



1. Les trois parties du corps qui interviennent dans la voix sont les poumons, les cordes vocales et les résonateurs (gorge, nez, bouche).
2. D'après le document 3, 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.

3. Reconnaître un mot sur un spectrogramme

- a. Fréquences des formants (intensités fortes, en gris/noir sur le spectrogramme) :

t (s)	0,3	0,5	0,7
fréquences des formants (Hz)	600 / 1500 / 2700	500 / 1500 / 2500	300 / 2500 / 3200

- b. En utilisant le tableau du document 3, on identifie les phonèmes :
0,3s \Rightarrow a 0,5s \Rightarrow eu 0,7s \Rightarrow i
- c. • Le premier son contient beaucoup de hautes fréquences et n'a aucune structure périodique : consonne fricative comme le "s".
• Le son perçu à la date 0,4s fait apparaître des formants ($F_1 = 300\text{Hz}$ et $F_2 = 1500\text{Hz}$). Ce type de son correspond à consonne vocalique telle que le "m".
• À la date $t = 0,6\text{s}$, un silence est observé : consonne occlusive comme le "d".
 \Rightarrow Le mot prononcé est donc "samedi".

4. Déterminer le sens d'une phrase sur un spectrogramme

- a. D'après le document 3,
 \Rightarrow les voyelles correspondent à des signaux périodiques : le spectre d'une voyelle est composé de pics correspondant à une suite discrète de fréquences $f_1, 2.f_1, 3.f_1, 4.f_1$ etc.
 \Rightarrow et les consonnes à des signaux apériodiques : 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.
- b. Le spectrogramme d'une voyelle fait apparaître des lignes horizontales régulièrement espacées correspondant au fondamental et à ses harmoniques ($f_1, 2.f_1, 3.f_1, 4.f_1$ etc.)
Le spectrogramme d'une consonne ne met pas en évidence des fréquences privilégiées : toutes les fréquences apparaissent avec plus ou moins d'intensité.
- c. \Rightarrow Distinction du "a" et du "o" :
- le son "a" est caractérisé par des formants vers 800Hz, 1200Hz et 2400Hz.
- le son "o" est caractérisé par des formants vers 400Hz, 900Hz et 2300Hz.
 \Rightarrow Le son "ch" se distingue du son "ss" car son spectrogramme est plus intense pour les fréquences élevées et peu intense pour les fréquences faibles et moyennes.

VOYELLES

signal périodique, spectre discret avec un fondamental de fréquence f_1 et des harmoniques de fréquences f_2, f_3 , etc qui se traduisent par des lignes horizontales sur le spectrogramme \Rightarrow

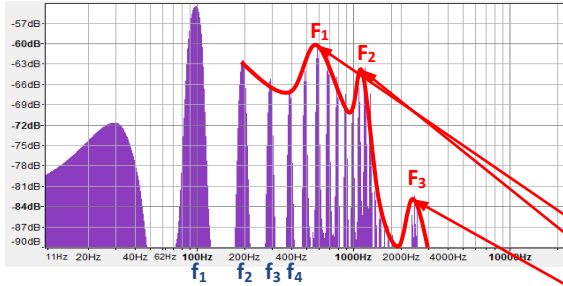
formants :

\Leftarrow maxima de l'enveloppe (entre 300Hz et 3kHz)
 formant 1 : F_1
 formant 2 : F_2
 formant 3 : F_3

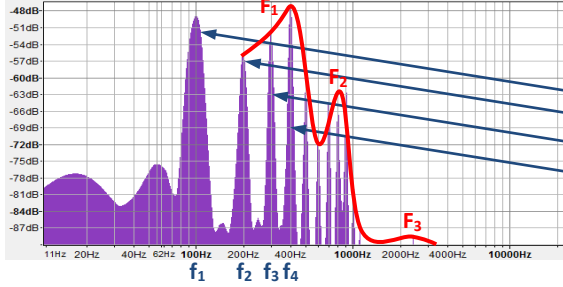
harmoniques :

$n^{\circ}1$: fondamental f_1
 $n^{\circ}2$: $f_2 = 2.f_1$
 $n^{\circ}3$: $f_3 = 3.f_1$
 $n^{\circ}4$: $f_4 = 4.f_1$

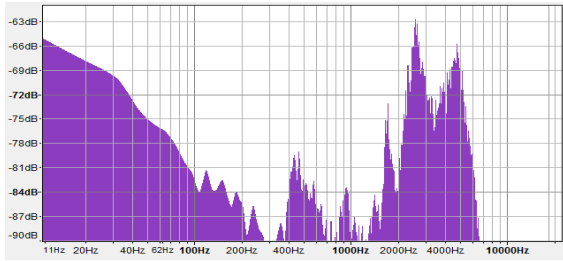
son "a" : spectre en fréquence (échelle logarithmique)



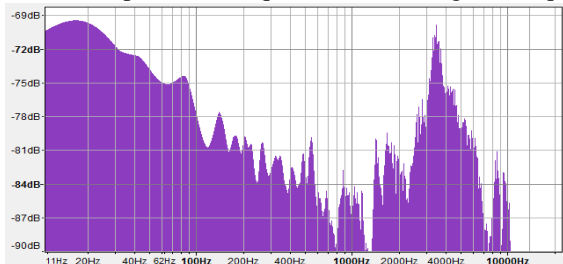
son "o" : spectre en fréquence (échelle logarithmique)



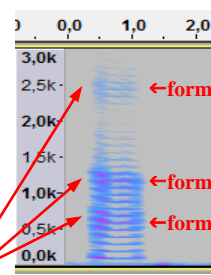
son "ch" : spectre en fréquence (échelle logarithmique)



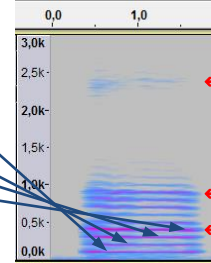
son "ss" : spectre en fréquence (échelle logarithmique)



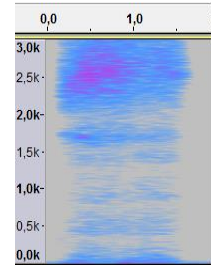
son "a" : spectrogramme



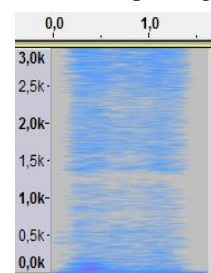
son "o" : spectrogramme



son "ch" : spectrogramme



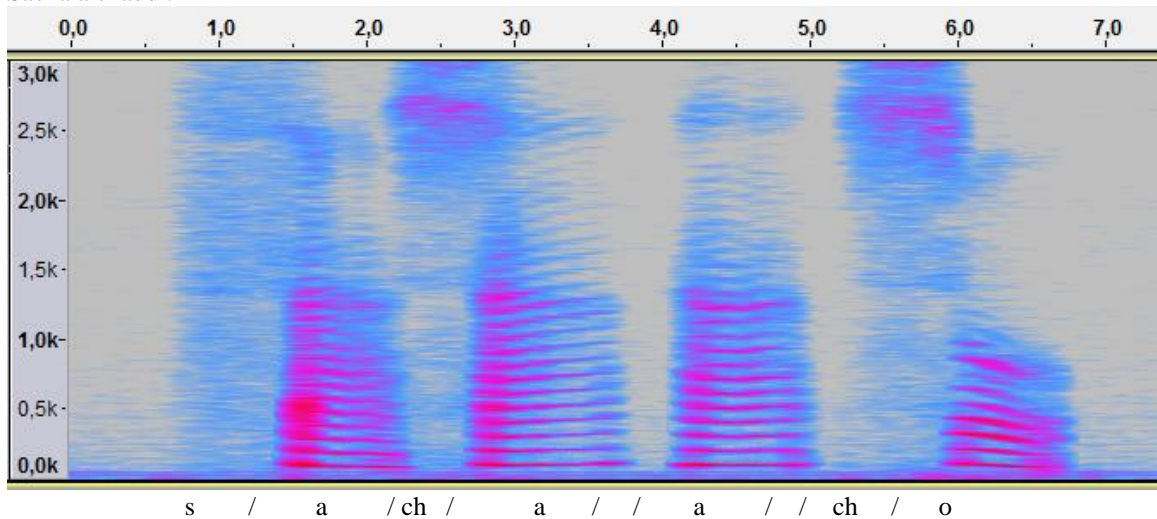
son "ss" : spectrogramme



CONSONNES

signal de type bruit, toutes les fréquences sont présentes avec des intensités variables

d. Sacha a chaud :



e. Les spectrogrammes d'une même phrase peuvent être sensiblement différents d'une personne à l'autre.

5. Conclusion :

La reconnaissance vocale est un procédé permettant d'analyser la parole et de la transformer en mots et en phrases interprétables par un ordinateur. Ce passage est cependant d'une grande complexité.

Après enregistrement par un microphone, le son est analysé : le spectrogramme est découpé en phonèmes qui sont alors comparés à ceux inclus dans une base de données. L'identification des formants permet par exemple d'identifier les voyelles.

Les spectrogrammes d'une même phrase peuvent être sensiblement différents selon la personne qui la prononce : les formants n'ont pas exactement la même valeur et il peut parfois y avoir des ambiguïtés entre plusieurs phonèmes et donc entre plusieurs mots. Parfois des mots, voire des groupes de mots peuvent se prononcer de la même façon. Ce sont des homophones : "la citerne est pleine d'eau" ou "la scie terne est plaine d'os" ?

Les accents régionaux, les liaisons entre les mots peuvent aussi perturber la reconnaissance. Des analyses contextuelles et sémantiques permettent alors de trancher. Ces logiciels nécessitent une puissance de calcul importante dont on dispose actuellement grâce aux progrès technologiques. Souvent, le système de reconnaissance vocale a besoin d'une phase d'apprentissage pour bien identifier les caractéristiques de la voix et des phonèmes du locuteur.

La reconnaissance vocale est nécessaire à l'homme pour communiquer simplement avec les machines qu'il utilise (GPS, smartphones, plateformes téléphoniques). Les récents progrès de la reconnaissance vocale et l'augmentation de la puissance de calcul des ordinateurs a permis l'apparition de ces techniques dans de nombreux objets du quotidien.