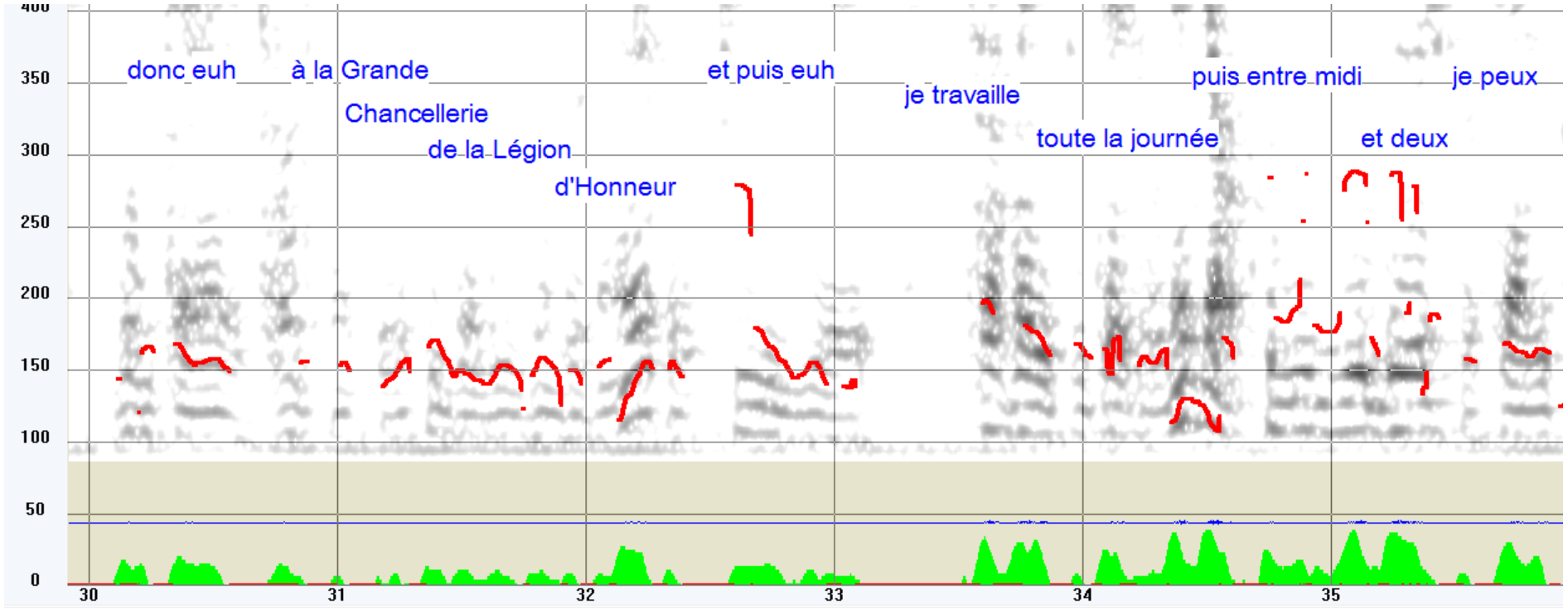


…et quand la musique n’est pas bonne…



Peut on se fier à cette courbe mélodique ?

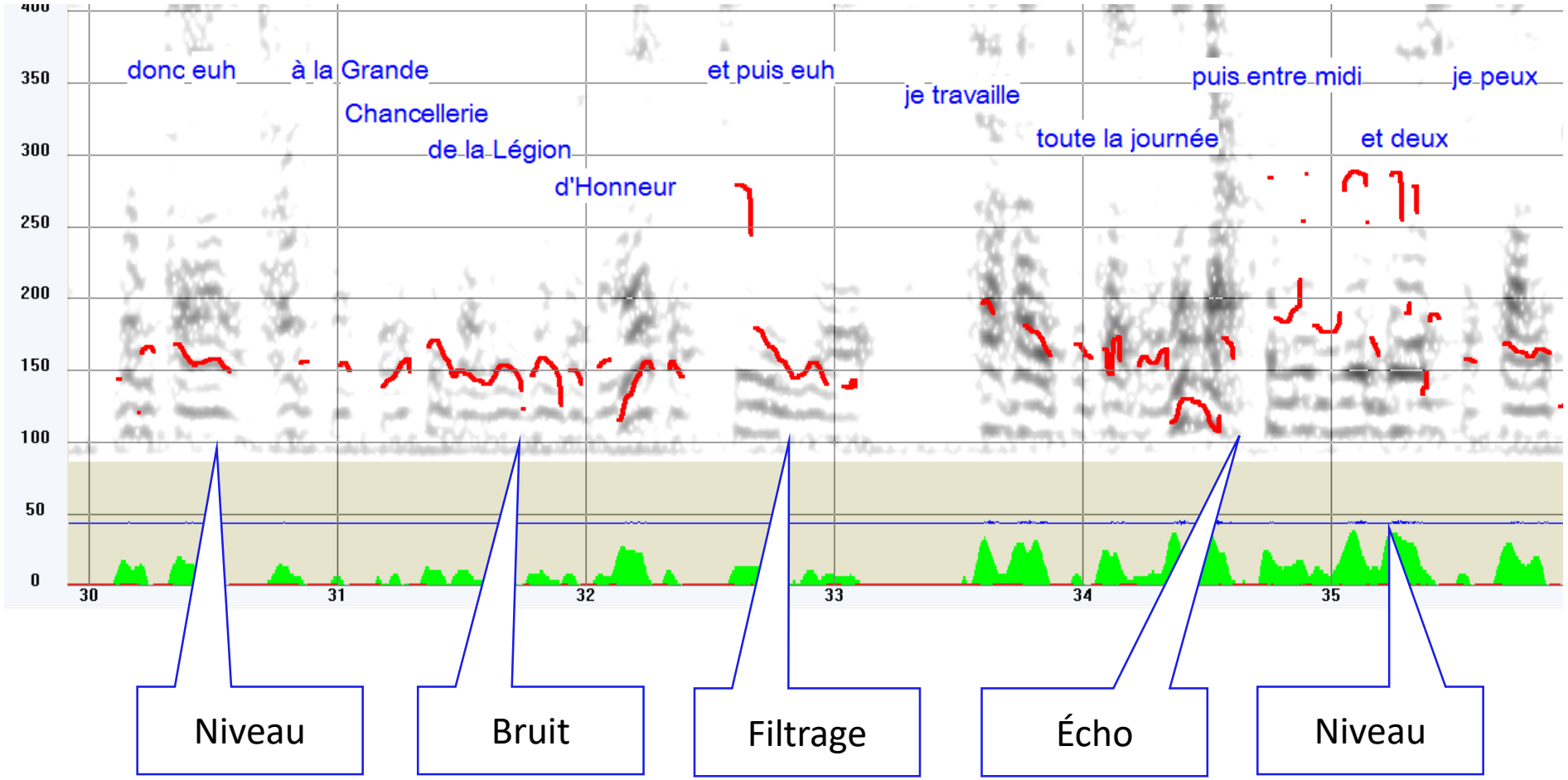
Vérification visuelle au spectrogramme



Spectrogramme « bande étroite »-> harmoniques

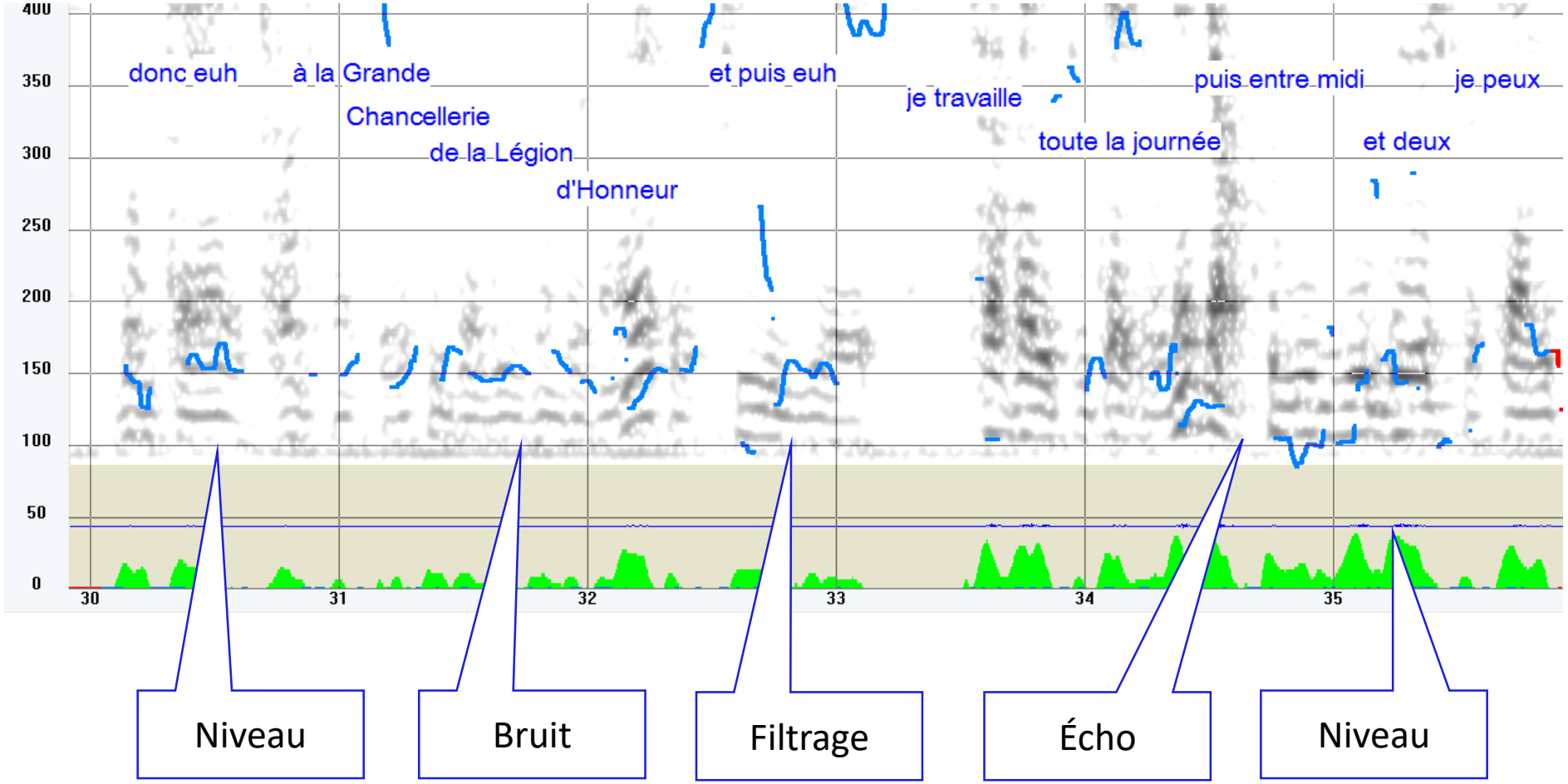
Comparaison de divers algorithmes

Le peigne spectral



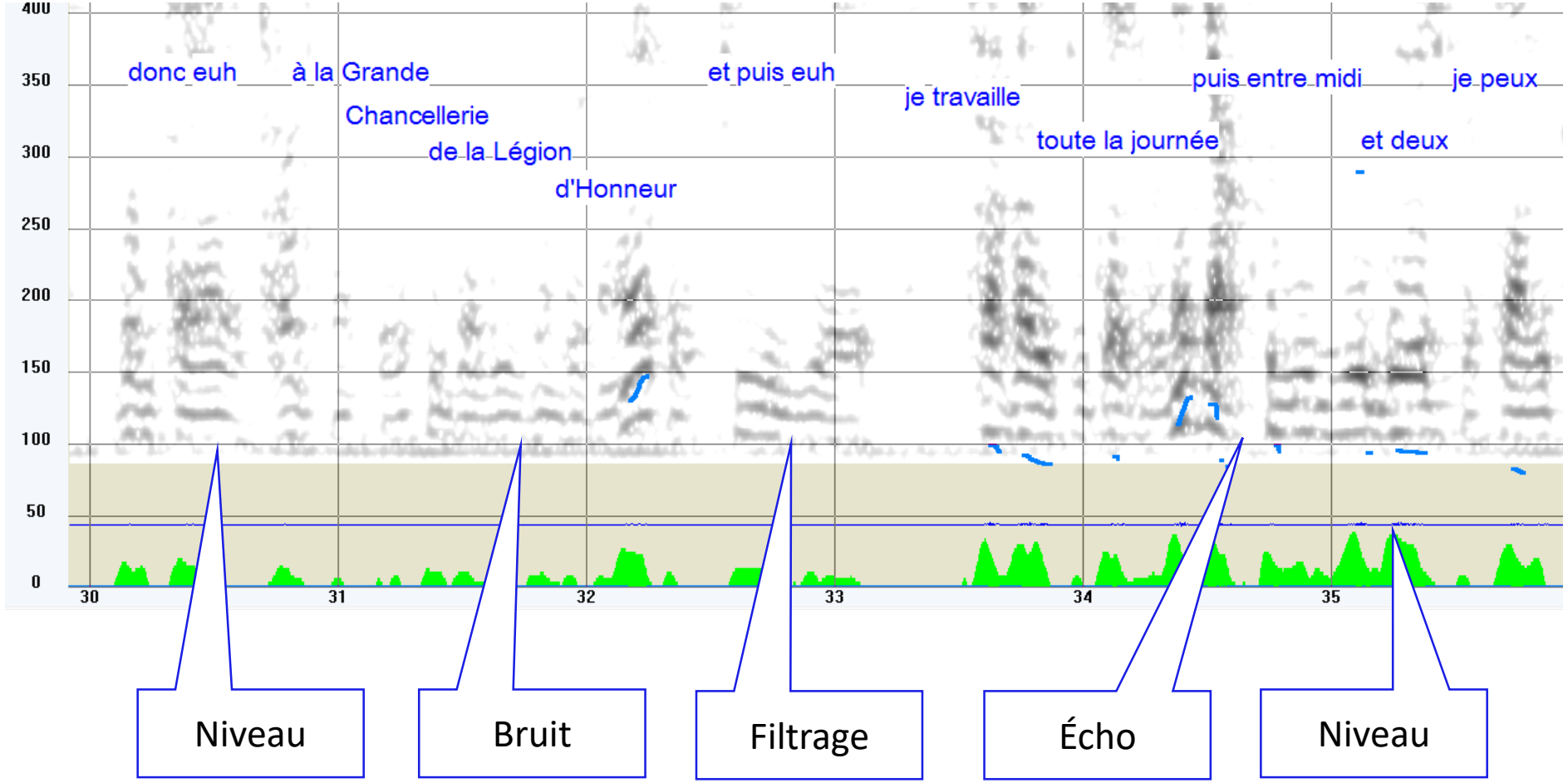
Comparaison de divers algorithmes

Autocorrélation classique



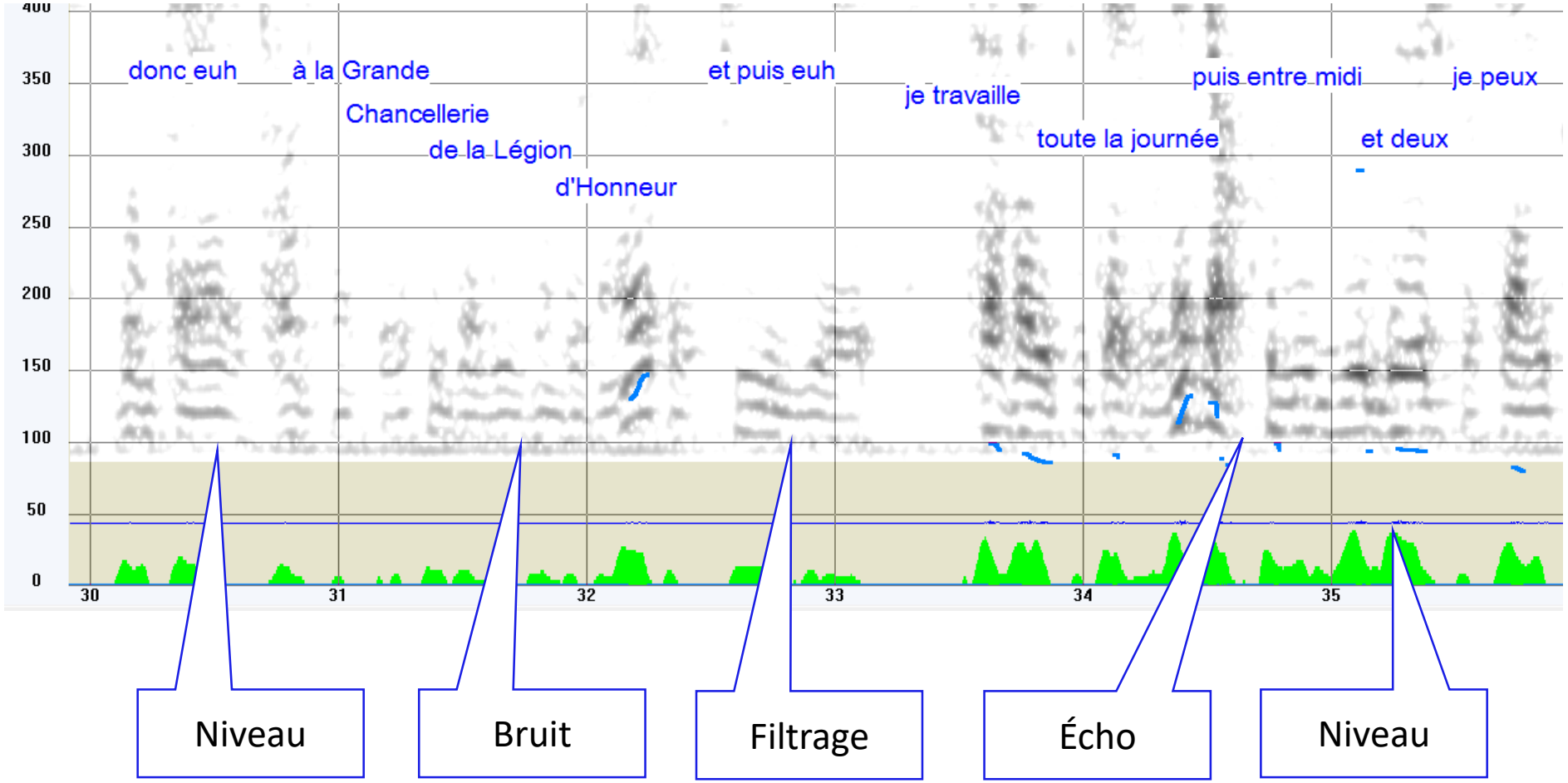
Comparaison de divers algorithmes

Autocorrélation Praat (Hann)



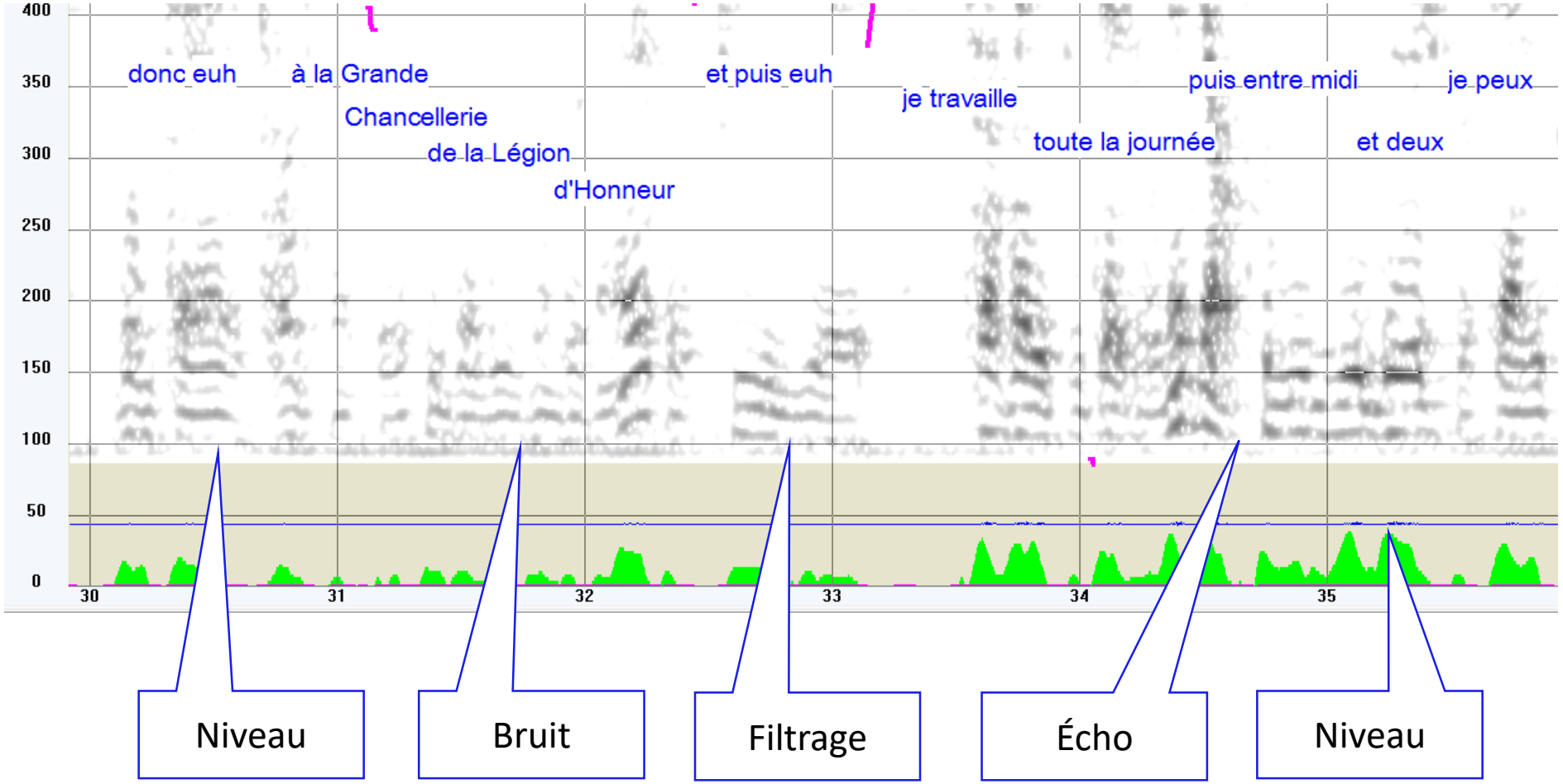
Comparaison de divers algorithmes

Autocorrélation Yin



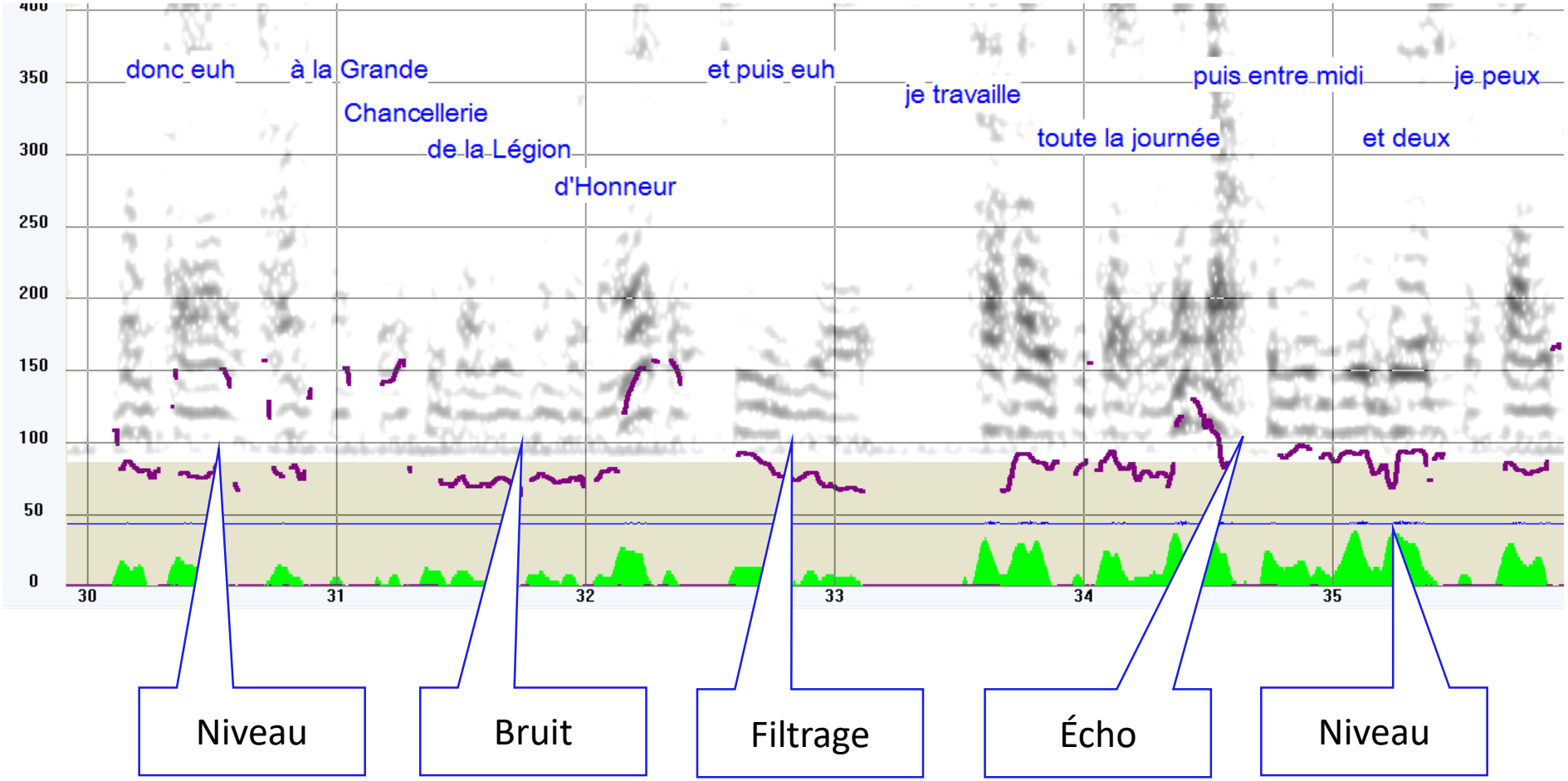
Comparaison de divers algorithmes

AMDF



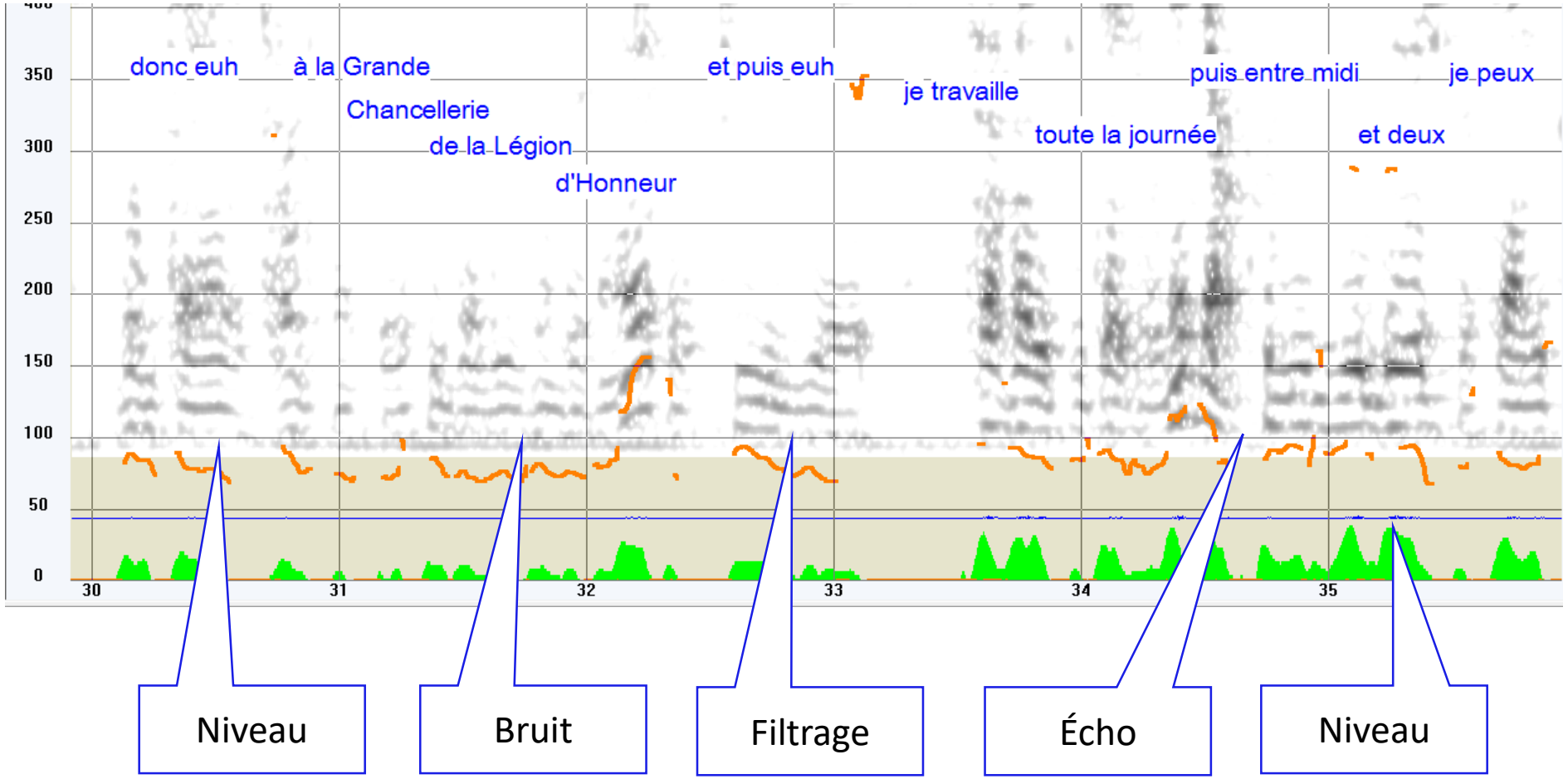
Comparaison de divers algorithmes

La brosse spectrale



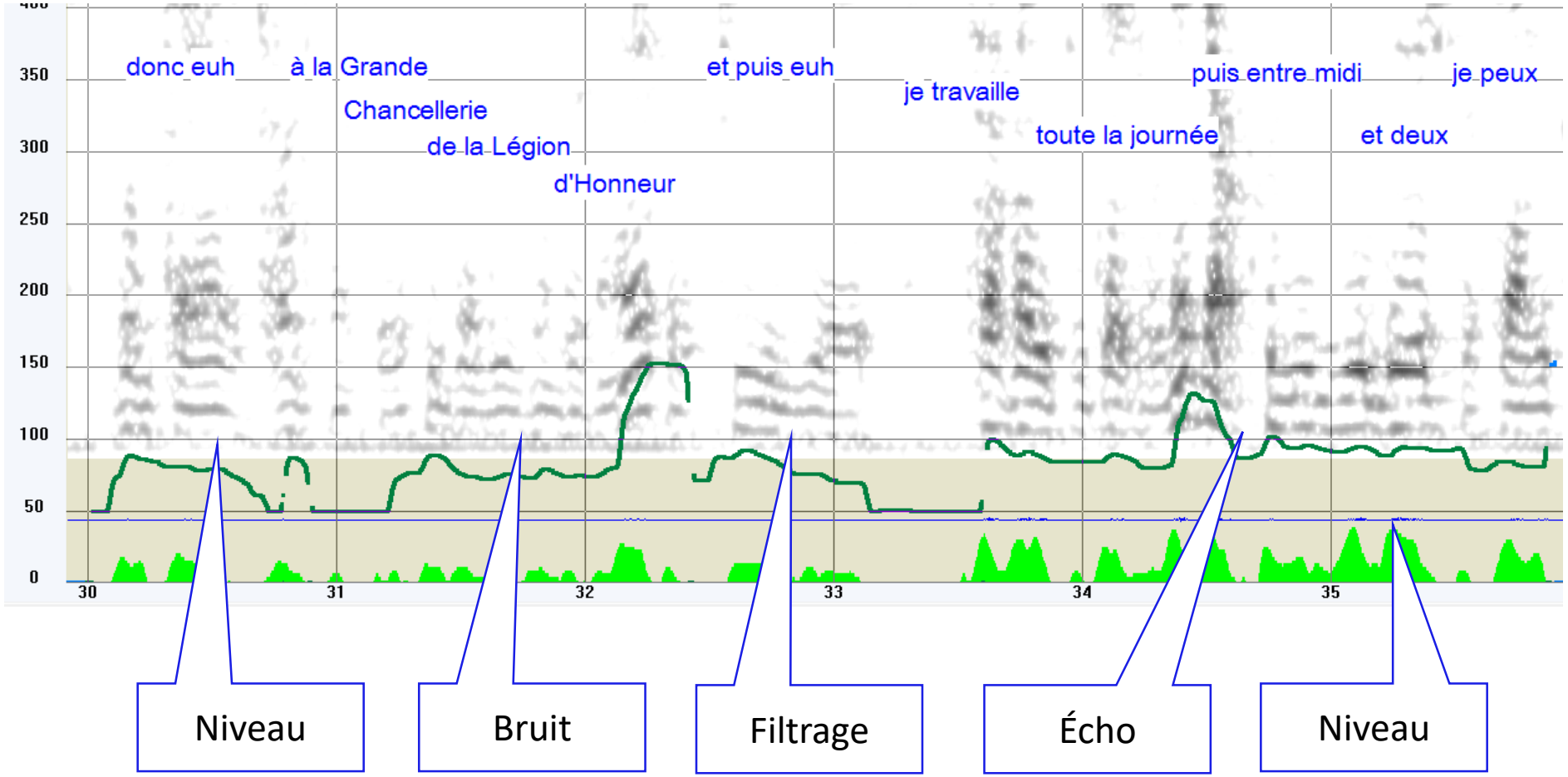
Comparaison de divers algorithmes

Sélection d'harmoniques

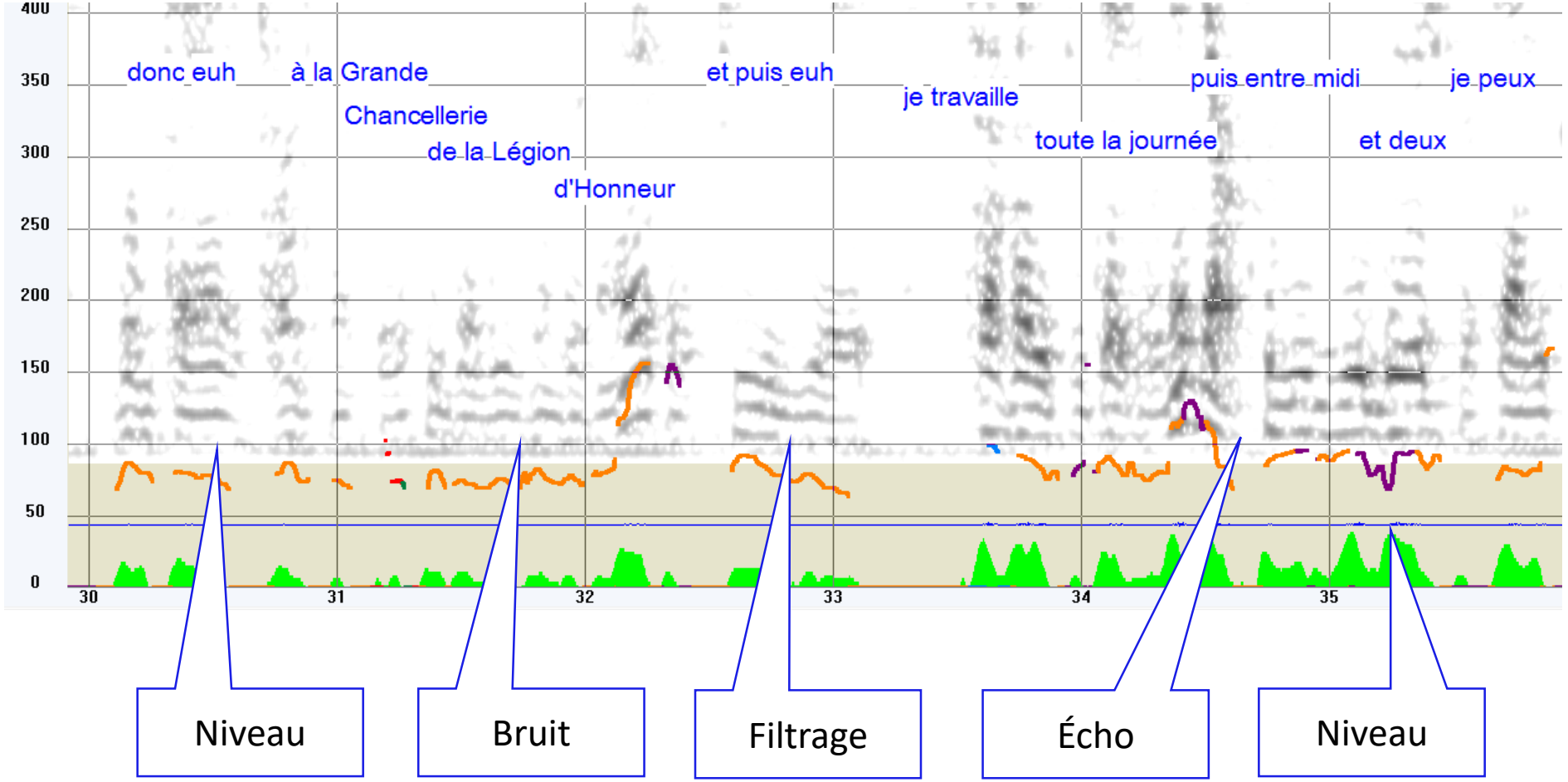


Comparaison de divers algorithmes

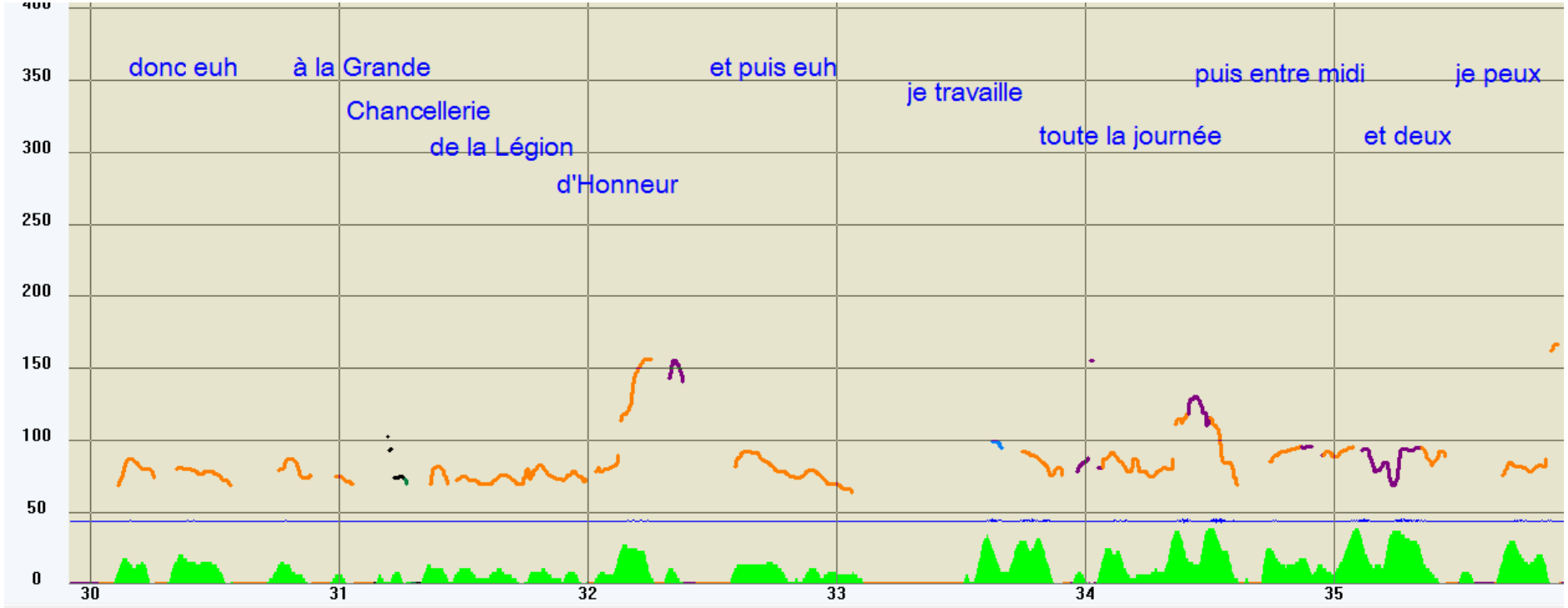
Swipep (Ircam)



Courbe nettoyée



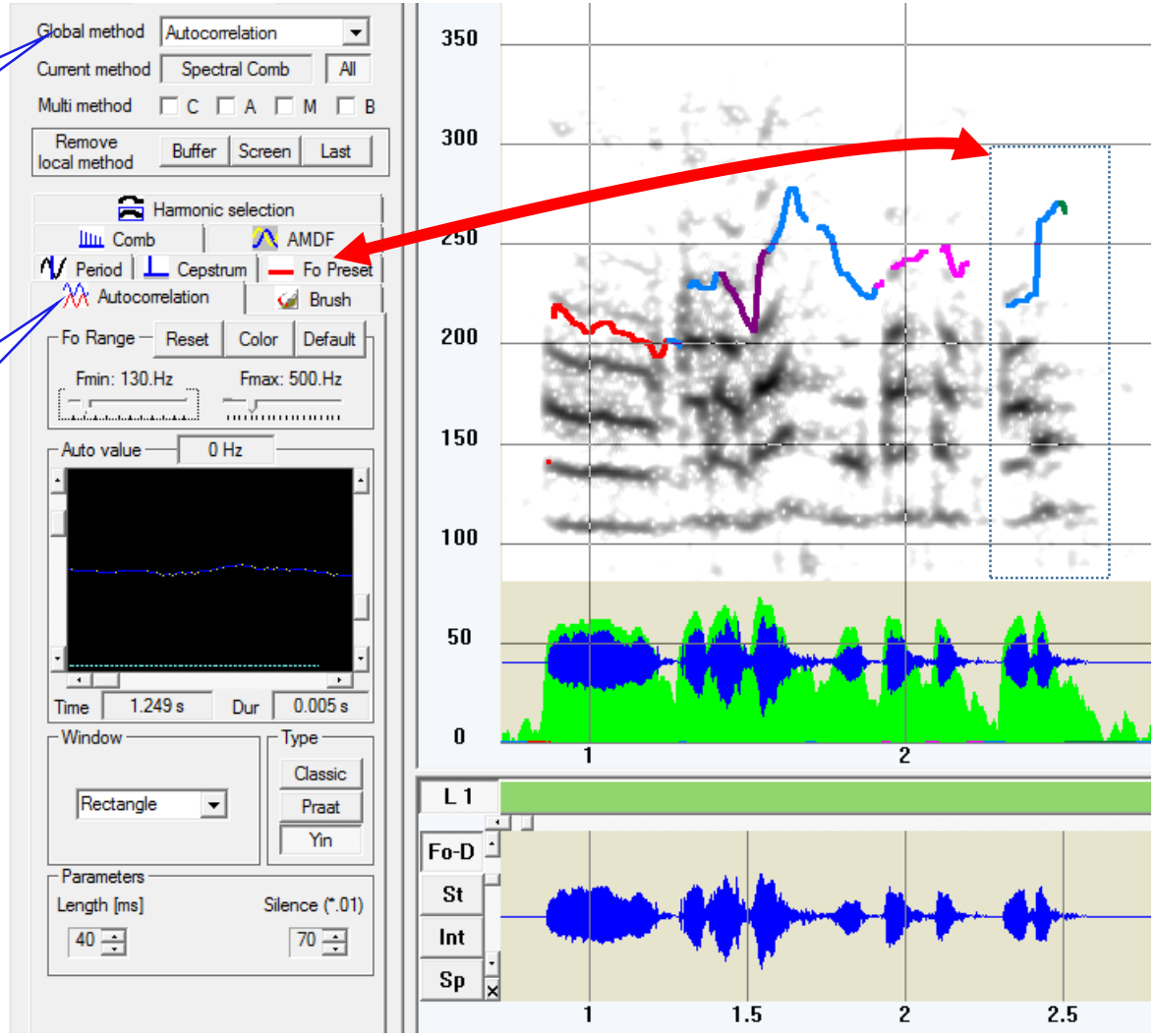
Courbe nettoyée



Nettoyage par sélection d'algorithmes sur des segments de parole

Algorithme global

Algorithme local

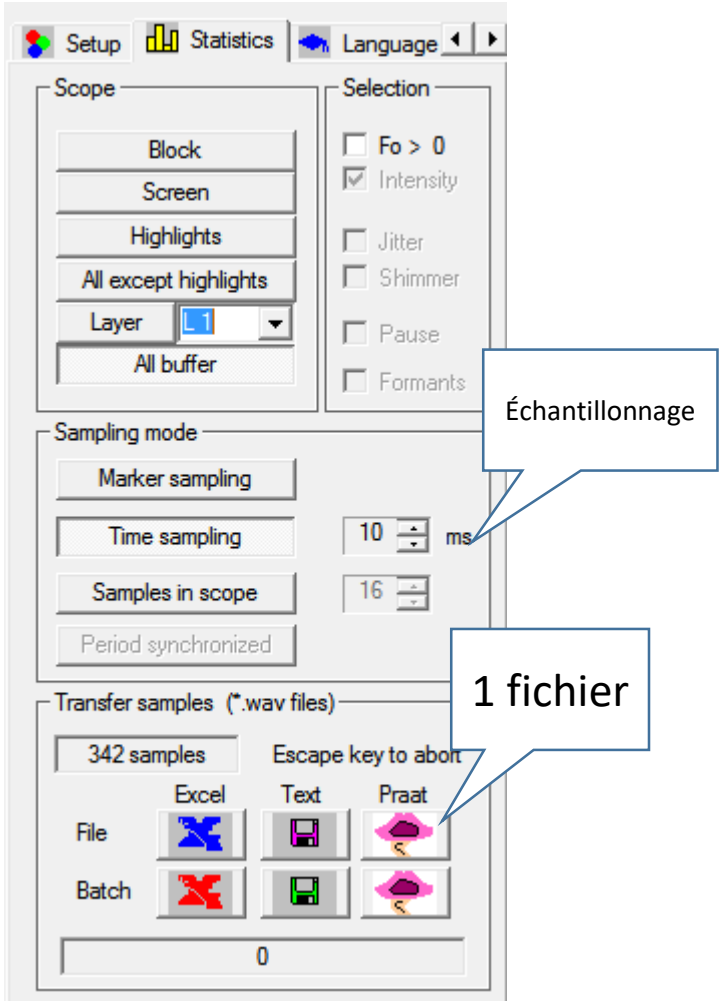


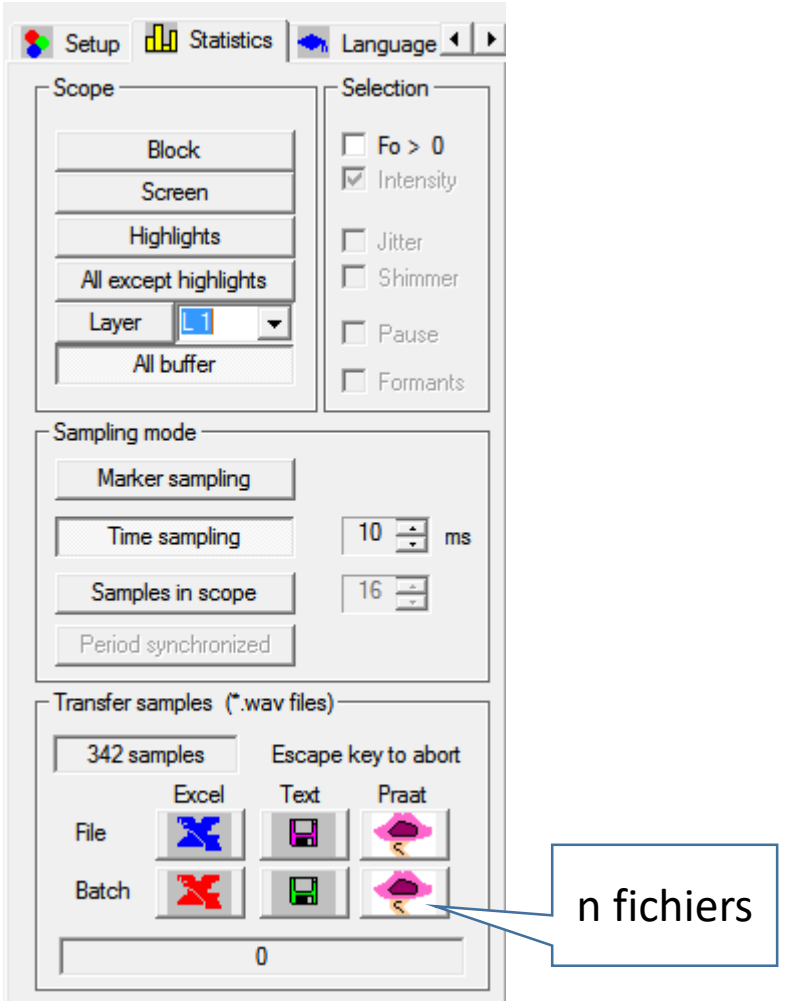
D103

Nettoyage par sélection d'algorithmes sur des segments de parole

```
<TrackingSelection>
        <TrackingRank>28</TrackingRank>
        <FZeroTracking>Null</FZeroTracking>
        <StartIndex>43492</StartIndex>
        <EndIndex>45490</EndIndex>
        <BottomFrequency>578</BottomFrequency>
        <TopFrequency>661</TopFrequency>
        <FZeroMin>117</FZeroMin>
        <FZeroMax>500</FZeroMax>
        <GenMinHarm>60</GenMinHarm>
        <GenMaxHarm>1000</GenMaxHarm>
        <NbPointSpectro>512</NbPointSpectro>
        <NbPointProcess>512</NbPointProcess>
        <LogNbPointProcess>9</LogNbPointProcess>
        <ProcessWindowType>Hanning</ProcessWindowType>
        <ProcessDeltaMax>40</ProcessDeltaMax>
        <ProcessEnhancement>0</ProcessEnhancement>
        <ProcessMultFactor>18</ProcessMultFactor>
        <ProcessZeroTreshold>-15</ProcessZeroTreshold>
        <SNRatio>12</SNRatio>
        <VoisSeuil>100</VoisSeuil>
        <IntensSeuil>18</IntensSeuil>
        <TauxHarm>33</TauxHarm>
        <Interpol>4</Interpol>
        <Group>2</Group>
        <Median>3</Median>
        <DynamicNbItem>16</DynamicNbItem>
        <DynamicNbValue>2</DynamicNbValue>
        <Turbulence>4</Turbulence>
        <Median2>3</Median2>
        <MinHarmon>50</MinHarmon>
        <MaxHarmon>1000</MaxHarmon>
        <MaxNbDent>6</MaxNbDent>
        <MaxNbPeak>25</MaxNbPeak>
        <DiffMinHarmon>50</DiffMinHarmon>
        <DiffMaxHarmon>4500</DiffMaxHarmon>
        <DiffMaxNbDent>1</DiffMaxNbDent>
        <DiffMaxNbPeak>25</DiffMaxNbPeak>
        <Diff2MinHarmon>1000</Diff2MinHarmon>
        <Diff2MaxHarmon>25</Diff2MaxHarmon>
        <Diff2MaxNbDent>0</Diff2MaxNbDent>
        <Diff2MaxNbPeak>25</Diff2MaxNbPeak>
        <AmdfLength>256</AmdfLength>
        <AmdfClipping>30</AmdfClipping>
        <AutoVoicingTreshold>40</AutoVoicingTreshold>
        <AutoSilenceTreshold>70</AutoSilenceTreshold>
        <AutoLength>40</AutoLength>
        <AutoSpan>50</AutoSpan>
        <AutoType>Yin</AutoType>
        <AutoWindowType>Rectangle</AutoWindowType>
        <FZeroPreset>PresetNull</FZeroPreset>
        <RecordIndex>0</RecordIndex>
        <NbImitation>0</NbImitation>
        <LTLStartSampleIndex>0</LTLStartSampleIndex>
        <LTLEndSampleIndex>0</LTLEndSampleIndex>
</TrackingSelection>
```

Export vers Praat, 1 à n fichiers





n fichiers dans le même répertoire (batch mode)

Génération d'un fichier .Pitch

```
File type = "ooTextFile"
Object class = "Pitch 1"

xmin = 0
xmax = 3.437
nx = 342
dx = 0.010
x1 = 0.2
ceiling = 800
maxnCandidates = 1
frame []:
    frame [1]:
        intensity = 0
        nCandidates = 1
        candidate []:
            candidate [1]:
                frequency = 0
                strength = 1
    frame [2]:
        intensity = 0
        nCandidates = 1
        candidate []:
            candidate [1]:
                frequency = 0
                strength = 1
    frame [3]:
        intensity = 0
        nCandidates = 1
        candidate []:
            candidate [1]:
                frequency = 0
                strength = 1
```

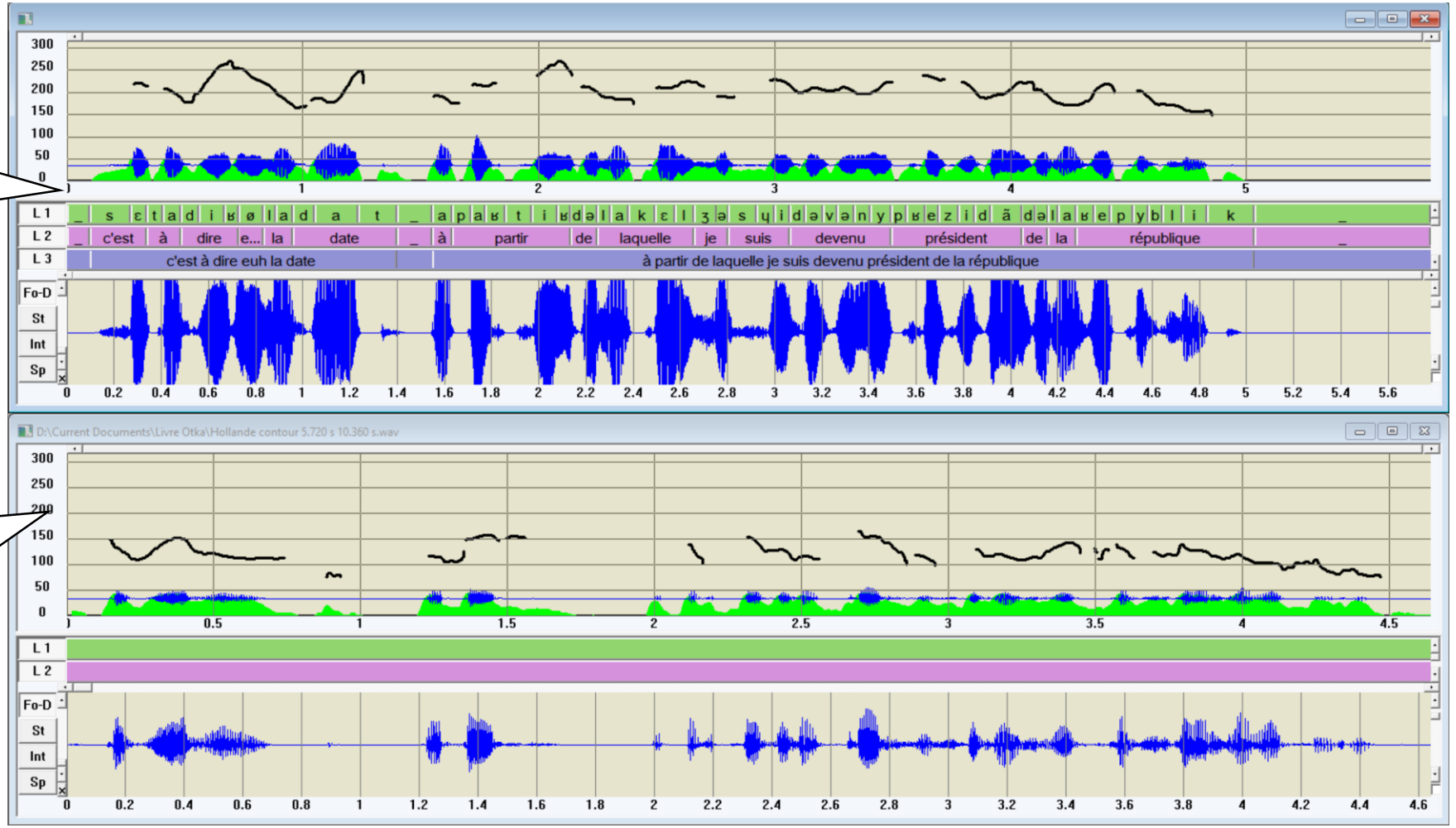
C’est fini ! ☹

Merci ! ☺

Object: map an existing speech segmentation on an unaligned speech segment

Display on the source sentence, with aligned text, API transcription, etc.

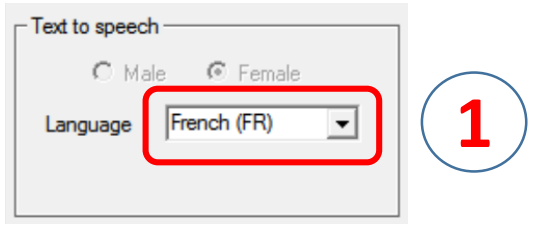
Result of the automatic aligment. The target sentence is now segmented according to the model

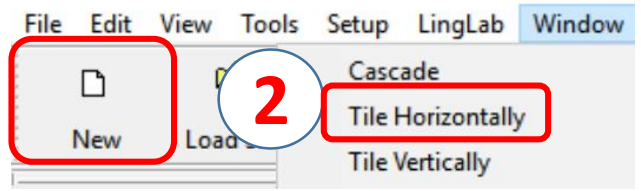


c'est à dire euh la date, à partir de laquelle je suis devenu président de la république.

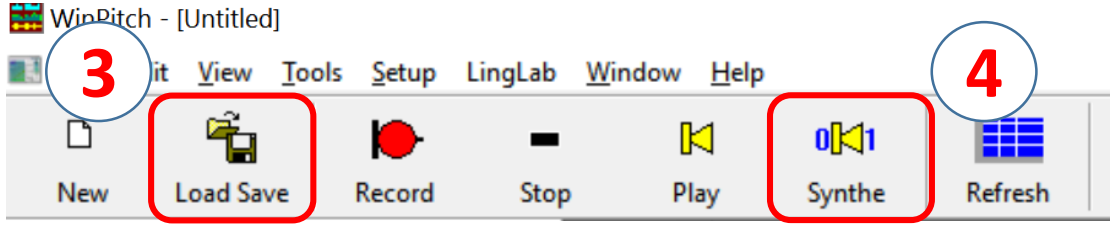
Sequence of operations

1. Load the source aligned sentence, possibly generated by the imbedded text-to-speech synthesizer. The orthographic input text must correspond exactly to the target text transcription. Make sure the synthesis language is correctly selected in the Setup / Screen command box





2. Create a new analysis window and display both windows horizontally



3. Load the target sound file in the empty lower window

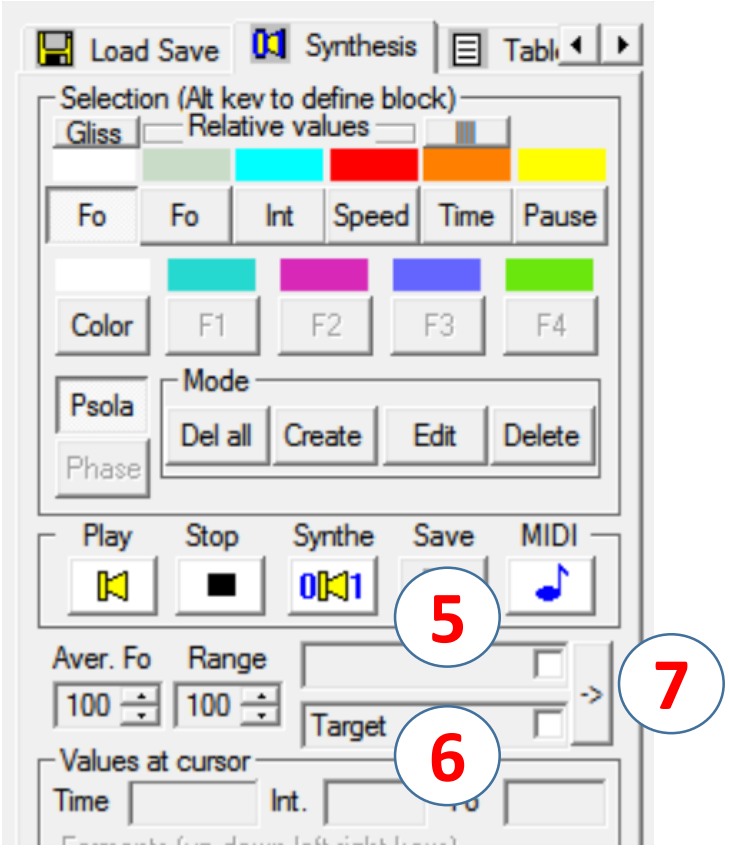
4. Select the Synthesis command box on the toolbar

Sequence of operations…

5. Select the source window (by clicking in it) and validate the source control box in the synthesis command box. The source is defined by whatever is displayed on the analysis screen, even if the complete file is much larger.

6. Select the target window (by clicking in it) and validate the target control box in the synthesis command box

7. Click on the arrow key on the right of the source and target selection boxes. This will initiate the automatic alignment process.



Sequence of operations…

8. All information entered in the source wp2 file will be mapped on the target file (transcription, screen text, highlight, etc.)