 **REPUBLIQUE DU BENIN** 

**\*\*\*\*\*\*\*\***

**UNIVERSITE D’ABOMEY-CALAVI**

**\*\*\*\*\*\*\*\*\*\*\*\*\***

**ECOLE POLYTECHNIQUE D’ABOMEY-CALAVI (EPAC)**

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

**DEPARTEMENT DE GENIE INFORMATIQUE ET**

**TELECOMMUNICATION**

**(GIT)**

**\*\*\*\*\*\*\*\*\*\*\*\*\***

**Option : Réseaux et Télécommunications (RT)**

**UE : Traitement de la parole**

**Production et perception de la parole**

**Réaliser par :**

**Réalisé par :**

**AMOUSSOU Carlos**

**DJOKI don de dieu**

**SOULEMANE Djafarou**

**Sous la supervision de:**

**Dr DJARA Tahirou**

**Année Académique : 2020-2021**

TABLE DES MATIERES

LISTE DES FIGURES

INTRODUCTION

I. Aperçu anatomique.

II. Mécanisme de la phonation.

III. Phonétique articulatoire.

IV. Acoustique de la phonation.

V. Modélisation de la production de la parole

VI. Mécanisme de l'audition.

VII. Psycho acoustique.

VIII. Masquage et bandes critiques.

CONCLUSION

LISTE DES FIGURES

***Figure1 :*** Anatomie de l’appareil vocal humain

***Figure2 :*** Vue générale de l’appareil vocal

***Figure 3 :*** Coupe latérale du larynx

***Figure 4 :*** Voyelle antérieure Vs postérieure

***Figure5 :*** Fermée Vs Ouverte

***Figure 6 :*** Arrondie Vs étirée (Non arrondie)

***Figure 7 :*** Classement des consonnes

***Figure 8 :*** Consonnes fricatives dorsales et latérales

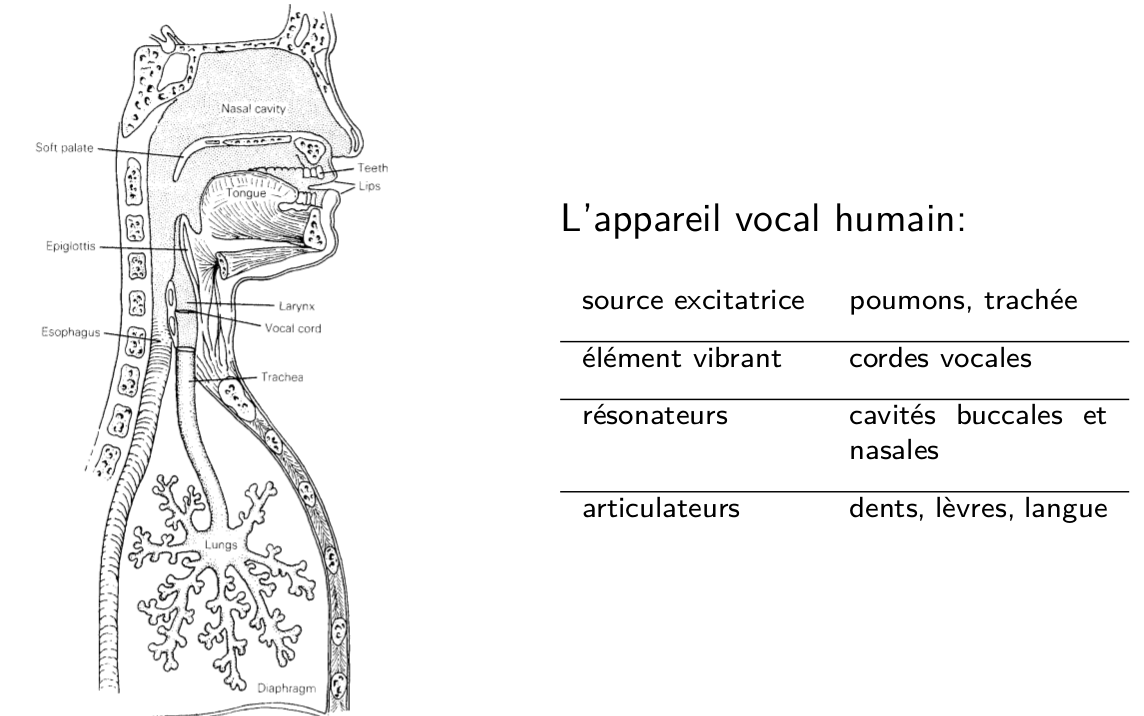
***Figure 9 :*** Sifflante (fricative) alvéolaire sourde/ sonore [s]/[z]

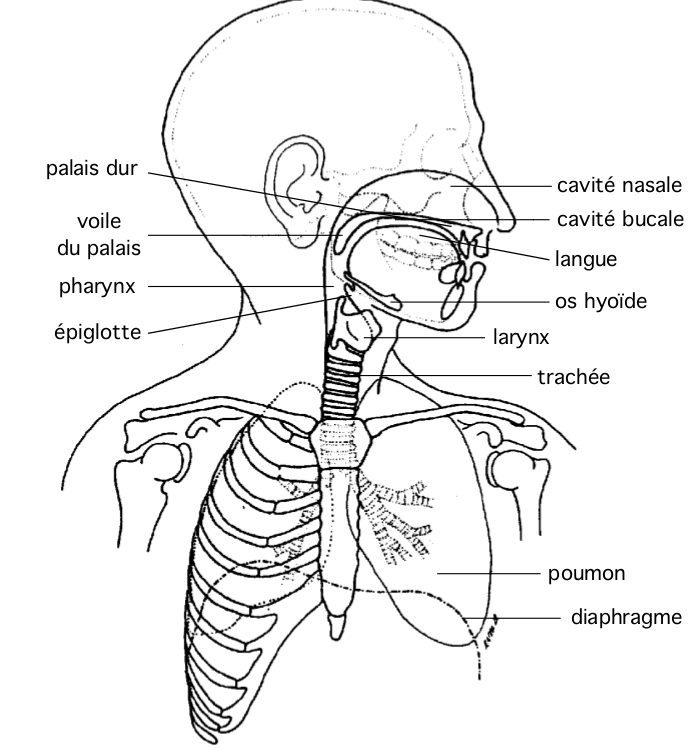
***Figure 10 :*** Position de la langue pour les fricatives et les spirantes latérales

***Figure 11 :*** Anatomie de l’oreille

***Figure 12 :*** l’enveloppe du son

**I. Aperçu anatomique**

***Figure1 :*** Anatomie de l’appareil vocal humain



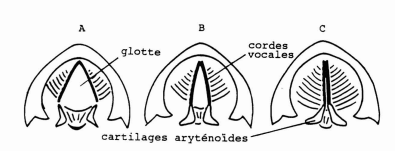
***Figure2 :*** Vue générale de l’appareil vocal

**II. Mécanisme de la phonation**

Le mécanisme phonatoire se compose en gros de trois composantes : l'appareil respiratoire, les cordes vocales, la cavité bucco-pharyngale. L'appareil respiratoire fonctionne comme un soufflet en fournissant l’énergie de départ sous forme d 'un souffle d 'air. Les cordes vocales agissent comme un générateur de sons ; pour la production des consonnes, la langue, les lèvres et les dents interviennent aussi et transforment ce souffle d’air en énergie sonore. Enfin, la cavité bucco-pharyngale, constituée du pharynx et de la bouche, a comme fonction d'amplifier cette énergie

Sonore.

Dans la phonation, tout débute par l'action des poumons. Ceux-ci libèrent, a un rythme qui est sous le contrôle volontaire du locuteur, un souffle d 'air, lequel passe par la trachée et traverse le larynx. Le larynx transforme alors le souffle en son glottique (ou laryngé), C'est, en effet, a l’intérieur du larynx, organe constitue de cartilages réunis entre eux par des ligaments et des muscles, que logent les cordes vocales. Celles-ci sont en fait des fibres musculaires contrôlées par des muscles qui ont comme fonction de les tendre, de les dilater, de les rétrécir ou encore de les allonger. Selon le type d’action des muscles et des cartilages sur les cordes vocales, l'espace entre elles peut être plus ou moins large, ou complètement fermé. C'est dans cet espace que passent les souffles d'air produits par I 'expiration. On appelle glotte cette zone entre les cordes vocales. Lorsque la glotte est largement ouverte, elle permet la respiration et aucun son n'est engendre ; lorsqu' elle n’est que faiblement ouverte, elle produit le chuchotement ; et quand elle est complètement fermée, il y a phonation (Figure 2).



***Figure 3 :*** Coupe latérale du larynx, représentant trois positions différentes des cordes vocales : a) glotte ouverte pour la respiration ; b) glotte entrouverte pour la voix chuchote ; c) glotte fermée par l'accolement des cordes vocales pour la phonation.

**III. Phonétique articulatoire**.

La distinction la plus pertinente entre voyelles et consonnes s’opère en tenant compte de la présence ou l’absence d’obstacle lors de la réalisation. On distinguera, alors, les deux cas suivants :

* Si le passage de l'air se fait librement à partir de la glotte, on a affaire à une voyelle ;
* Si le passage de l'air à partir de la glotte est obstrué, complètement ou partiellement, en un ou plusieurs endroits, on a affaire à une consonne.

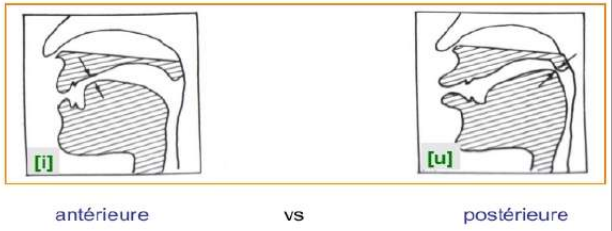
Avant d'aller plus loin, on signalera que le passage des consonnes aux voyelles ne se fait pas de manière abrupte, mais sur un continuum. On distinguera ainsi des articulations intermédiaires, comme par exemple les semi-voyelles ou les spirantes.

**III-1. Les voyelles**

Les voyelles résultent des simples vibrations des cordes vocales, quand l'air s'échappe sans n’être freiné ni arrêté. La réalisation de chacune dépend principalement de la position de la langue ou de l'ouverture de la bouche, de toute la modulation que nous pouvons opérer avec les organes phonateurs (arrondissement des lèvres, etc.). Les principes généraux de classement des voyelles se basent sur les caractéristiques articulatoires suivantes :

**III-1.1 Antérieure Vs Postérieure**

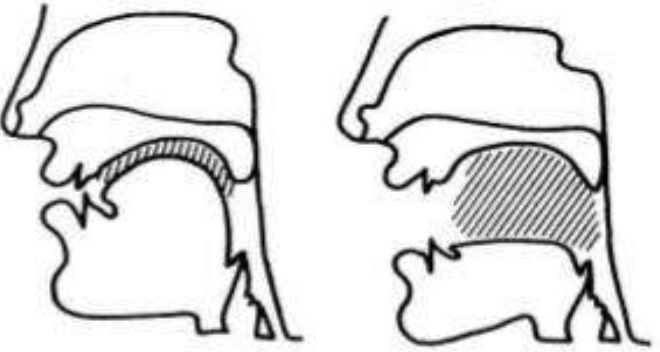
Lors de la réalisation d’une voyelle, le point d'émission ou la zone de vibration dans la cavité buccale se situe soit vers l'avant soit vers l'arrière : c'est vers cette zone que la langue se rapproche plus ou moins en direction du palais. Ainsi, les voyelles correspondant à "i" ou "é" se situent vers l'avant, elles sont antérieures. Il en est de même du "a" le plus courant. Les voyelles correspondant à "o", "ou", "on" se situent vers l'arrière, car la vibration se situe quasiment dans la gorge, elles sont donc postérieures. De même, le "â", avec accent circonflexe, comme dans âme. La voyelle correspondant à "eu" (ex : eux) est généralement décrite comme antérieure, mais elle est en fait plutôt centrale tout comme "e".

***Figure 4 :*** Voyelle antérieure Vs postérieure

**III-1.2 Ouvertes Vs fermées**

Lors de la réalisation des voyelles, les mâchoires sont plus ou moins écartées, la masse de la langue plus ou moins éloignée du palais : c'est ce qu'on appelle le degré d'aperture. On trouve ainsi des voyelles fermées / ouvertes : le /a/ est la voyelle la plus ouverte du français (surtout le "â"), le /i / est la voyelle la plus fermée. Il suffit tout simplement de prononcer successivement i / é / è / a (mis

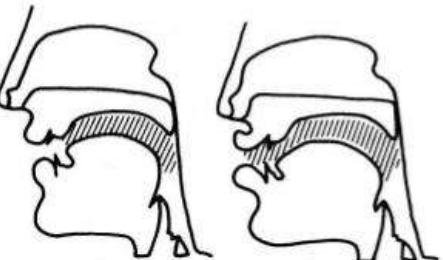
/ mes / mère / ma) pour tester les degrés d'aperture, on se rend compte tout de suite qu'on ouvre la bouche d’une manière graduelle. A noter aussi que la position qu’occupe la voyelle dans la syllabe (orale) détermine souvent son degré d'aperture.



***Figure5 :*** Fermée Vs Ouverte

**III-1.3 Arrondies Vs non arrondies**

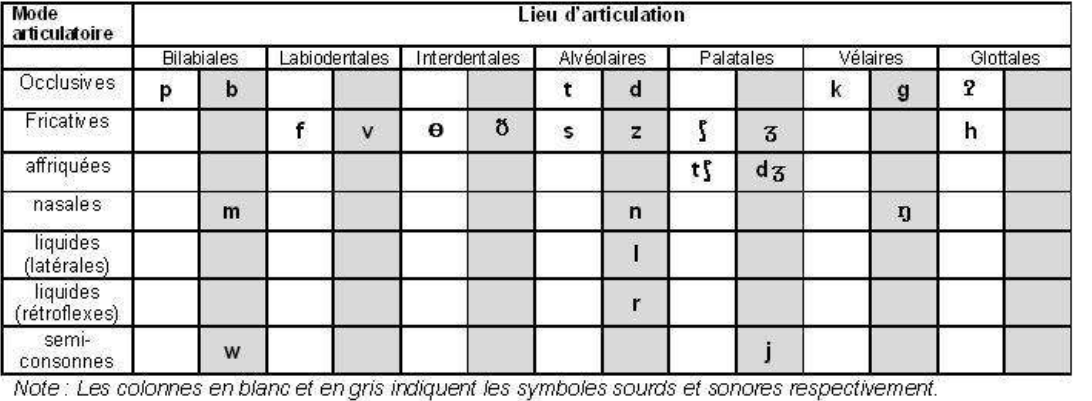
Pour la réalisation de certaines voyelles, l’on remarque que les lèvres s'arrondissent et ce projettes quelque peu en avant, comme dans le cas de la réalisation de certains phonèmes : Les voyelles o / ou / eu/ on, sont arrondies (ou labiales), on constate l'arrondissement et la contraction des lèvres en prononçant ces voyelles.



***Figure 6 :*** Arrondie Vs étirée (Non arrondie)

**III-2 Les consonnes**

En ce qui est des consonnes, selon l’articulation on distingue plusieurs grandes catégories : les occlusives, les fricatives, les spirantes, les latérales.

***Figure 7 :*** Classement des consonnes

**III-2-1 Les Occlusives**

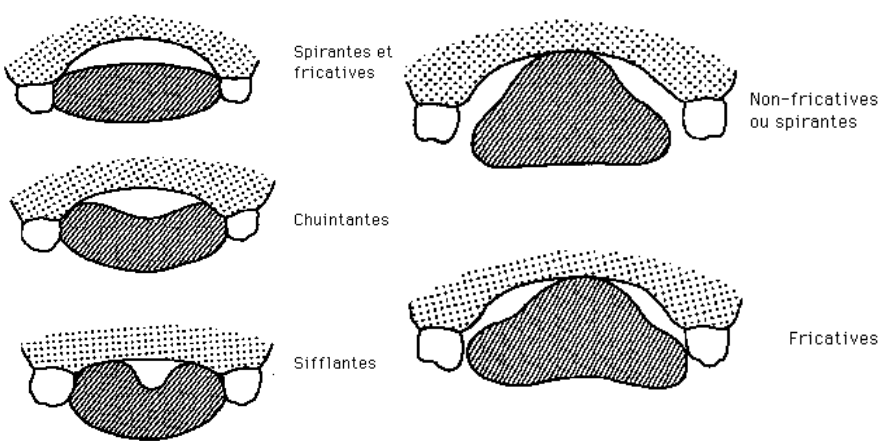
Les consonnes occlusives sont produites par une fermeture complète du canal vocal, et non un simple rétrécissement, ce qui les différencie des continues.

L'occlusion passe par trois phases :

* Fermeture totale de la cavité buccale ;
* Retenue de l’air ;
* Relâchement brusque et libération de l'air contenu dans la cavité buccale

Selon le mode et la zone d’articulation, on distingue les occlusives orales et les occlusives nasales.

**III-2-2 Les fricatives**

Les consonnes fricatives se concrétisent grâce à un resserrement du canal vocal qui ne va pas jusqu’à la fermeture totale. Ce sont essentiellement les lèvres et la langue qui, selon leur position et leur tension musculaire particulière, conditionnent le type de friction réalisée.

***Figure 8 :*** Consonnes fricatives dorsales et latérales

**III-2-3 Sifflantes alvéolaires**

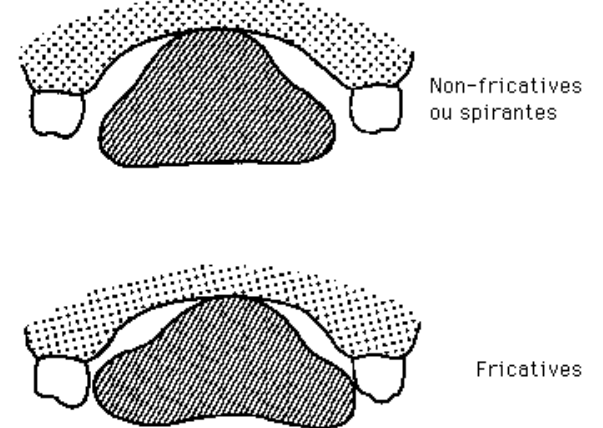
Les sifflantes alvéolaires sont produites par le rapprochement de la pointe de la langue vers la région alvéolaire. On peut diviser cette variété de sifflantes en trois catégories, selon que le dessus de la langue (anglais), son extrémité (castillan) ou la partie antérieure de son dos (français) entre en jeu ; la qualité du son en est sensiblement altérée, mais il n'existe pas de notation particulière pour marquer ces nuances dans l'alphabet phonétique de l'API (Alphabet Phonétique International).

****

***Figure 9 :*** Sifflante (fricative) alvéolaire sourde/ sonore [s]/[z]

**III-2-4 Les Latérales**

On appelle ces articulations latérales car, lors de leur réalisation, le dos de la langue prend contact avec le palais, alors que l'avant de celle-ci s'affaisse pour laisser s'écouler l'air interne par un canal latéral ou parfois bilatéral (cf. figure ci-dessous). À l'inverse, pour les articulations dorsales, la langue prend appui sur les molaires, et l'air s'écoule par un canal médian, sur le dos de la langue.)



***Figure 10 :*** Position de la langue pour les fricatives et les spirantes latérales

**VI. Mécanisme de l'audition**

Tout commence par l’émission d’un son qui est une vibration de l’air dans lequel il se propage et parvient jusqu’à l’oreille. L’onde est captée par l’oreille et met en vibration le tympan. Les vibrations du tympan sont transmises aux parties internes de l’oreille, puis transformées en influx nerveux. Par le nerf auditif, l’influx nerveux arrive jusqu’aux aires auditives du cerveau qui donnent naissance à la sensation auditive. Celle-ci est interprétée par d’autres éléments du cerveau.

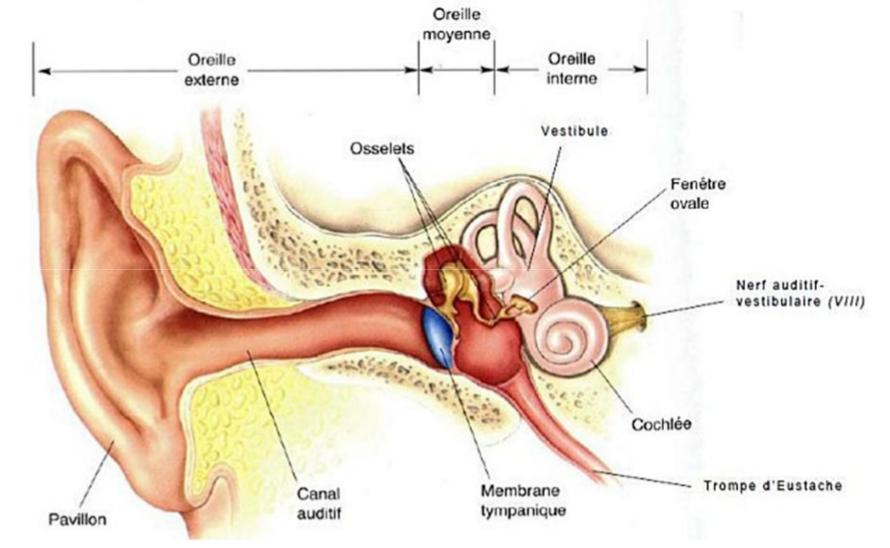
Les sons sont caractérisés par leur niveau et leur fréquence. Le niveau ou intensité donne la sensation de son faible et de son fort ; il est mesuré en décibels (dB). En dessous de 0 décibel, seuil de l’audition humaine, la perception est impossible, tandis qu’au-dessus de 120 décibels, le son est douloureux et nocif pour l’oreille. La plupart des sons de l’environnement sont compris entre 30 et 90 décibels. A chaque fois que le son augmente de 10 décibels, il est perçu deux fois plus fort.

Voici quelques exemples d’intensité de son :

* Des chuchotements ont une intensité de 25 à 35 dB ;
* Une conversation a une intensité de 55 à 75 dB ;
* Un concert peut avoir une intensité de 110 dB.

La fréquence, elle, donne la sensation de son grave et de son aigu ; c’est le nombre de vibrations par seconde ; elle s’exprime en hertz (Hz). La fréquence du son est exprimée en Hertz (Hz). L’oreille humaine peut entendre des sons compris entre 20 et 20 000 Hz. Au-dessous de 20 Hz, ce sont des infrasons, au-dessus de 20 000 Hz, ce sont des ultrasons. Le système auditif est l’organe de l’audition. Il comprend l’oreille et une partie nerveuse complexe, dont le fonctionnement est mal connu, appelée système auditif central. L’oreille est composée de trois parties :

* L’oreille externe qui capte les sons au niveau du pavillon et les amène jusqu’au tympan par le conduit auditif ;
* L’oreille moyenne qui a un rôle de transmission ; par l’intermédiaire de la chaîne des osselets, elle transporte les vibrations du tympan à l’oreille interne ;
* L’oreille interne renferme à la fois l’organe de l’équilibre, le labyrinthe, et la cochlée qui est l’organe de l’audition. Celle-ci abrite les cellules sensorielles ou cellules ciliées qui transforment le son en influx nerveux, transmis au cerveau par le nerf auditif ;



***Figure 11 :*** Anatomie de l’oreille

**VII. Psycho acoustique.**

La psycho acoustique se définit comme la : « psychologie auditive ».

C’est une science qui met en relation l’acoustique, l’anatomie humaine et la « conscience » auditive. C’est-à-dire qu’elle prend en compte l’aspect biologique, liée à la physiologie et aux mécanismes du fonctionnement de l’oreille d’une part et d’autre part, la perception subjective, les « idées reçues » de l’oreille et la mémoire auditive.

**VII-1** **L’orientation auditive**

**VII-1-1 la provenance**

Il ne nous parvient jamais exactement les mêmes sons aux deux oreilles. C'est pourquoi on peut localiser les sons par :

- la différence de temps (le son met un temps différent pour atteindre chacune des 2 oreilles)

- la différence d’intensité (la tête jouant le rôle d'écran, le son s’amortit d’une oreille à l’autre, surtout pour les hautes fréquences)

- la différence de phase (les oreilles n'étant pas à égales distance de la source, les deux signaux reçus ne sont pas en phase)

**VII-1-2 la distance**

Pour évaluer la distance d'une source sonore, l'oreille ne dispose d'aucun autre moyen d'évaluation que l'expérience. En effet, outre une différence de niveau, il existe également une différence de timbre entre un son proche et le même son plus lointain. Avec la distance, les composants aigus du spectre sont plus facilement absorbés que les composants graves, d'où évidemment altération du timbre. C'est par conséquent à travers la reconnaissance de ces nuances du timbre que l'oreille apprécie les distances, au fur et à mesure de son expérience.

**VII-2** **LA PERCEPTION DU TIMBRE**

Contrairement à la hauteur, à la durée ou à l'intensité, il n'existe pas d'unité de mesure pour caractériser le timbre. Celui-ci dépend de trois éléments :

- le spectre

- l'enveloppe

- l'environnement d'écoute

**VII-2-1 le spectre**

Le spectre d'un son se caractérise par :

- le nombre de composants (un seul pour un signal sinusoïdal, quelques-uns pour certains sons de synthèse, des centaines pour des signaux complexes)

- la répartition des composants dans le spectre (harmoniques équidistants pour des sons périodiques, partiels non équidistants pour des signaux non périodiques)

- l’intensité de chaque composants (certains auront peut-être tellement d'énergie qu'ils seront des formants)

**VII-2-2 l’enveloppe**

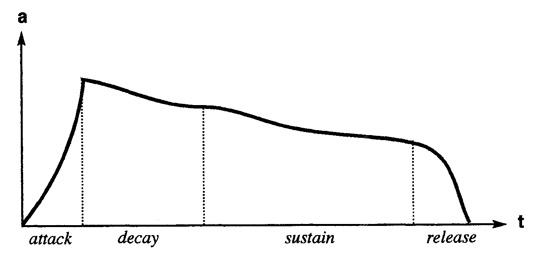
L'enveloppe se caractérise par :

- le transitoire d'attaque (attack)

- la phase de décroissance (decay)

- la phase de soutien (sustain)

- le transitoire d''extinction (release)



***Figure 12 :*** l’enveloppe du son

**VII-3** **LA PERCEPTION DES HAUTEURS**

La tonie est la hauteur subjective que l'on attribue au son (alors que la fréquence est une mesure objective du son).

**VII-3-1** **la tonie et la hauteur**

**La hauteur spectrale**

L'oreille peut entendre un fondamental absent en reconstituant la fréquence comprise entre deux harmoniques de rang consécutif.

**La hauteur formantique**

Si l'on crée un formant au niveau par exemple de l'harmonique 3 d'un son, on entendra peut-être ce son à la quinte de sa fondamentale. Alors qu'un appareil de mesure électronique indiquerait à juste titre la fréquence du fondamental comme étant la véritable fréquence du son, l'oreille, elle, percevrait une tonie différente.

**VII-3-2 la tonie et l’intensité**

Lorsque l'intensité s'accroît, les sons graves semblent baisser, alors que les sons aigus semblent monter.

**VII-3-3 la tonie et la durée**

Cette théorie est hypothétique, mais cependant intéressante. On dit souvent que la seconde psychologique s'écoule plus ou moins vite en fonction du degré d'ennui ou de stress de l'individu. Prenons le cas d'un chanteur en situation de concert, donc en état d'excitation : sa seconde psychologique se trouvera raccourcie. Pendant cette seconde plus courte, une note chantée correspondant par exemple au la 3 n'aura pas le temps d'accomplir 440 vibrations comme elle devrait, mais peut-être seulement 430. Ainsi le chanteur aura l'impression d'être trop bas et chantera instinctivement plus haut. Est-ce la raison pour laquelle on constate souvent une montée du diapason lorsque le drame s'accentue ?

**VII-3-4** **le vibrato**

Un vibrato dépend de 2 paramètres :

- la profondeur de modulation (l’intervalle parcouru à chaque fluctuation)

- la vitesse (nombre de fluctuations par unité de temps)

Si la vitesse du vibrato est inférieure à 6 Hz (6 fluctuations par seconde), l’oreille entend une ondulation dont la variation de hauteur est tout à fait nette. Au-delà de 10 Hz, on perçoit un trémolo de notes différentes. En revanche, si la vitesse est comprise entre 6 et 10 Hz, ce n'est plus une variation de hauteur que l'oreille entend, mais une variation d'intensité.

**VII-4 les intervalles mélodiques**

**L’attraction**

Dans certaines musiques, des notes donnent parfois l'impression d'être attirées par d'autres, comme la sensible vers la tonique dans un contexte tonal. Ce phénomène, uniquement culturel, incite l'oreille à ne pas considérer, par exemple, comme fausse (mais simplement comme tendue) une note sensible trop haute de 5 ou 10 savarts avant sa résolution.

**La tolérance**

Parfois l'oreille, consciemment tout d'abord, puis par automatisme, arrive à s'accommoder de notes perçues comme fausses par rapport à une échelle donnée (un piano mal accordé, un tempérament inégal entendu pour la première fois). Précisons que la plupart des notes produites par un chanteur seraient fausses (par rapport au tempérament égal) pour un instrument de mesure électronique.

**VII-5 les traumatismes auditifs**

Un bruit intense ou incessant entraîne chez l’individu des troubles du métabolisme (fatigue, agressivité...). Certains sons de basse fréquence, à très forte intensité, peuvent même entrer en résonance avec des parties du corps humain et les affecter. Outre le vieillissement, l’écoute d’une musique à un niveau sonore important (comme dans certaines discothèques) altère les facultés auditives. La plupart des personnes âgées ne perçoivent plus les fréquences supérieures à 10 000 Hz, c’est-à-dire qu’elles entendent un spectre moins large dont le timbre est par conséquent modifié. C’est la disparition de ces composants aigus (où est contenue l’intelligibilité de la parole) qui rend parfois difficile la compréhension d’un interlocuteur.

**VII-6 Illusion auditive**

Les illusions auditives proviennent d’un paradoxe entre la réalité physique des stimuli de l’oreille et leur perception. Véritables hologrammes audio, on constate par l’expérience que ces phénomènes suscitent des réactions bien différentes dans un panel de plusieurs individus.

**VIII. Masquage et bandes critiques.**

Le masquage auditif est souvent observé dans la vie quotidienne : en marchant dans une rue, un camion qui passe peut perturber notre conversation. De même, on ne voudrait pas qu'un téléphone portable sonne dans les parties calmes d'un concert classique, mais il pourrait même ne pas être entendu dans un concert de rock bruyant. Les composants d'un son interagissent avec les composants d'un autre son de fréquence et de temps similaires et les rendent inaudibles. C'est ce qu'on appelle le masquage. Un masquage partiel peut également se produire : les composants ne sont pas inaudibles, mais leur volume est réduit

Bibliographie

http://www.vivaarte.fr/bordarts/pages/licence3/acoustique/acoustique.htm