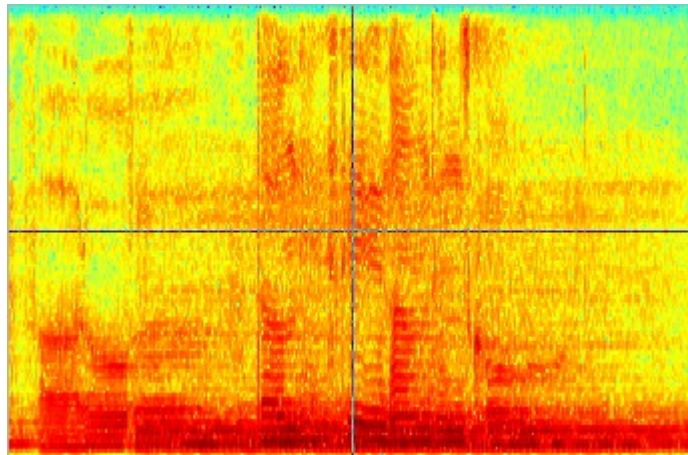


Le Vocodeur : analyse des résultats.

Avant toutes choses, notons que, pour des raisons pratiques, nous avons dû nous restreindre à des fichiers wav échantillonnés en 8000 Hz, à comparer aux 40 000 Hz usuels.

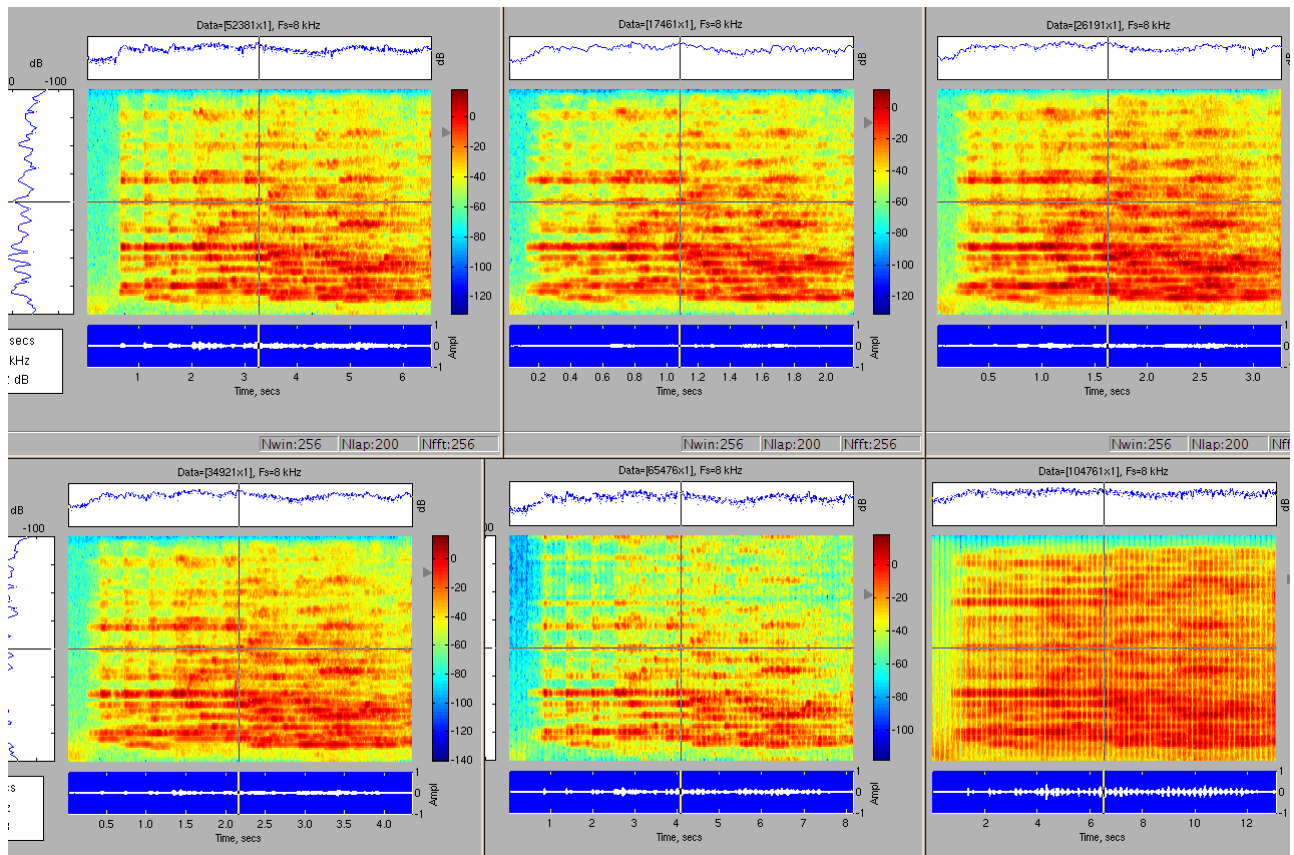
Cela donne un son renfermé.

Nous avons essayé de nombreuses variations de paramètres ; voici un petit récapitulatif de nos résultats expérimentaux. Voici le spectrogramme du signal qui a servi de base aux autres spectrogrammes du rapport : un homme qui chante, d'une voix grave, « Far away, the misty mountains cold ».

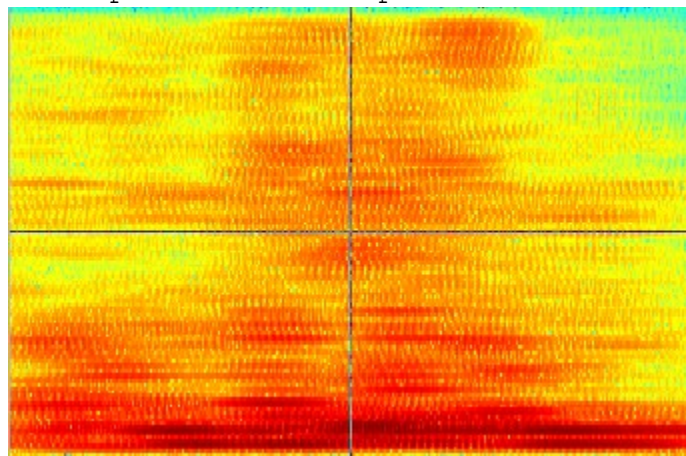


En ce qui concerne les fenêtres :

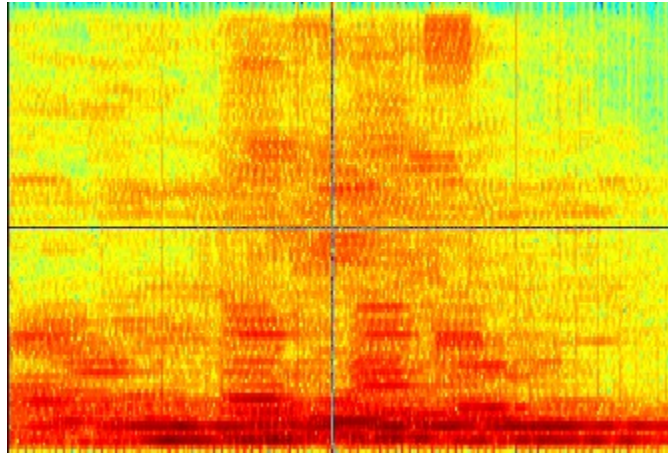
- La fenêtre sinus cardinal, qui a été notre fenêtre « type », donne des voix un peu métalliques, et induit un certain écho. Par contre, sur de la musique simple (extraits de flûte), on ne perçoit pas de modification. On notera toutefois que, sur un exemple de riff de guitare, on a une impression de perte de précision des attaques, et ce même en ralentissant (quand on accélère, on obtient une sorte de bouillie, mais cela paraît inévitable). Notons enfin que, du fait de la décroissance de sinus cardinal, augmenter la taille de la fenêtre a un effet bénéfique. Ci dessous, les spectrogrammes obtenus pour un morceau joué à la flûte traversière, avec en haut à gauche le signal original, puis un signal accéléré de moins en moins : $\times 3$, $\times 2$, $\times 1.5$, $\times 0.8$, $\times 0.5$. On voit alors sur les spectrogrammes que les attaques, qui étaient nettes (en haut à gauche, sur le signal original, avec des lignes verticales bien marquées), sont plus étalées sur le signal accéléré. À notre avis, ces tâches sont dûes à la largeur de la fenêtre.



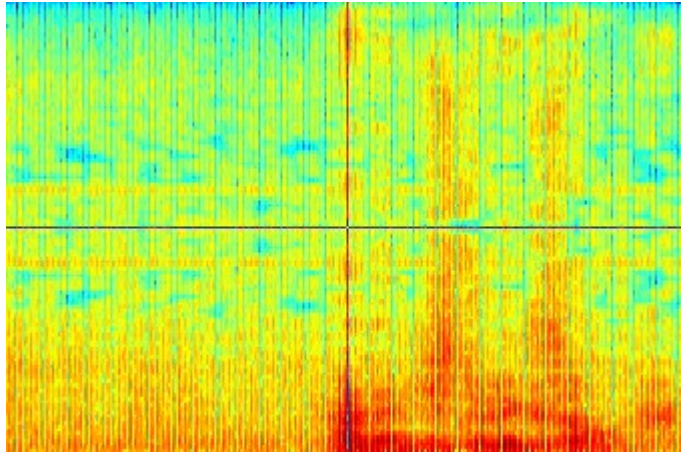
- Ce n'est pas le cas de la fenêtre de Hamming et de la fenêtre triangulaire : si la fenêtre est petite, la qualité est similaire à celle obtenue avec sinc, mais, si l'on élargit trop la fenêtre, on a un effet de moyennage temporel beaucoup trop important, et on perd toute compréhension.



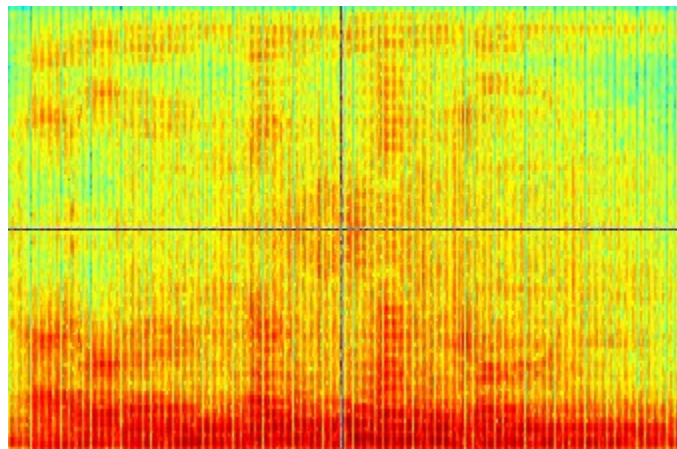
- Les fenêtres créneaux ont le même problème, mais on détecte en plus un artefact lors de l'accélération : des craquement apparaissent (ce sont les barres verticales rouges sur le spectrogramme).



En ce qui concerne le bitrate : plus il est élevé, mieux c'est. En effet, avec les fichiers de qualité médiocre, l'écho induit par le fenêtrage a une influence catastrophique sur l'intelligibilité de la parole. L'image ci-dessous montre que le phénomène d'ondulation, dont on parle plus bas, passe presque inaperçu lorsque le fichier est de bonne qualité : les oscillations sont trop rapides pour être perçues.



Le problème principal que nous avons rencontré est celui de l'ondulation lors de la décélération. Un simple spectrogramme (ralentissement x3) est éloquent :



Le son obtenu est difficilement intelligible, c'est un peu comme si le micro était écarté de la source sonore périodiquement, plusieurs fois par seconde. Nous attribuons ce problème à notre interpolation linéaire lors de la décélération du signal. Il aurait peut-être été plus judicieux de décélérer le signal avant de le traiter, de manière à ne pas avoir ce genre de problèmes.