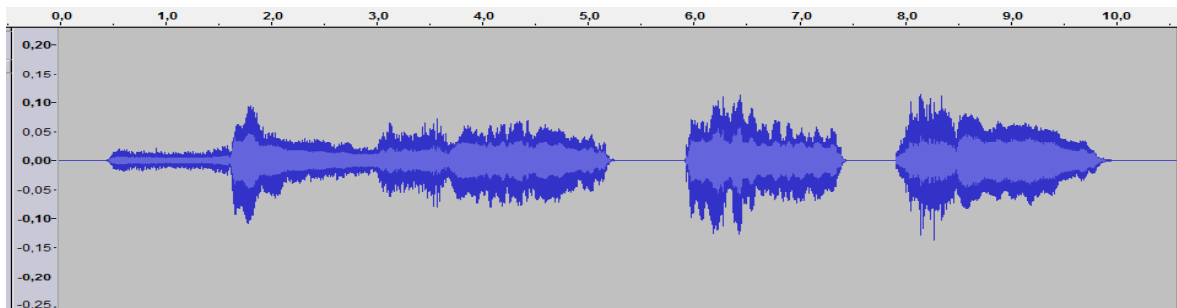


**La reconnaissance vocale est une technique permettant d'analyser une parole et de la retranscrire sous une forme exploitable par une machine.**

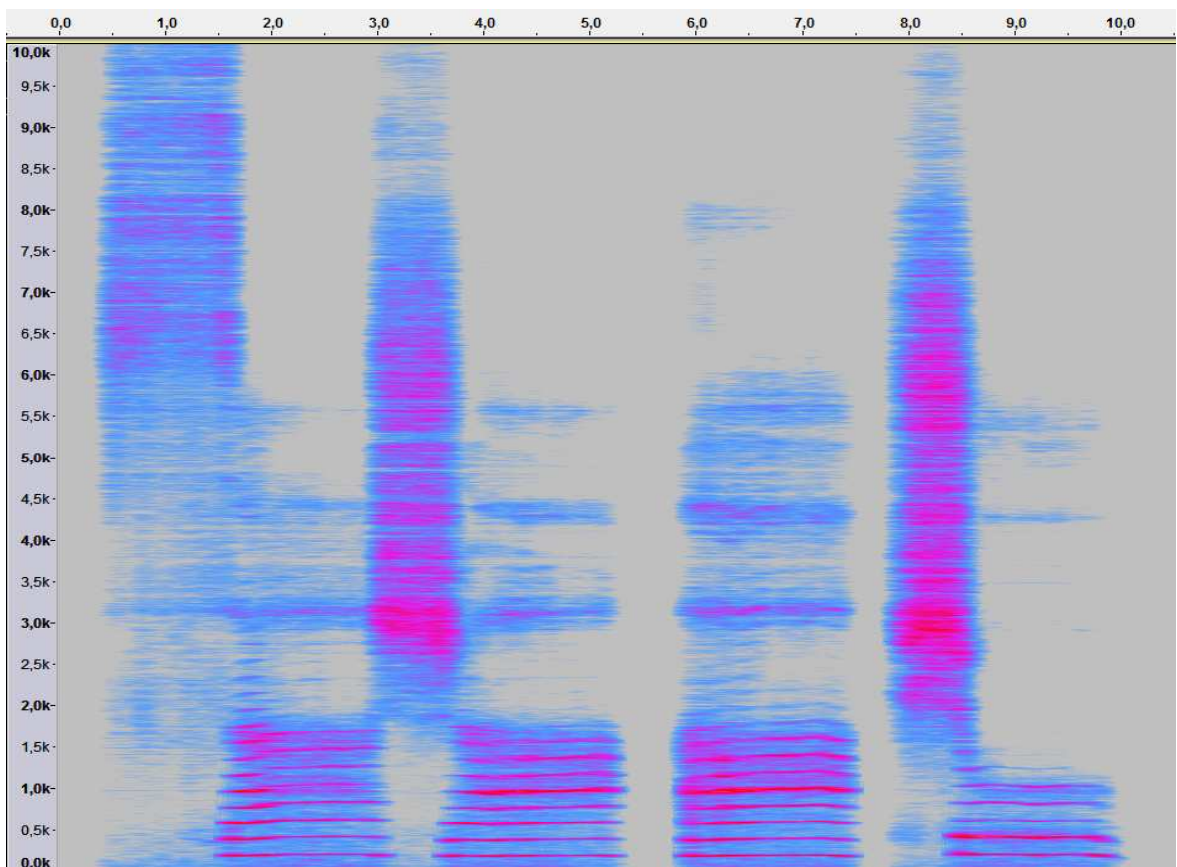
**Que signifie la phrase suivante ?**

*Allure temporelle de la phrase prononcée lentement et distinctement enregistrée à l'aide du logiciel Audacity:*



*Spectrogramme de la même phrase :*

(Spectrogramme : le temps est représenté en abscisse ( en s ) , la fréquence en ordonnée (en Hz), le niveau sonore est matérialisé par un niveau de couleurs)



## Questions :

1. Quelle(s) information(s) peut-on déduire de la représentation temporelle précédente?

Après avoir pris soigneusement connaissance des documents joints, répondre aux questions suivantes :

2. Dans le texte du document 1 que signifient les termes repérés en gras ?

3. Quelle(s) partie(s) de l'appareil phonatoire est (sont) responsable(s) des différences entre les différentes lettres prononcées ?

À l'aide d'audacity , enregistrer les signaux sonores en prononçant lentement : aaaa , oooo , chhhh, ssss , aaaa , oooo , chhhh , ssss . Visualiser le signal temporel.

4. À l'aide du signal temporel enregistré, vérifier la distinction entre les voyelles et les consonnes relevée dans le texte du document 1.

Expliquer cette distinction. On illustrera à l'aide de schémas.

5. Avec le logiciel d'acquisition visualiser le spectre en fréquence de chaque son prononcé.

Noter les fréquences des 4 premiers pics de chaque spectre pour les différents sons lorsque cela est possible.

Indiquer les différences entre les voyelles et les consonnes, les illustrer à l'aide de schémas.

6. Sur la gauche du signal temporel, cliquer sur « Piste audio » puis choisir « Spectrogramme ».

(S'assurer que la définition d'affichage est 16384 dans Edition - Préférences - Spectrogrammes)

Expliquer ce qui permet de distinguer les voyelles des consonnes sur un spectrogramme.

7. Sur le doc.4, on a prononcé les 4 sons dans le désordre. Localisez sur l'axe des temps

aaaa, oooo, chhhh , ssss.

Qu'est ce qui permet de distinguer le « a » et le « o », le « ch » et le « s » ?

8. Quelle est la phrase prononcée ? Repérer les différents sons sur l'axe des temps.

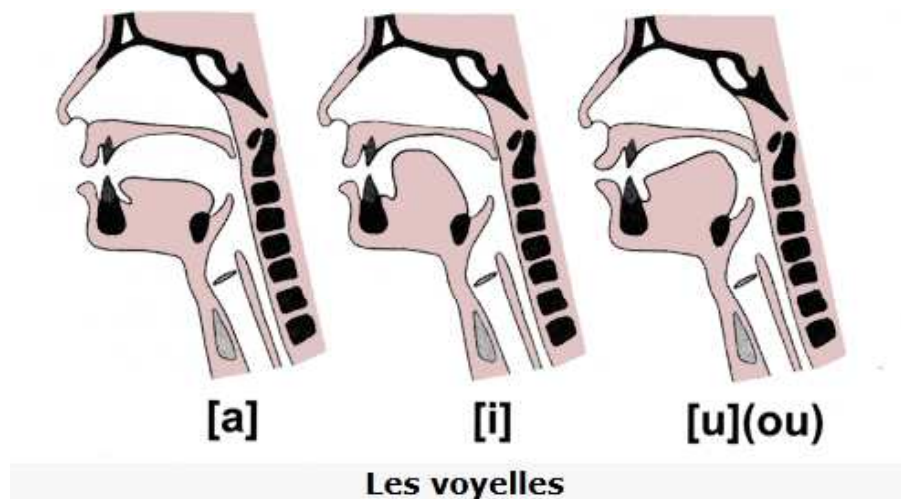
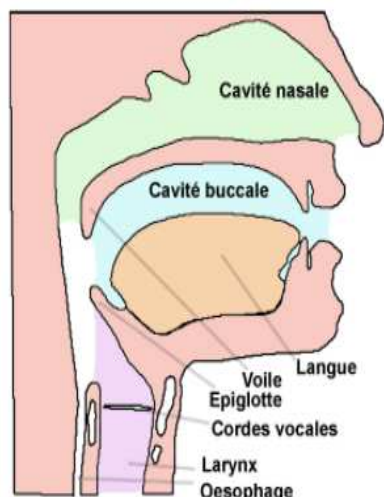
9. Quelles peuvent être les difficultés rencontrées par un système de reconnaissance vocale ?

## DOCUMENTS A DISPOSITION

### Document 1 : Les sons de la parole

« La distinction des sons de la parole en deux catégories bien distinctes, les voyelles et les consonnes, est principalement liée à leur caractère périodique ou apériodique. Les sons **vocaliques** sont périodiques et classés en termes de répartition d'énergies dans le spectre, tandis que les sons **consonantiques** sont assimilables à des bruits, avec une structure moins ordonnée et apériodique. »

## Document 2 : Les résonateurs



Le spectre du son produit par nos cordes vocales est modifié par les résonances des différentes harmoniques dans les cavités situées entre les cordes vocales et l'extérieur: le pharynx, la bouche et les cavités nasales. En fonction de la forme de ces cavités, certaines fréquences seront atténuées alors que d'autres resteront amplifiées et seront plus audibles.

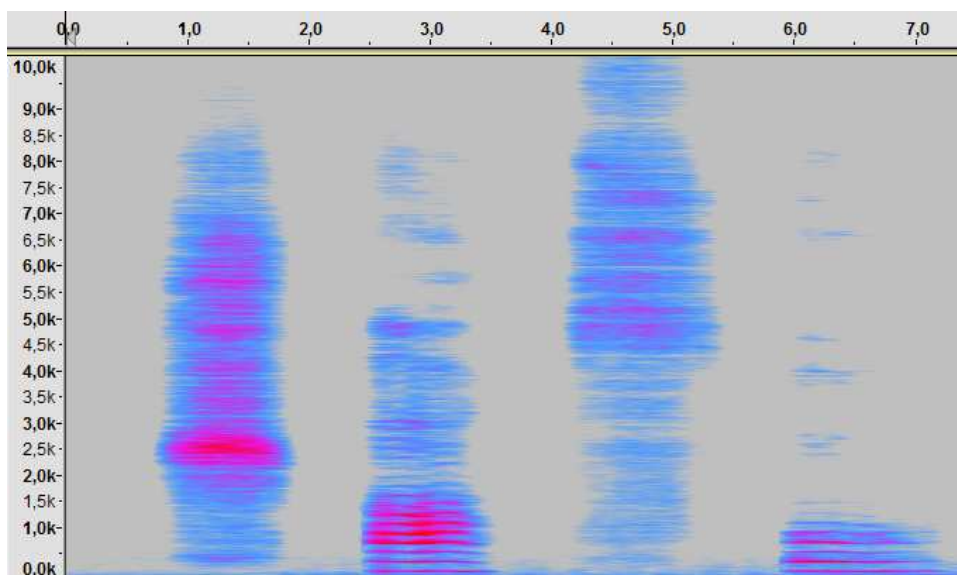
<http://www.chaumetsoftware.com/timbre.html>

## Document 3 : Sons , sons musicaux et bruits

Contrairement au spectre d'un son périodique, qui est *discret* c'est-à-dire composé d'une série définie de fréquences (*fondamental et harmoniques* :  $f$ ,  $2f$ ,  $3f$ ,  $4f$ , etc.), le spectre d'un bruit est *continu*. Autrement dit, il comporte théoriquement toutes les fréquences (avec des amplitudes plus ou moins fortes, voire nulles). Les chocs brefs sont un exemple de bruit. (Cependant, un son peut être non périodique, sans pour autant être un bruit. Dans ce cas, le spectre est composé de fréquences qui ne sont pas multiples d'une même valeur.)

*Initiation à l'acoustique – A. Fischetti*

## Document 4 :



## Document 5 : Techniques de la reconnaissance vocale

Pour identifier le langage parlé, deux techniques peuvent être employées en fonction de l'application utilisée.

- L'approche globale s'intéresse directement aux mots. Une image acoustique d'un certain nombre de mots est enregistrée dans une base et comparée à l'image acoustique du mot prononcé. Cette technique est efficace, mais le nombre de mots est limité, à cause de la capacité de stockage et de traitement de la machine.
- L'approche analytique est basée sur la reconnaissance des phonèmes<sup>(1)</sup> et syllabes. L'avantage est incomparable : tous les mots de la langue française peuvent être prononcés avec seulement 36 phonèmes. Pour pallier les différences d'accent, de timbre, ou de rythme de parole, on « entraîne » le système de reconnaissance en faisant lire au locuteur des phrases choisies. La plus grande difficulté est que la machine doit interpréter le sens du discours pour ne pas confondre deux phrases phonétiquement identiques (comme par exemple : « la citerne est pleine d'eau » et « la scie terre est plaine dos »).

*Bordas - Physique Chimie Spécialité - Collection Espace- Edition 2012*

<sup>(1)</sup> Un **phonème** est la plus petite unité discrète ou distinctive (c'est-à-dire permettant de distinguer des mots les uns des autres) que l'on puisse isoler par segmentation dans la chaîne parlée.

# Éléments de correction

1. Quelle(s) information(s) peut-on déduire de la représentation temporelle précédente?

D'après la représentation temporelle, la phrase est constituée de 3 mots. On observe en effet 3 parties séparées par des silences sur cette représentation.

2. Dans le texte du document 1 que signifient les termes repérés en gras ?

Un son vocalique correspond à une voyelle. Un son consonantique correspond à une consonne.

3. Quelle(s) partie(s) de l'appareil phonatoire est (sont) responsable(s) des différences entre les différentes lettres prononcées ?

D'après le document 2, le son émis par les cordes vocales est modifié par la forme des cavités buccale, nasale et pharyngée. Sur le schéma, on observe que la position de la langue et des lèvres modifie la forme de la cavité buccale, amplifiant et atténuant certaines fréquences du spectre du son produit ce qui permet d'émettre des voyelles différentes.

4. À l'aide du signal temporel enregistré vérifier la distinction entre les voyelles et les consonnes relevée dans le texte du document 1.

Expliquer cette distinction. On illustrera à l'aide de schémas.

D'après le document 1, les voyelles correspondent à des signaux périodiques et les consonnes à des signaux apériodiques.

Représentations temporelles :  
d'une voyelle



d'une consonne



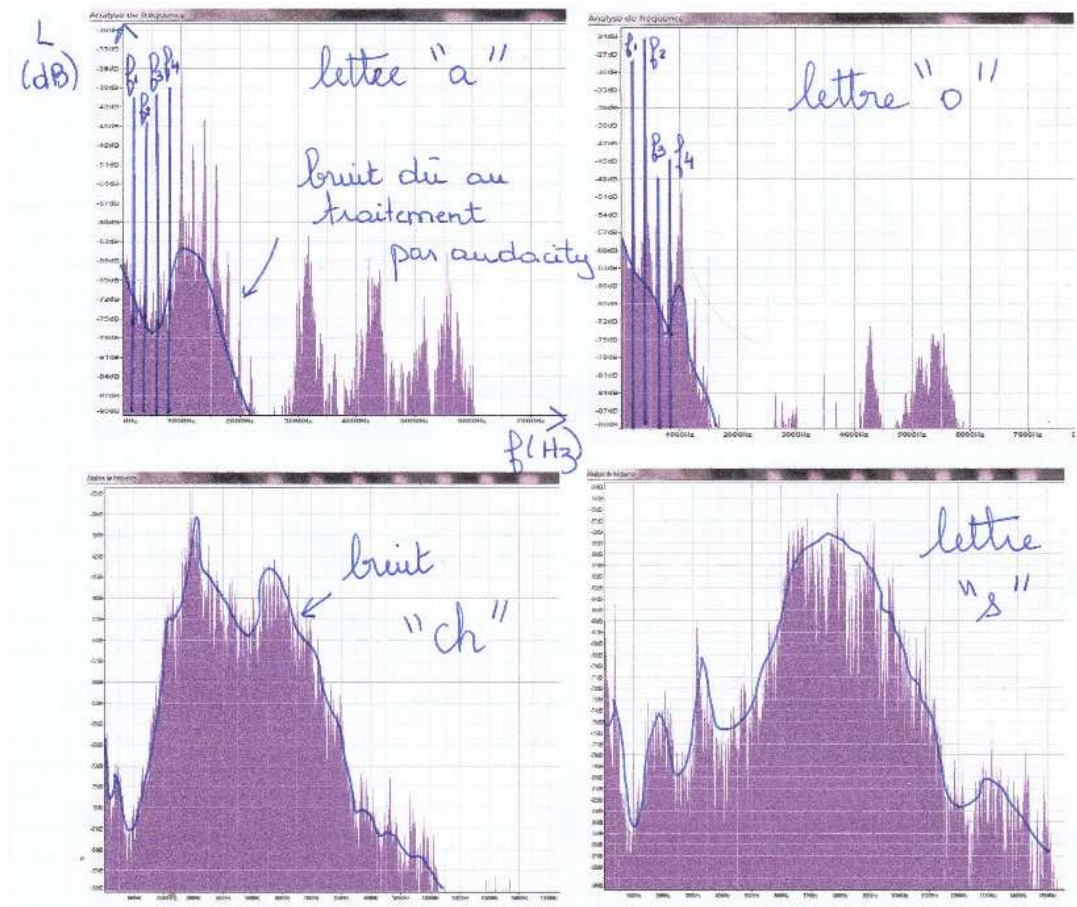
5. Avec le logiciel d'acquisition visualiser le spectre en fréquence de chaque son prononcé. Noter les fréquences des 4 premiers pics de chaque spectre pour les différents sons lorsque cela est possible. Indiquer les différences entre les voyelles et les consonnes. Illustrer à l'aide de schémas.

L'analyse spectrale d'une voyelle se distingue de celle d'une consonne par l'allure de la courbe :

- Le spectre d'une voyelle est composé de pics correspondant à une suite discrète de fréquences de valeurs  $f_1$ ,  $2 f_1$ ,  $3 f_1$ ,  $4 f_1$ , ... ce qui correspond à un son musical (périodique).

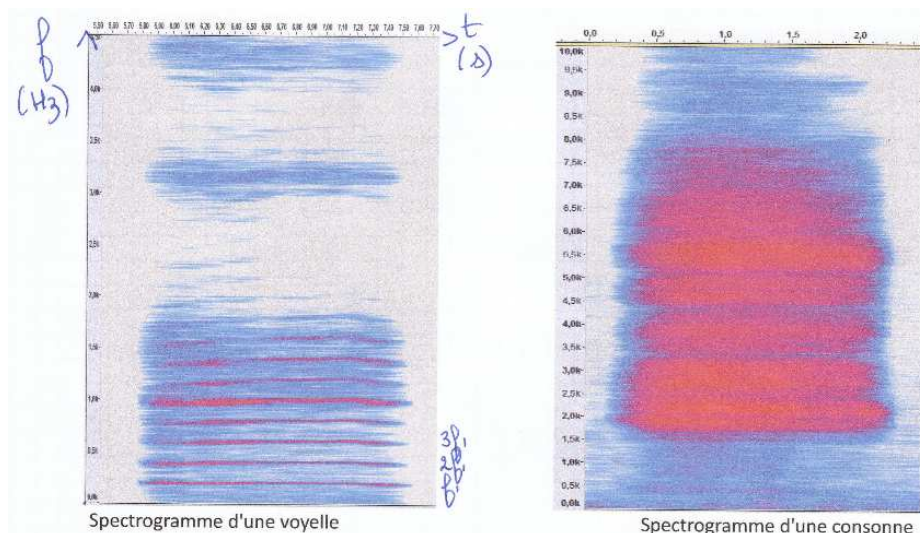


- Le spectre d'une consonne est continu, il comporte toutes les fréquences avec des amplitudes plus ou moins fortes, ce qui correspond à un bruit (apériodique). Aucune fréquence ne se distingue clairement des autres.

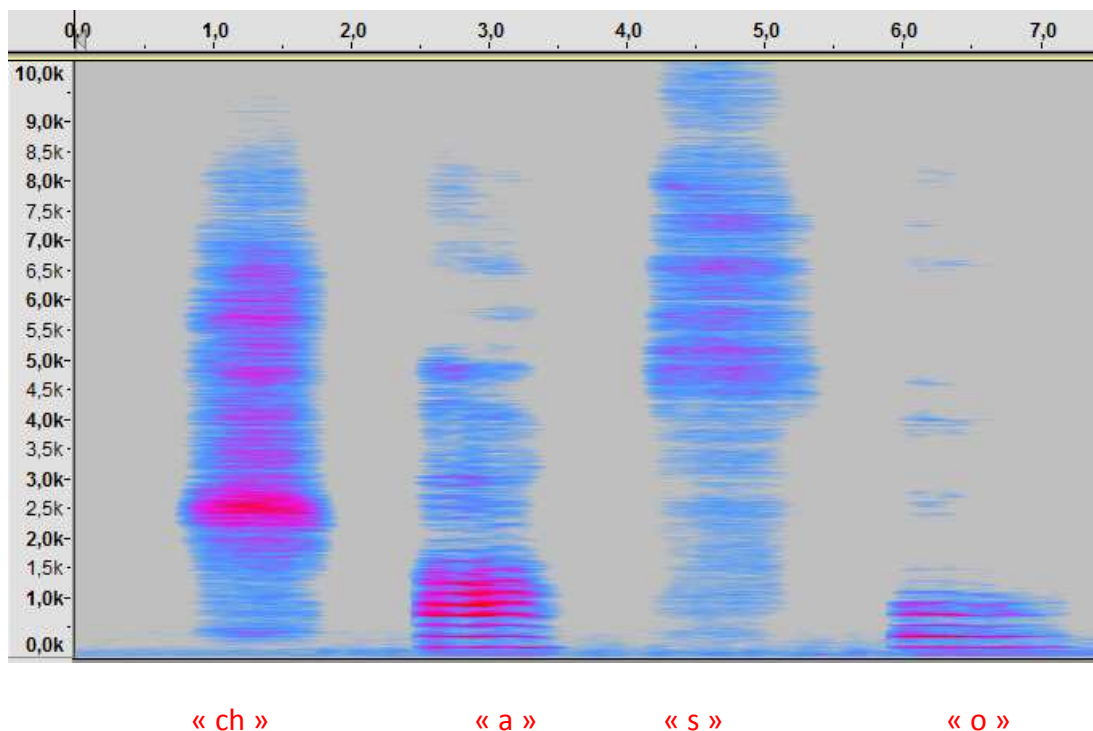


- Sur la gauche du signal temporel, cliquer sur « Piste audio » puis choisir « Spectrogramme ».  
Expliquer ce qui permet de distinguer les voyelles des consonnes sur un spectrogramme.

Le spectrogramme d'une voyelle fait apparaître des lignes horizontales régulièrement espacées correspondant au fondamental et à ses harmoniques ( $f_1$ ,  $2 f_1$ ,  $3 f_1$ ,  $4 f_1$ , ...)  
le spectrogramme d'une consonne ne met pas en évidence des fréquences privilégiées : toutes les fréquences apparaissent avec plus ou moins d'intensité.



7. Sur le doc.4, on a prononcé les 4 sons dans le désordre. Localisez sur l'axe des temps aaaa, oooo, chhhh, ssss.  
Qu'est ce qui permet de distinguer le « a » et le « o », le « ch » et le « s » ?

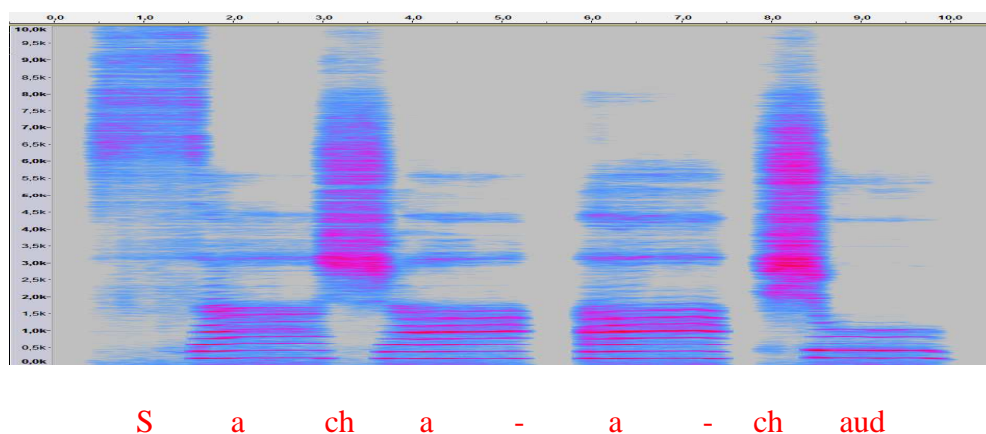


Le son « a » est caractérisé par la présence d'harmoniques de rangs faibles et moyens ce qui le différencie du son « o » composé d'harmoniques de rangs faibles uniquement.

Le son « s » se distingue du son « ch » car son spectrogramme est plus intense pour les fréquences élevées et peu intense pour les fréquences faibles et moyennes.

Le son « ch » est composé principalement de fréquences moyennes.

8. Quelle est la phrase prononcée ? Repérer les différents sons sur l'axe des temps.



9. Quelles peuvent être les difficultés rencontrées par un système de reconnaissance vocale ?

Cette question ouverte a pour but de montrer la complexité de la technique de reconnaissance vocale.