



**HAL**  
open science

## Partie 5. Etudes de groupes de cas. Evaluation et Réhabilitation de la Voix

Anne Menin-Sicard, Etienne Sicard

► **To cite this version:**

Anne Menin-Sicard, Etienne Sicard. Partie 5. Etudes de groupes de cas. Evaluation et Réhabilitation de la Voix. Evaluation Et Rehabilitation De La Voix, Deboeck Solal, 2016, 978-2-35327-318-8. hal-02306658

**HAL Id: hal-02306658**

**<https://hal.science/hal-02306658>**

Submitted on 14 Oct 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Partie 5. Etudes de groupes de cas

Dans ce chapitre, nous étudierons l'analyse des **caractéristiques de groupes de voix**, en se basant sur des **outils statistiques** afin de dégager des tendances. Ces outils ont été élaborés pour faciliter les études statistiques sur un grand nombre de cas dans le but de valider les indicateurs d'altération de la voix choisis pour Vocalab. Après avoir détaillé les bases théoriques qui ont été le fondement de ce travail, nous évoquons les **corrélations**, indispensables pour la sélection des indicateurs pertinents. Des études de cas en fonction de pathologies fréquemment retrouvées en orthophonie seront décrites ainsi que le traitement de données concernant la presbyphonie. Ces données sont en partie issues de **mémoires d'orthophonie** donc nous avons assuré l'encadrement ou le co-encadrement

## Sommaire

1. Bases théoriques
2. La base de données de voix
3. L'outil statistique de VOCALAB
4. Statistiques sur les indicateurs
5. Conclusion

## Contenu

Partie 5. Etudes de groupes de cas .....	1
Bases théoriques .....	2
Notion de moyenne .....	2
Histogrammes .....	3
Corrélation entre données .....	6
La base de données de voix .....	8
Motivation .....	8
La base de données de voix du LURCO .....	9
L'outil Statistiques de VOCALAB .....	10
Fonctionnalités proposées .....	10
Fondamental .....	12
Etude de cas – Variabilité pour un même patient .....	14
Etude de cas – Nodules .....	15
Statistiques sur les indicateurs .....	16
Calculs sur 5 indicateurs .....	16
Etude de cas – Nodules .....	18

Etude de cas – Voix Cancer.....	19
Etude de cas – Voix âgées.....	20
Etude de cas - Presbyphonie.....	23
Corrélation entre indicateurs.....	23
Analyse long terme.....	25
Conclusion.....	27

## Bases théoriques

### Notion de moyenne

La notion de moyenne nous est très familière, il s'agit souvent de la moyenne « **arithmétique** ». Celle-ci est la somme des valeurs numériques  $x_i$  de la liste, divisée par le nombre  $n$  de ces valeurs (Equation 5-1). Bien que très simple à calculer, cette moyenne est sensible aux valeurs extrêmes des valeurs de la liste. Par exemple, si une valeur d'un indicateur est très supérieure à la moyenne, et si le nombre  $n$  de cas étudiés est relativement peu élevé (ce qui est souvent le cas dans le contexte de l'analyse de voix pathologiques), cette seule valeur extrême influence fortement la valeur moyenne et peut induire une interprétation erronée sur la tendance de l'ensemble des données.

$$\overline{x_{ARITH}} = \frac{1}{n}(x_1 + x_2 + \dots x_n) \quad (1)$$

Une solution à ce problème consisterait à enlever de la série les valeurs extrêmes, en l'occurrence, la plus basse et la plus haute tout en disposant d'un nombre significatif de points pour que la moyenne reste pertinente. Si le nombre  $n$  d'échantillons initial est peu important, cette approche est à éviter.

Une alternative, implémentée dans VOCALAB4, est la **moyenne géométrique**, décrite par l'équation 2. Cette moyenne est moins sensible que la moyenne arithmétique aux valeurs les plus élevées d'une série de données. Elle donne, par conséquent, une meilleure estimation de la tendance centrale des données, surtout lorsque les valeurs varient fortement et que le nombre de cas étudié est peu important.

$$\overline{x_{GEO}} = \sqrt[n]{x_1 \times x_2 \times \dots x_n} \quad (2)$$

Un exemple de calcul de moyenne est celui du Fondamental Usuel Moyen (FUM) de l'outil **Evaluation > FO+Etendue Vocale**. Il s'agit de la moyenne géométrique des fréquences fondamentales instantanées calculées sur les sons voisés. Sa valeur est donnée en haut à droite de l'écran, soit en temps réel si le bouton « Fondamental » est activé, soit en temps différé d'après lecture d'un fichier son (Figure 1).



Figure 1 : exemple de calcul de moyenne des F0 sur une voix parlée permettant d'extraire le fondamental usuel moyen (FUM)

Sous EXCEL, le calcul d'une moyenne de valeurs s'exprime grâce à la fonction « =**MOYENNE(A2:A12)** » par exemple, où les données sont positionnées dans les cases allant de A2 à A12. Le terme **MOYENNE.GEOMETRIQUE** permet d'accéder à la moyenne géométrique.

## Histogrammes

Les histogrammes sont une représentation d'une probabilité en fonction d'une valeur. Un histogramme important pour caractériser la voix est celui du fondamental, que l'on trouve dans l'outil **Evaluation > FO+Etendue Vocale**. Il s'agit ici d'une probabilité d'apparition du F0 instantané dans un petit intervalle de fréquences, de l'ordre d'un quart de ton, compromis entre un intervalle trop restreint qui « diluerait » la figure, et un intervalle trop large qui ne permettrait pas de dégager des informations significatives. A noter que l'échelle des fréquences n'est pas linéaire, mais exponentielle, afin de reproduire la perception de l'oreille. Un quart de ton vaut environ 2 Hz autour du Do1, 4 Hz autour du Do2, 8 Hz autour du Do3, 16 Hz autour du Do4, etc.

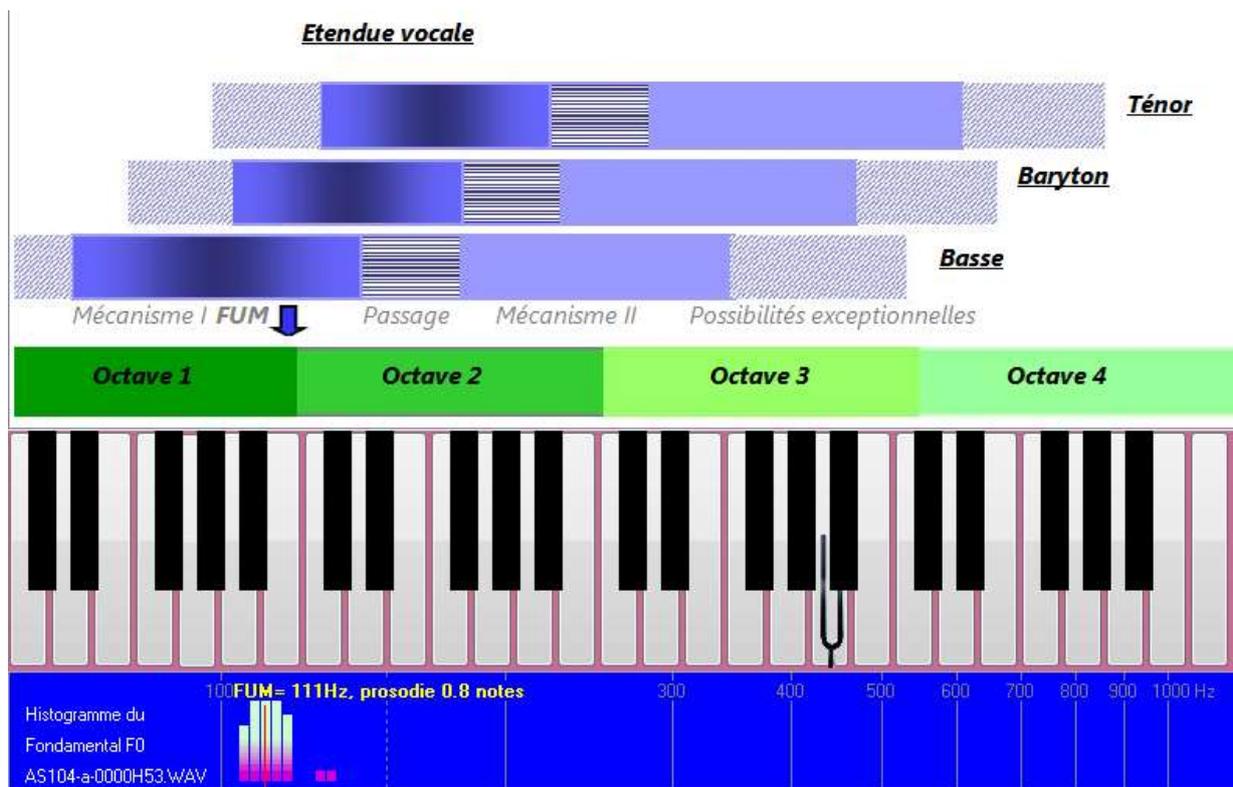


Figure 2: Histogramme d'un /a/ stable (AS104), qui donne quasiment un pic autour du FUM

Lorsque le fondamental est parfaitement stable, la probabilité vaut 100% à la valeur donnée, et devient nulle ailleurs. Un exemple d'histogramme d'un /a/ d'une voix normale est donné figure 2. Les valeurs successives du fondamental sont très proches, et se répartissent principalement dans 5 intervalles de fréquence. L'histogramme du /a/ est dans ce cas très resserré.

Autant l'histogramme d'un /a/ n'est pas très informatif en soi, autant l'histogramme de la voix parlée (par exemple sur la base d'un texte lu de 10-20 secondes, comme illustré figure 3) est très intéressant à analyser et pertinent cliniquement. En plus de permettre l'extraction du FUM grâce au calcul de moyenne explicite plus haut, il donne des indications sur :

- La plage de variation de F0, de part et d'autres du FUM, ce qui est relié en statistiques à la notion d'écart type, ou encore de moment d'ordre 2. Cette notion est illustrée figure 3. La notion de prosodie exprimée en notes dans VOCALAB vaut environ 2 fois l'écart type, et regroupe plus des 2/3 des valeurs.
- La dissymétrie de l'histogramme par rapport au FUM, est reliée en statistiques à la notion de Skewness, ou encore de moment d'ordre 3, illustrée figure 4. La dissymétrie donne une indication des possibilités plus limitées vers les basses ou les hautes fréquences.

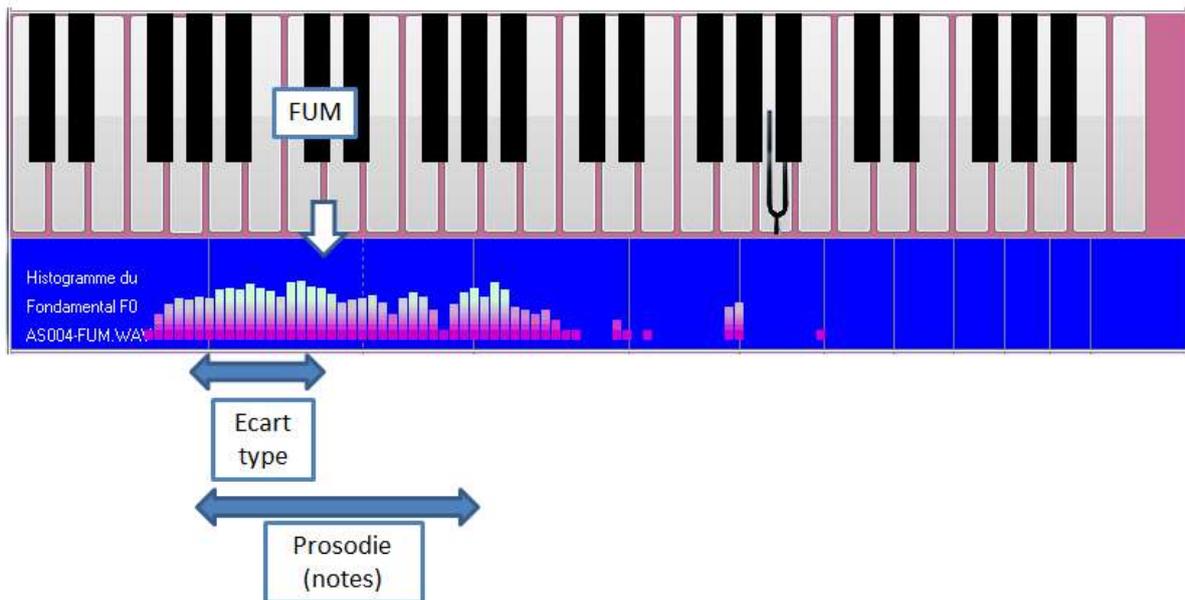


Figure 3 : FUM et prosodie d'un texte lu, voix normale (cas AS004 de la base de données)

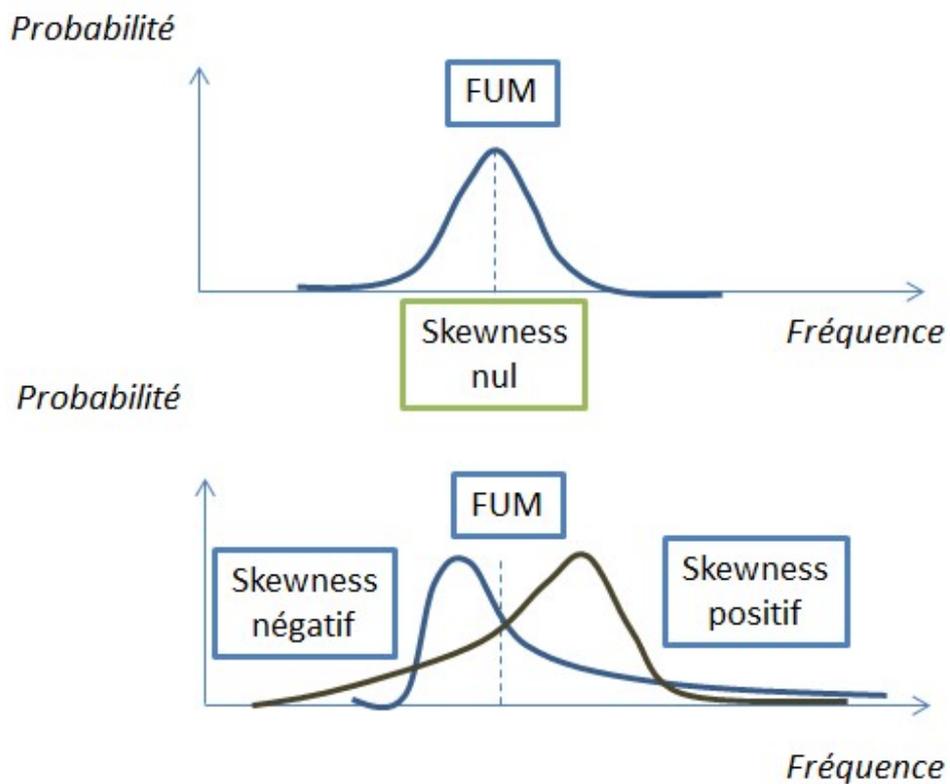


Figure 4: Histogramme de la voix parlée normale (en haut), et déformation vers la gauche ou vers la droite (notion de skewness ou moment d'ordre 3)

Anne Menin-Sicard, Etienne Sicard. (2016) Partie 5. Etudes de groupes de cas. Evaluation et Réhabilitation de la Voix, Deboeck Solal, ISBN 978-2-35327-318-8

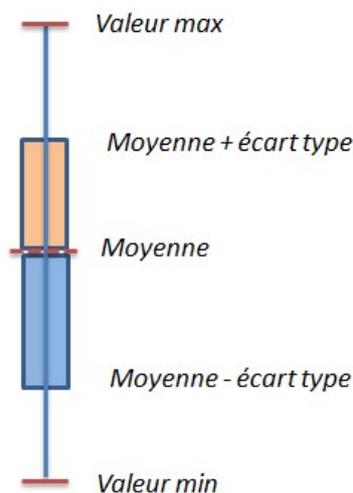


Figure 5 : représentation de l'histogramme sous forme min/moyenne/Max avec écart type

Une représentation simplifiée d'un histogramme peut être faite en ne faisant figurer que le minimum, la moyenne et le maximum, avec une représentation symbolique de l'écart type (Figure 5). Nous utiliserons cette représentation pour illustrer la variation d'un indicateur pour un ensemble d'échantillons.

## Corrélation entre données

La corrélation, ou plus exactement le *coefficient de corrélation linéaire* entre deux variables aléatoires X et Y, se calcule par la formule ci-dessous.

$$Cor(X, Y) = \frac{1}{N \cdot \sigma_X \cdot \sigma_Y} \sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y}) \quad (3)$$

avec

$$\begin{aligned} \bar{X} &= \text{moyenne } X \\ \bar{Y} &= \text{moyenne } Y \\ \sigma_X &= \text{écart type } X \\ \sigma_Y &= \text{écart type } Y \\ N &= \text{nombre d'échantillons} \end{aligned}$$

Ce calcul donne un résultat compris entre -1.0 et +1.0.

- si *Cor* est proche de 0, il n'y a **pas de lien** entre X et Y. Les deux variables sont indépendantes.
- si *Cor* est proche de -1, **X est lié à Y** de manière négative. Les deux variables sont corrélées, si l'une augmente, l'autre diminue.
- si *Cor* est proche de 1, **X est lié à Y** de manière positive. Les deux variables sont corrélées et dépendantes l'une de l'autre avec la même tendance.

En pratique, X et Y sont deux ensembles de données (par exemple, une série de valeurs d'un indicateur, une série de valeurs de Handicap Index, le FUM, la prosodie, etc.), dont on cherche à savoir s'il existe une forme de dépendance entre elles.

### Exemple

- ▶ On suppose ici que l'on dispose de deux séries de données X et Y (Barbera & Frenay 2016), l'une de l'auto-évaluation du patient (de 5/25 à 25/25, 25 étant le degré de satisfaction le plus élevé), l'autre du Voice Handicap Index (de 0/40 à 40/40, 40 étant le handicap le plus élevé). Le tableau de données est donné ci-après.
- ▶ On cherche à évaluer s'il existe une corrélation entre ces deux séries de données, autrement dit si ces valeurs expriment des phénomènes similaires.
- ▶ Sous EXCEL, la commande permettant de calculer le coefficient de corrélation est **=COEFFICIENT.CORRELATION(A2:A17; B2:B17)**, où **A2 :A17** représente la 1<sup>ère</sup> série de données située en colonne A, lignes 2 à 17, **C2 :C17** la seconde, correspondant à la colonne B.
- ▶ Le résultat vaut **-0.90** ce qui veut dire que les données sont très fortement corrélées, mais que la tendance va dans le **sens inverse** : plus le score d'auto-évaluation est élevé, plus le VHI est faible.

Série de données X Auto-évaluation ( /25)	Série de données Y VHI-10 ( /40)
10	34
11	23
14	14
15	21
15	19
15	13
16	13
16	9
17	7
17	11
18	8
18	10
19	8
19	9
20	8
21	6
22	5
23	3
24	0
(colonne A)	(Colonne B)

En principe, l'emploi de la formule donnée par l'équation 3 (encore appelé coefficient de Pearson) ne donne des résultats fiables que si les séries de données d'origine X et Y ont une distribution statistique de type gaussien, c'est-à-dire des courbes « en cloche » comme dans la figure 4 en haut et comportent peu de valeur exceptionnelles.

Le coefficient de corrélation de Spearman permet lui aussi de détecter l'existence de relations entre 2 séries de données X et Y, lorsque leur relation s'éloigne d'une dépendance linéaire, c'est-à-dire d'une simple droite, la distribution statistique de X ou Y n'est plus gaussienne, est dissymétrique, comme illustré figure 4 en bas avec un *skewness* non nul, ou encore comporte un nombre significatif de valeurs exceptionnelles. Ces conditions se retrouvent hélas souvent dans l'étude d'un nombre restreint et hétérogène de voix pathologiques, ce qui invite à une interprétation prudente des données de corrélation.

## La base de données de voix

### Motivation

Les études statistiques doivent porter sur un nombre de cas assez grands pour pouvoir être significatives. De nombreuses banques de données de voix sont citées dans les publications scientifiques traitant de l'évaluation et la rééducation de la voix. Nous listons une sélection d'articles traitant de ce sujet dans l'étude bibliographique en ligne (Sicard, 2013). La plupart de ces études portent sur une centaine de cas de voix, en général pour moitié normales et pour l'autre moitié des cas de pathologies variées.

Certaines études de synthèse telles que Wuyts (2000), Ghio (2008) ou Werth (2010) portent sur plusieurs centaines jusqu'à plus d'un millier de cas de voix. Dans la plupart des cas, les échantillons sonores et les données de base caractérisant le patient ne sont pas accessibles gratuitement, sont soit associés à une licence logicielle (comme pour le *Multi-Dimensional Voice Program*, Deliyski 2012), soit classés confidentiels et restreints à un usage interne en milieu hospitalier, donc difficilement accessible pour les orthophonistes.

Dès les premières versions de VOCALAB, nous avons regroupé au sein d'une médiathèque différents cas de voix, allant des voix normales, chantées aux voix altérées et pathologiques (Figure 6). Ces fichiers permettent de conduire des études de cas, mais pas des statistiques. Les dénominations utilisées dans la médiathèque ont été établies par une combinaison de jury d'écoute et d'analyse objective. Il est à noter que plusieurs cas de voix proposés dans la médiathèque donnent des indications d'altération multiples, dont nous ne retenons, avec une certaine subjectivité, que la caractéristique qui nous semble la plus significative.

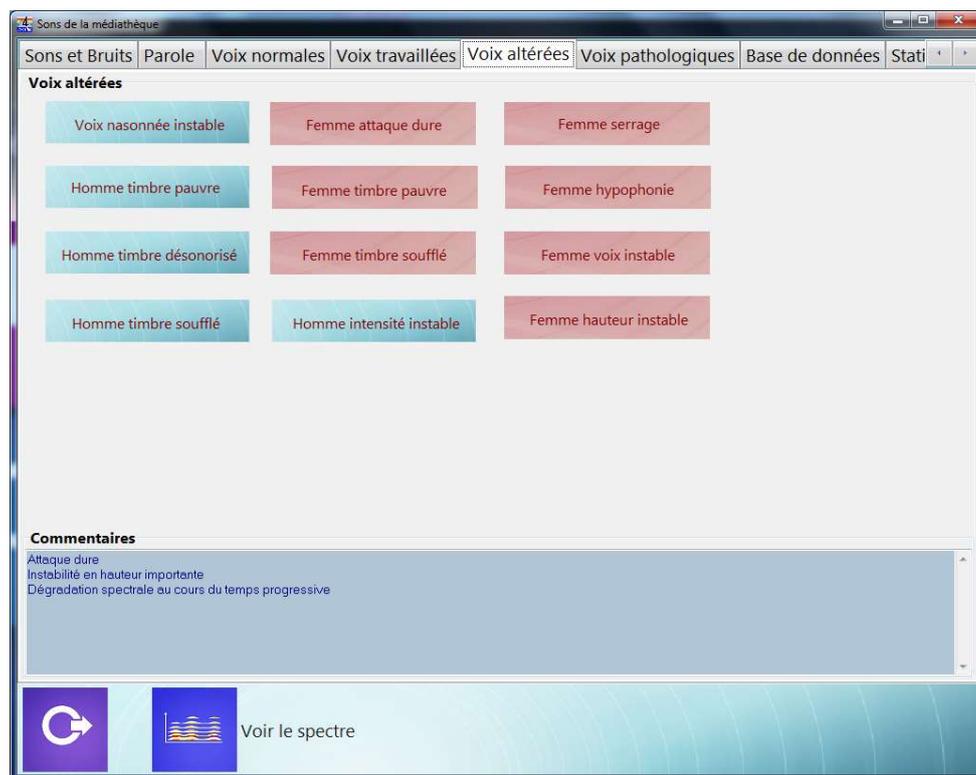


Figure 6 : La médiathèque de VOCALAB comporte quelques exemples de voix normales, travaillées, altérées et pathologiques.

## La base de données de voix du LURCO

Dans le cadre de l'ERU 15 du laboratoire LURCO, 200 cas de voix ont été regroupés, dont 150 environ sont pathologiques, avant et après rééducation orthophonique. Les échantillons mis en ligne concernent principalement la production d'une voyelle /a/ en voix conversationnelle, d'une durée de quelques secondes, et des sirènes.

Dans le respect de la confidentialité de l'identité des patients selon les règles de la CNIL, l'objectif est mettre les échantillons sonores librement accessibles sur Internet (Sicard, 2012). Les orthophonistes partenaires de l'ERU 15 ont effectué un enregistrement de voyelle /a/ et de sirène (/a/, /ou/), tout d'abord en bilan, puis en fin de rééducation. Les spectres des voix ont été analysés et commentés, complétés par une description des axes thérapeutiques et des résultats fournis par les indicateurs du logiciel VOCALAB (Sicard, 2016) dans une fiche de synthèse construite à cet effet. Un exemple de fiche est détaillé en deuxième partie de ce chapitre.

Le but de la base de données est d'intégrer les pathologies de voix les plus diverses, tout en couvrant des âges, sexe et contextes d'enregistrement différents. Les échantillons sonores sont de qualité variable, plus ou moins bruités, tels que produits dans des cabinets d'orthophonie et non dans des cabines audiométriques.

Bien qu'enregistrés selon des consignes similaires, les échantillons sont capturés avec des microphones et chaînes d'acquisition pouvant altérer la qualité de différentes façons :

- bruit 50 Hz ;

- sifflements ;
- bruit d'ordinateur ;
- coupure basse ou haute fréquence ;
- résonance.

Cette grande variabilité de conditions d'enregistrement constitue un défi pour l'interprétation des résultats issus d'outils d'évaluation objectifs, dans la mesure où le bruit, la faible dynamique, les interférences diverses influent sur les résultats eux-mêmes. Une voix enregistrée avec différentes chaînes d'acquisition (Delyski 2005) et un même logiciel, et plus encore avec des logiciels différents donne des résultats pouvant varier de l'ordre du 1/4 de ton pour le FUM, et de l'ordre de plusieurs tons pour l'étendue vocale. Il est donc recommandé de conduire des études statistiques avec une configuration matérielle et logicielle unique, et des réglages identiques, sous peine de faire porter des calculs et exploiter des résultats scientifiquement non valides, et par conséquent des interprétations erronées.

Nous devons aussi connaître les limites des outils et évaluer les biais méthodologiques de la collecte des données, permettant de développer un regard critique sur les résultats des indicateurs recueillis en situation clinique.

## L'outil Statistiques de VOCALAB

Analyser rapidement et efficacement les données d'un grand ensemble de fichiers de voix nécessite un outil spécifique accessible dans VOCALAB dans l'outil **Médiathèque > Sons > Statistiques**. L'écran en haut à gauche de choisir le répertoire de travail. Par défaut, il s'agit de **div\_vocalab > stats**. Nous avons sélectionné dans cet exemple un répertoire **stats\_normal** dans lequel nous avons regroupé un ensemble de /a/ de voix normales extrait de la base de données. La liste des fichiers WAV du répertoire ciblé est donnée en bas à gauche.

### Fonctionnalités proposées

Cliquer « **Calculer Tout** » pour obtenir les données primaires extraites de l'analyse des fichiers sons présents dans le répertoire actif (Figure 7). Cela peut prendre plusieurs minutes, selon la durée des fichiers sons et la puissance de calcul de l'ordinateur.

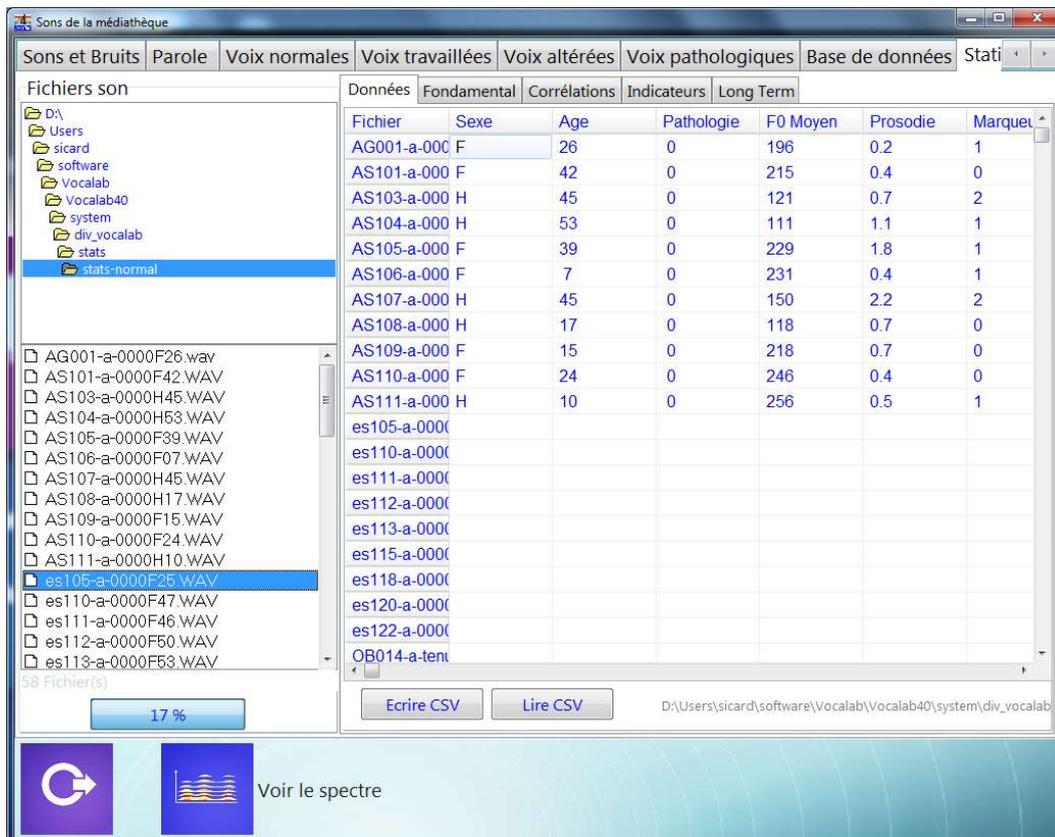


Figure 7 : Lancement du calcul des paramètres génériques pour la série de fichiers incluse dans un répertoire donné (ici, les voix normales disponibles dans la base de données en ligne)

## Données primaires

- ▶ Sexe : décodé d'après la lettre H/F/E identifiés dans le nom du fichier
- ▶ Age : décodé d'après les 2 derniers chiffres du nom du fichier
- ▶ Pathologie : décodé d'après les 4 chiffres identifiés dans le nom du fichier
- ▶ F0 Moyen (en Hz): moyenne des F0 instantanés de l'échantillon
- ▶ Prosodie : en notes, calculée d'après l'écart type des F0 instantanés
- ▶ Marqueurs : nombre d'indicateurs dont la valeur est > 1.0
- ▶ Altération de l'attaque
- ▶ Instabilité en hauteur
- ▶ Instabilité en amplitude
- ▶ Bruit/Signal
- ▶ Pauvreté harmonique
- ▶ Instabilité en hauteur 1 seconde après l'attaque
- ▶ Instabilité en amplitude 1 seconde après l'attaque
- ▶ Bruit/Signal 1 seconde après l'attaque
- ▶ Pauvreté harmonique 1 seconde après l'attaque
- ▶ Autoévaluation ou Handicap Index, décodé d'après AE ou HI dans le nom du fichier

A noter que plusieurs données sont décodées d'après le nom du fichier. Par exemple, si le nom du fichier est **sp015-a-AE12-1211F27.WAV**, les données telles que l'auto-évaluation (12/25), la pathologie (1211), le sexe (F) et l'âge (27 ans) sont automatiquement extraits.

L'écran de droite comprend 5 onglets.

- **Données** : tableau qui regroupe les informations primaires sur chaque fichier : nom du fichier, sexe, âge, pathologie, prosodie, nombre d'indicateurs pathologiques, valeur des indicateurs. Ces données peuvent être sauvegardées au format CSV (Bouton « **Ecrire CSV** ») en vue d'une exploitation par Microsoft EXCEL™ ou OpenOffice CALC. Les données sont calculées en cliquant « **Calculer Tout** ». Elles peuvent aussi être importées d'après un fichier tabulé grâce au bouton « **Lire CSV** »
- **Fondamental** : positionne les fichiers dans un espace Age/Fréquence.
- **Corrélations** : calcule les corrélations entre indicateurs
- **Indicateurs** : dessine les valeurs d'indicateurs pour un ensemble de fichiers
- **Long Term** : calcule le spectre moyen de l'ensemble des fichiers

## Fondamental

La figure 8 donne la répartition de l'âge en fonction de F0 pour 60 voix normales (homme, femme, enfant de 5 à 95 ans). La courbe typique d'évolution de F0 en fonction de l'âge est dessinée à gauche pour les hommes, à droite pour les femmes. De 1 an à 18 ans, le fondamental décroît de 350 à 125 Hz en moyenne pour les hommes, et de 350 à 225 Hz pour les femmes /RefKent/. La tendance est ensuite à la stabilité de 20 à 65 ans, avec une distinction homme/femme autour de 170 Hz. A partir de 65 ans, une tendance au resserrement pour les voix âgées est observée : le F0 des voix d'homme augmente, tandis que celui des voix de femme diminue.

Les 60 cas de voix normales suivent plus ou moins la tendance « théorique » telle que proposée par Baken (1999) ou Kent (2000), les F0 étant cependant fortement dispersés dans l'espace des fréquences.

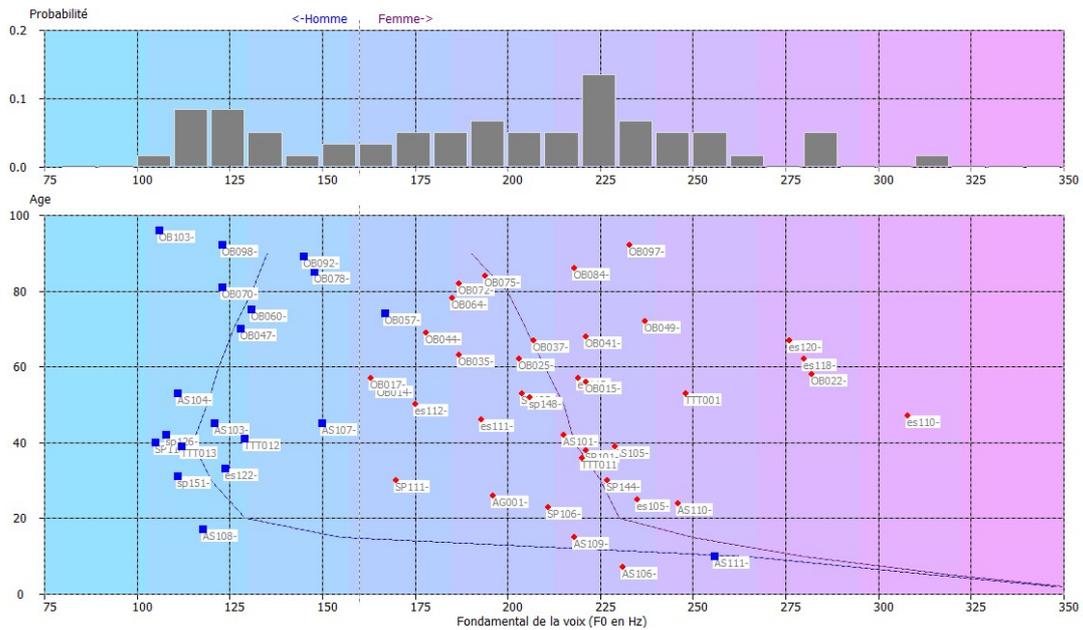


Figure 8 : répartition du F0 et de l'âge de 60 /a/ de voix normales

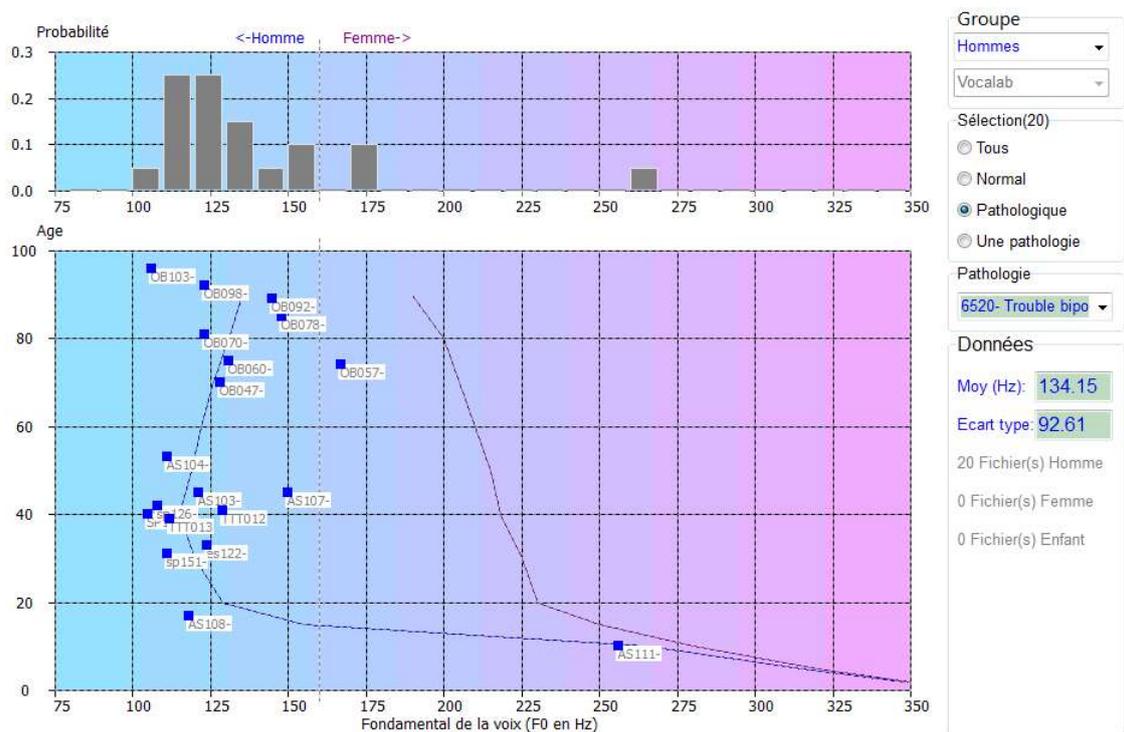


Figure 9 : sélection du groupe « hommes » et actualisation des graphiques et données correspondantes

L'écran de droite permet d'affiner la sélection, par exemple ne considérer que les voix d'hommes, ou encore considérer une pathologie spécifique (ne s'applique pas ici car ce sont des voix normales). L'écran se redessine alors et effectue les calculs de moyenne et d'écart type associés.

## Etude de cas – Variabilité pour un même patient

Dans le cadre de son mémoire d'orthophonie, Marie Daumet (2015) s'est enregistrée chaque semaine avec un même équipement, avec une distance bouch-micro et un réglage identique, en prononçant un /a/ impliqué, afin d'analyser notamment la variabilité du fondamental et des indicateurs sur une longue constante de temps. Le fondamental moyenné sur l'ensemble des 23 échantillons considérés ici est autour de 243 Hz, avec une variabilité de 2 tons environ (on encore 4 ½ tons), ce qui est remarquablement faible.

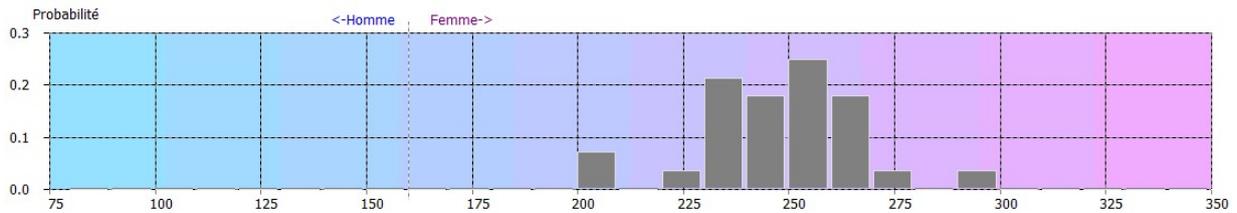


Figure 10 : Variabilité de F0 pour une même personne qui s'est enregistrée plus de 20 semaines consécutives (Daumet 2015)

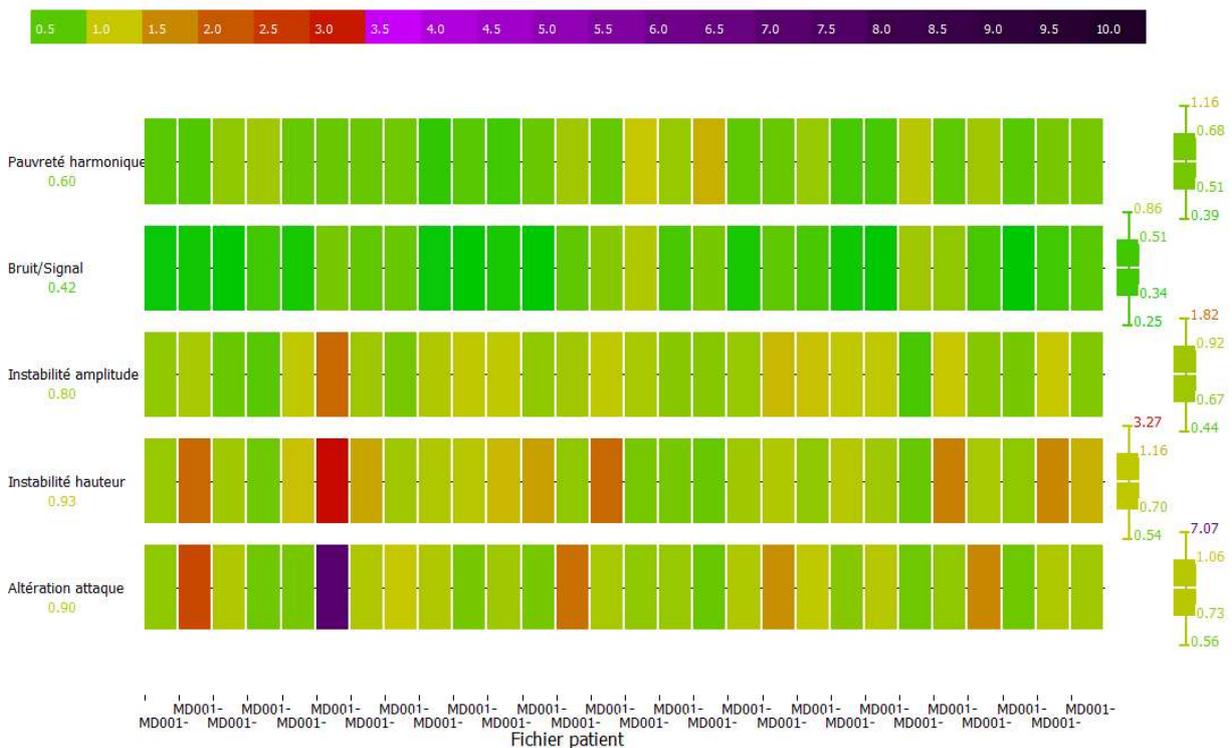


Figure 11: Variabilité des indicateurs pour une même personne qui s'est enregistrée plus de 20 semaines consécutives (Daumet, 2015)

Au cours du temps, M. Daumet a produit des /a/ aux caractéristiques très proches, à l'exception notable d'un /a/ avec une altération significative de l'attaque (6<sup>ème</sup> enregistrement en partant de la gauche, pouvant être lié à une atteinte ORL passagère). Il ressort de cette étude que l'indicateur

bruit/signal ainsi que pauvreté harmonique évoluent très peu. L'écart type est plus significatif pour l'instabilité en hauteur, en amplitude, et pour l'altération de l'attaque. Plusieurs enregistrements pour un même patient avec la même consigne et le même matériel donnent des résultats proches en termes d'indicateurs, voire très proches en ce qui concerne le bruit et la pauvreté harmonique. C'est le paramètre d'instabilité en hauteur qui varie le plus ainsi que l'attaque, deux paramètres dont on sait qu'ils sont particulièrement sensibles à l'émotion et aux humeurs.

## Etude de cas – Nodules

Les nodules déforment les cordes vocales et perturbent leur vibration. Ils sont en général la conséquence de l'accumulation de micro traumatismes des cordes vocales en lien avec le forçage vocal. Ils apparaissent en général à la jonction 1/3 antérieur - 2/3 postérieur des CV, là où l'amplitude vibratoire est la plus grande. L'objectif de notre étude de cas est de comparer la moyenne des F0 du /a/ entre voix normales et voix atteintes de nodules, et quantifier une éventuelle diminution du fondamental.

### Comparer deux groupes

- ▶ Créer un sous-répertoire dans lequel vous placez les cas de voix des témoins
- ▶ Effectuez le calcul de F0 moyen
- ▶ Créer un sous-répertoire dans lequel vous placez les cas de voix pathologiques
- ▶ Effectuez le calcul de F0 moyen
- ▶ Comparez les résultats obtenus

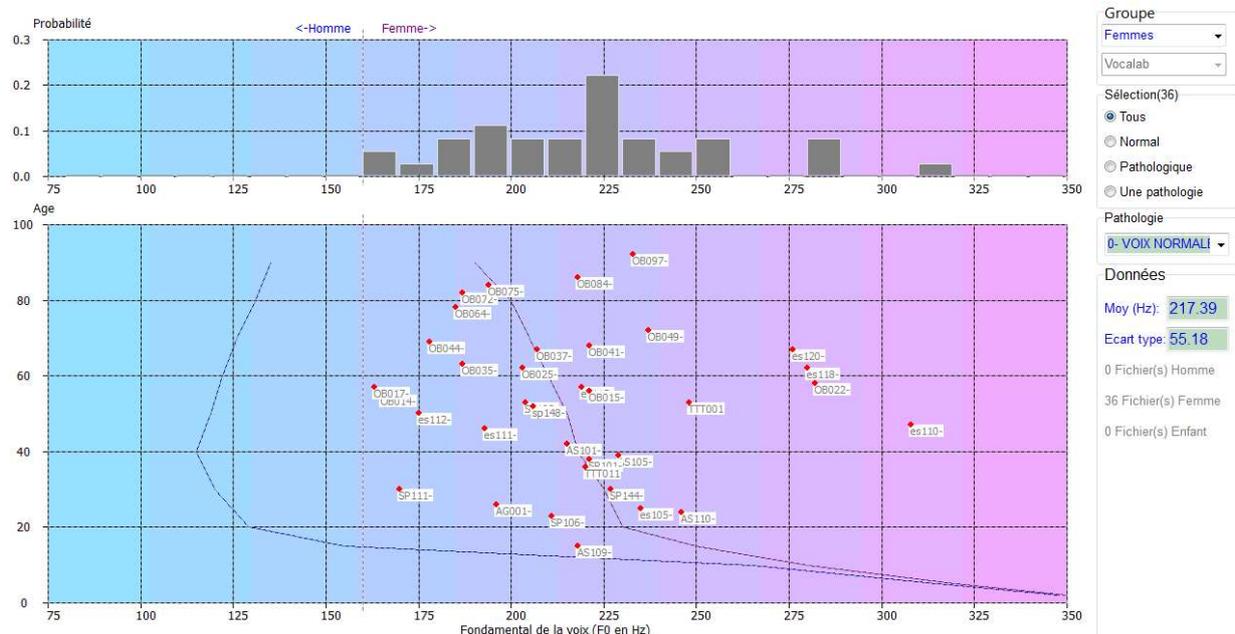


Figure 12: sélection du groupe « femmes » des 36 cas de /a/ de voix normales permettant d'extraire la moyenne du F0 (218 Hz)

Nous avons utilisé comme références les /a/ de voix normales disponibles dans la base de données, en ne prenant que le groupe des femmes. Nous avons ici 36 fichiers, dont les âges sont répartis entre 18 et 90 ans (Figure 12). Le F0 moyenné est autour de 218 Hz. A noter qu'il s'agit de /a/ et non de voix parlée. On mesure ici non pas le Fondamental Usuel Moyen, qui est calculé d'après un texte, mais le F0 moyen sur un /a/.

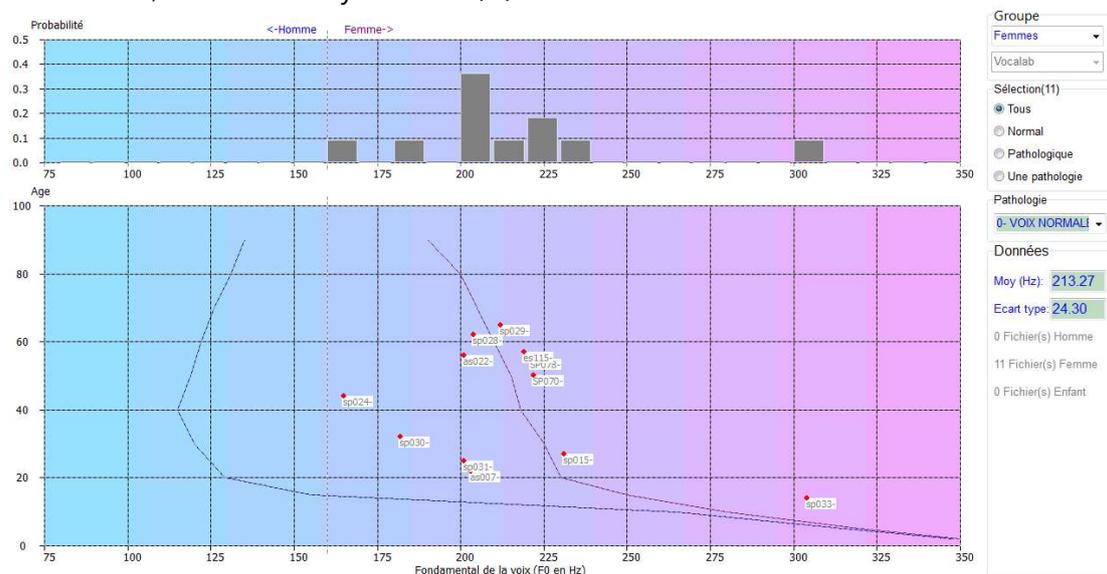


Figure 13: sélection du groupe « femmes » des /a/ de voix atteintes de nodules permettant d'extraire la moyenne du F0

Nous copions dans un nouveau répertoire 11 fichiers de /a/ avec la pathologie de type nodules, avant rééducation. Ces fichiers extraits de la base de données en ligne ont un nom contenant « **a-avant** ». Nous effectuons de nouveau les calculs et l'extraction du F0 moyen. Nous trouvons 214 Hz (Figure 13), soit une très légère diminution de F0. Cet écart est relativement peu significatif ici de par le faible nombre de fichiers analysés. Le nombre de cas nécessaire pour dégager une tendance dépend de la marge d'erreur acceptable et du niveau de confiance du résultat. Pour une marge d'erreur de l'ordre de 15% et un niveau de confiance de 7/10, l'ordre de grandeur du nombre de fichiers nécessaires serait autour de 50.

## Statistiques sur les indicateurs

### Calculs sur 5 indicateurs

L'onglet **Indicateurs** donne accès à une mise en graphique des valeurs des indicateurs pour l'ensemble des fichiers, selon le code de couleur rappelé en haut de l'écran de la figure 14. La moyenne est donnée à gauche, l'écart type, les valeurs min et max à droite. Il s'agit ici de 8 cas de maladie de Wilson transmis par Michaela Pernon (2015), orthophoniste au Centre National de Référence Wilson coordonné par le Dr Woimant, Service de Neurologie du Pr Chabriat, accessibles

via la base de données en ligne. Les excellentes conditions d'enregistrement permettent d'obtenir des échantillons sonores avec un niveau de bruit environnant très faible. Ce qui frappe dans cet exemple est l'augmentation considérable des indicateurs de pauvreté harmonique (moyenne >2.5) et dans une moindre mesure de l'indicateur bruit/signal (>1.5). Les autres indicateurs sont proches des voix normales.

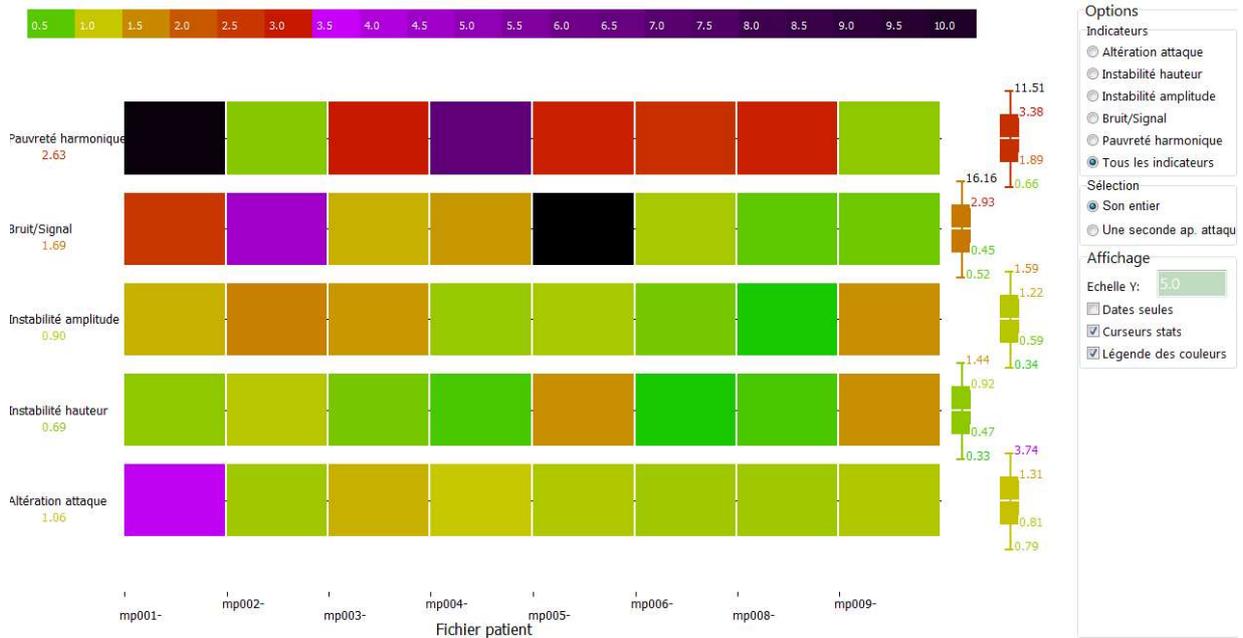


Figure 14: Représentation graphique des 5 indicateurs de VOCALAB4, avec moyennes, écarts types et valeurs min/max associées, pour des cas de maladie de Wilson (Pernon 2015)

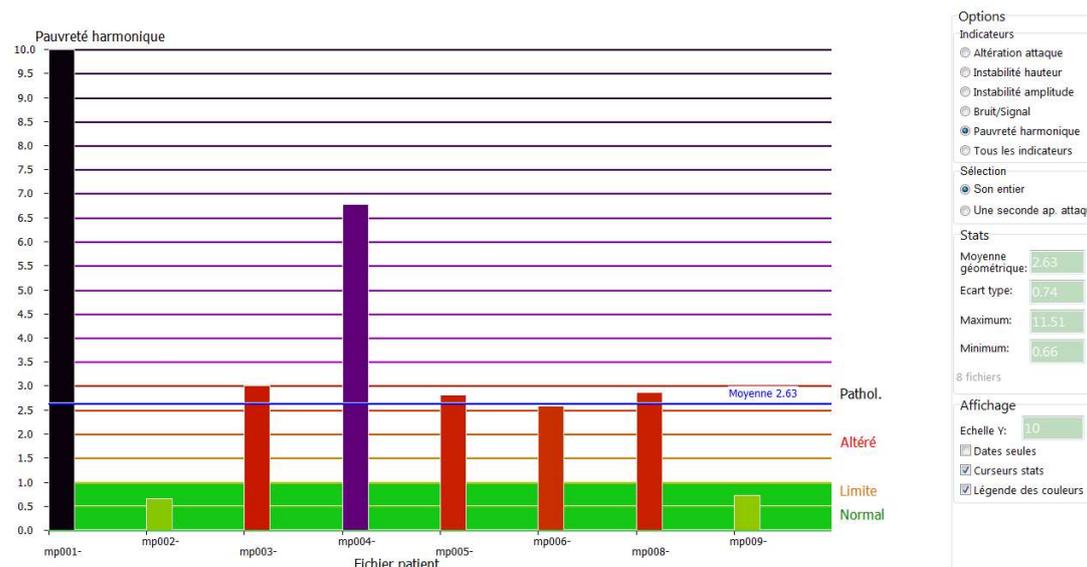


Figure 15: représentation graphique de l'indicateur de pauvreté harmonique sur 8 cas de maladie de Wilson, montrant une élévation significative pour 6 patients, 2 autres étant dans les valeurs normales.

Il est possible de visualiser un indicateur en particulier, comme la pauvreté harmonique des cas de la figure 15, pour mieux apprécier la répartition des valeurs pour chaque patient.

## Etude de cas – Nodules

Le calcul des indicateurs pour l'ensemble des fichiers de /a/ en lien avec la pathologie de type nodules avant la prise en charge orthophonique est illustré figure 16. On ne note pas de tendance significative comme dans le cas de la maladie de Wilson étudiée au paragraphe précédent. On note toutefois quatre indicateurs un peu élevés : bruit/signal, instabilité en amplitude, instabilité en hauteur et altération de l'attaque.

Ces résultats sont assez comparables à ceux publiés par (Jiang, 2009) qui note aussi, que ce soit avec les outils MDVP ou CSpeech une augmentation du 50% du *Jitter* et 40% du *Shimmer* entre le groupe normal et le groupe Nodules composé de 21 sujets chacun.

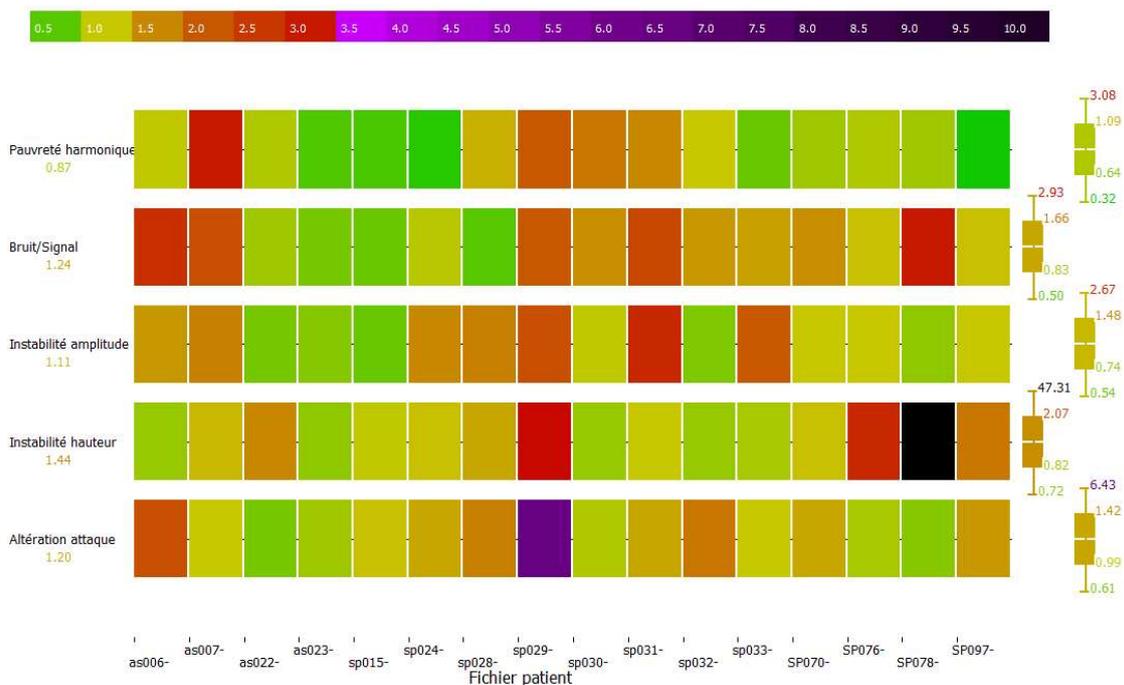


Figure 16: représentation graphique des indicateurs portant sur 11 cas de /a/ avec pathologie de type nodule.

Un problème important concernant ce type d'analyse est la variabilité des conditions d'enregistrement, dans la mesure où 2 orthophonistes différentes ont contribué à enrichir la base de données de cas de nodules, avec des équipements informatique et chaîne d'acquisition différents. Les niveaux de bruit, niveaux de puissance et consignes peuvent être disparates, réduisant d'autant le degré de confiance que l'on peut accorder aux résultats. Les calculs sont bien plus probants si le matériel, logiciel, lieu et protocole sont identiques pour l'ensemble des patients.

## Etude de cas – Voix Cancer

Dans le cadre du projet C2SI coordonné par le Docteur Woisard de l'IUCT (Sicard, 2016), nous avons pu analyser plusieurs dizaines de /a/ de patients ayant été atteints de cancer de la sphère ORL, de différents grades de sévérité allant de T2N0M0 à T3N2bM0, selon la classification TNM de Sobin (2009).

Les enregistrements sont effectués au moins 6 mois après la fin du traitement, avec un protocole, matériel, logiciel et réglages identiques, limitant la variabilité liée aux conditions d'enregistrement. Comme le montre la répartition âges/F0 de la figure 17, les cas étudiés sont principalement dans la tranche 55-75 ans, avec 17 cas d'hommes et 7 de femmes. La synthèse des données (Figure 18) montre une élévation significative de tous les indicateurs, à l'exception de l'instabilité en amplitude. L'indicateur le plus élevé est la pauvreté harmonique. Ces résultats recourent les analyses de Verdolini (2006) et Osta (2010).

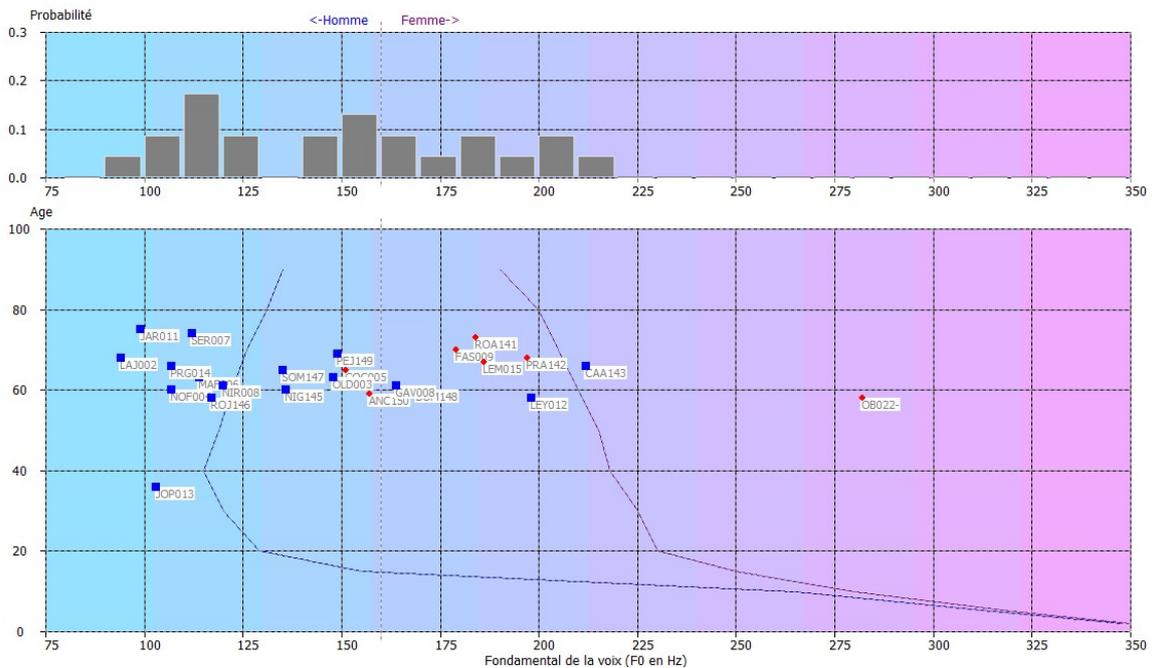


Figure 17: L'âge des patients est principalement dans la tranche 55-75 ans, avec 17 hommes et 7 femmes (Sicard, 2016)

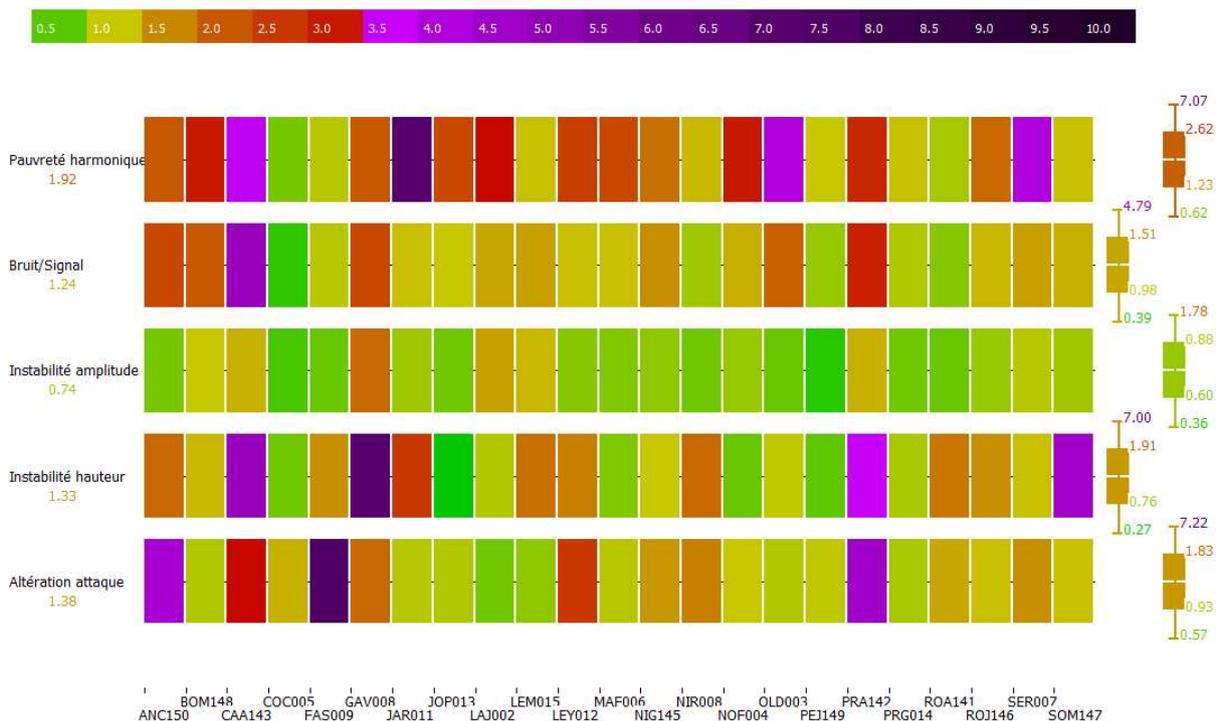


Figure 18: Augmentation significative des indicateurs sur 24 cas de /a/ de patients atteints de cancer ORL (Sicard, 2016)

## Etude de cas – Voix âgées

Un ensemble de 70 cas de /a/ de voix d'hommes et de femmes âgées de 45 à 90 ans a été étudié dans le cadre du mémoire d'orthophonie d'Océane Barbera et Camille Fresnay (2016). L'âge du groupe étudié est réparti de 45 à 95 ans (Figure 19). Un des objectifs du mémoire était de dégager des tendances en termes d'élévation des indicateurs d'altération de la voix au-delà de 60 ans.

Un microphone Shure SM-58 directionnel, une carte-son externe AVID reliée à l'ordinateur par port USB ont été utilisés pour l'ensemble des cas. Les enregistrements ont été réalisés dans une pièce calme, à l'aide du logiciel VOCALAB. La distance bouche-micro était de l'ordre de 10cm pour tous les participants, grâce à une réglette, et le protocole comportait 6 types de production sonore, dont les /a/ tenus étudiés ci-après. On peut donc ici considérer que l'incertitude liée aux conditions expérimentales est faible, donc le niveau de confiance dans les résultats obtenus élevé.

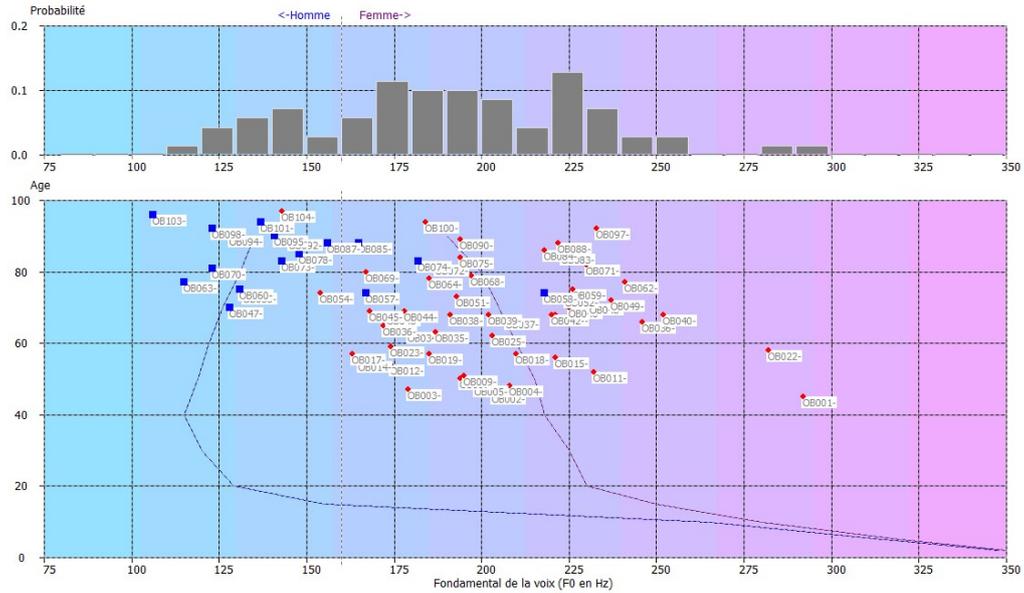


Figure 19: Etude de 70 cas de voix âgées, répartition du /a/ des cas étudiés (Barbera & Fresnay, 2016)

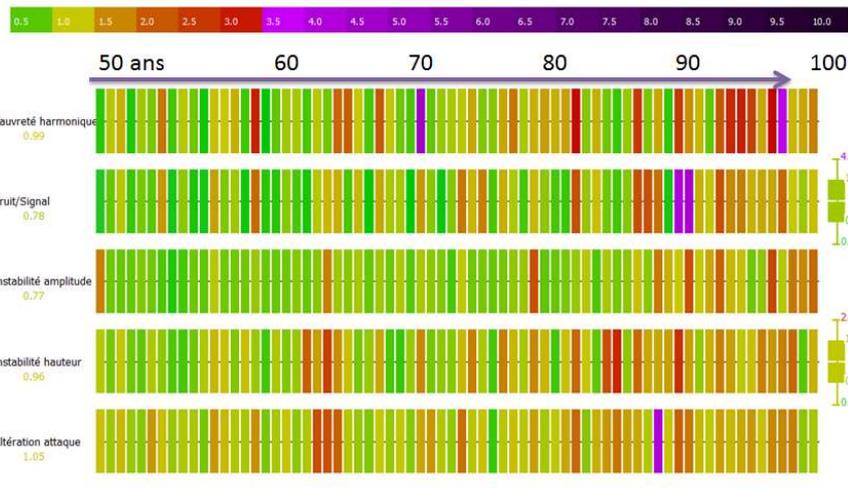


Figure 20: Evolution des indicateurs avec l'âge, pour des /a/ de voix normales. On note une « coloration » de l'ensemble des indicateurs avec l'âge (Barbera & Fresnay, 2016)

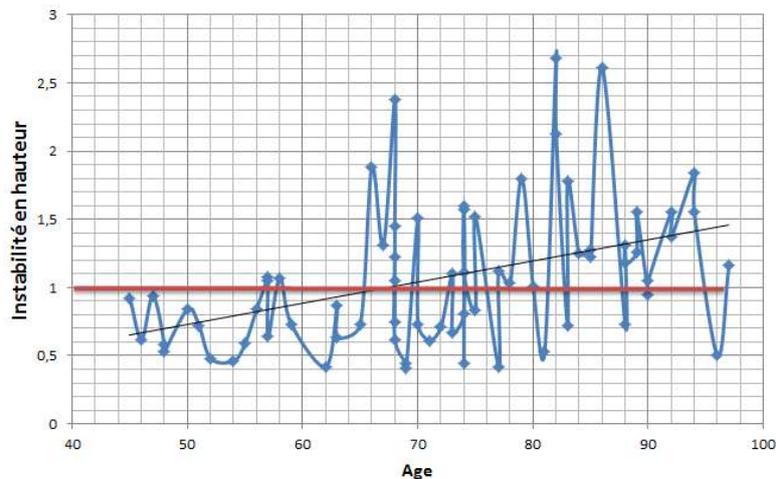


Figure 21: Evolution de l'instabilité en hauteur avec l'âge pour des voix normales et courbe de régression obtenue avec EXCEL montrant une pente autour de 0.2 points tous les 10 ans (Barbera & Fresnay, 2016)

Les noms de fichier ont été définis de telle sorte que le classement alphabétique corresponde à l'augmentation de l'âge (Figure 19). Par contre, hommes et femmes sont mixés. L'étude citée porte sur un total de 44 hommes et 56 femmes. On note une coloration des cas de /a/ en voix normale à partir de 65 ans sur la figure 20, tendance confirmée avec une courbe de régression linéaire de la figure 21. Ce graphe ne distingue pas les hommes et les femmes

Les résultats des calculs issus de VOCALAB ont été exportés sous EXCEL™ à l'aide du bouton « **Ecrire CSV** » en vue de déterminer une loi approximative d'élévation des valeurs seuil (Figure 21). On peut noter une tendance à l'augmentation, avec un seuil qui s'élève d'environ 0,2 points par tranche de 10 ans. Des résultats similaires sont observés pour les quatre autres indicateurs. Cette tendance a été modélisée dans VOCALAB avec un « recoloriage » des seuils de couleur de manière à s'adapter à l'âge du patient. Par exemple, une instabilité en hauteur autour de 1.0 est colorée par défaut en jaune, mais passe au vert si le patient est âgé de plus de 75 ans.

### Age non précisé ou âge <60 ans

- ▶ Valeur < 0.8 : **vert**
- ▶ 0.8 < Valeur < 1.5 : **orange**
- ▶ Valeur > 1.5 : **Rouge**

### Age > 60 ans

- ▶ Seuil **vert/ orange** augmente avec l'âge
- ▶ Seuil **orange / Rouge** augmente avec l'âge
- ▶ Augmentation de **0,1 à 0,25/10 ans** selon les indicateurs
- ▶ Taux d'élévation de **0,2/10 ans** implémenté dans VOCALAB

## Etude de cas - Presbyphonie

Dans le cadre du mémoire d'orthophonie d'Océane Barbera et Camille Fresnay (2016), 6 cas de presbyphonie (4 hommes, 2 femmes) ont été insérés dans la base de données en ligne. On note sur le graphique de la figure 22 des valeurs d'indicateurs particulièrement élevées concernant la pauvreté harmonique, le bruit/signal et l'instabilité en hauteur.

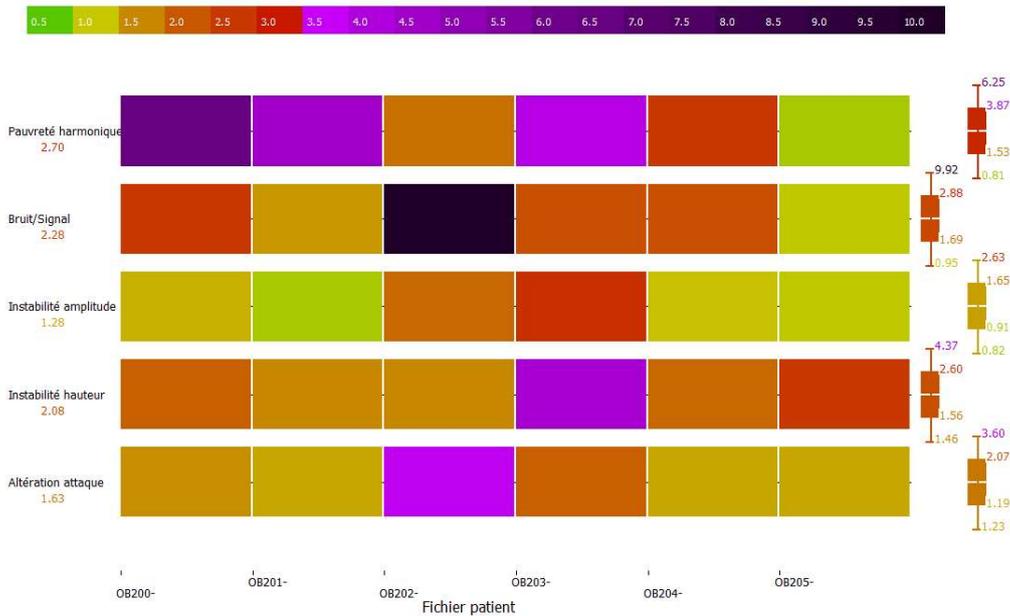


Figure 22: Augmentation significative des indicateurs sur 6 cas de /a/ de patients atteints de presbyphonie /RefOC/

## Corrélation entre indicateurs

Nous pouvons évaluer la corrélation entre deux indicateurs, comme par exemple la pauvreté harmonique, dont la série de valeurs est reportée en X, et l'instabilité en hauteur, dont la valeur correspondante est dessinée en Y (Figure 23). On observe un nuage de points qui occupe l'espace X,Y sans dégager de tendance claire. Le coefficient de corrélation est proche de 0, ce qui indique que les variables traitent de phénomènes différents.

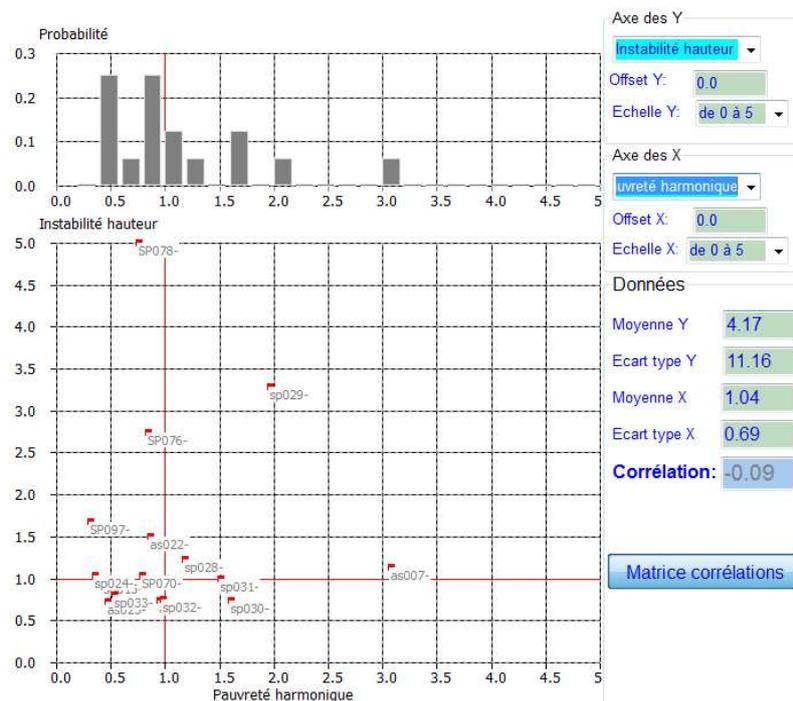


Figure 23: Un coefficient de corrélation proche de 0 indique que les deux grandeurs sélectionnées évaluent des phénomènes différents.

Les indicateurs sur le son entier sont assez fortement corrélés au son sur 1 seconde après l'attaque, notamment pour le bruit/signal et la pauvreté harmonique. En d'autres termes, un son de longue durée n'apporte que peu d'informations supplémentaires sur les notions de bruit et de pauvreté harmonique (Figure 20). Par contre, les résultats en instabilité en amplitude et en hauteur peuvent différer selon que le calcul porte sur une portion de son après l'attaque, où le patient est en général en zone de confort, ou sur la totalité du son /a/. En effet, divers phénomènes d'instabilité long terme et de terminaison sonore peuvent se cumuler à l'instabilité moyen terme, dégradant presque systématiquement les indicateurs d'altération de la voix.

Il est possible de calculer la matrice de corrélation entre l'ensemble des variables, en cliquant le bouton « **Matrice corrélations** ». VOCALAB est capable de passer au crible l'ensemble des couples de variables et calcule le coefficient de corrélation associé. Le fichier de sortie est un tableau dont la diagonale donne la valeur 1.0 (Figure 24). En d'autres termes, une série de données ressemble toujours à elle-même.

Pour des séries de données différentes, plus le coefficient est bas, moins les données sont corrélées. C'est le cas par exemple de l'attaque et l'instabilité en hauteur (-0.11), ou encore l'instabilité en hauteur et la pauvreté harmonique (-0.09). On constate en revanche une corrélation non négligeable entre bruit/signal et pauvreté harmonique (0.44), ou encore instabilité en hauteur et bruit/signal (0.54).

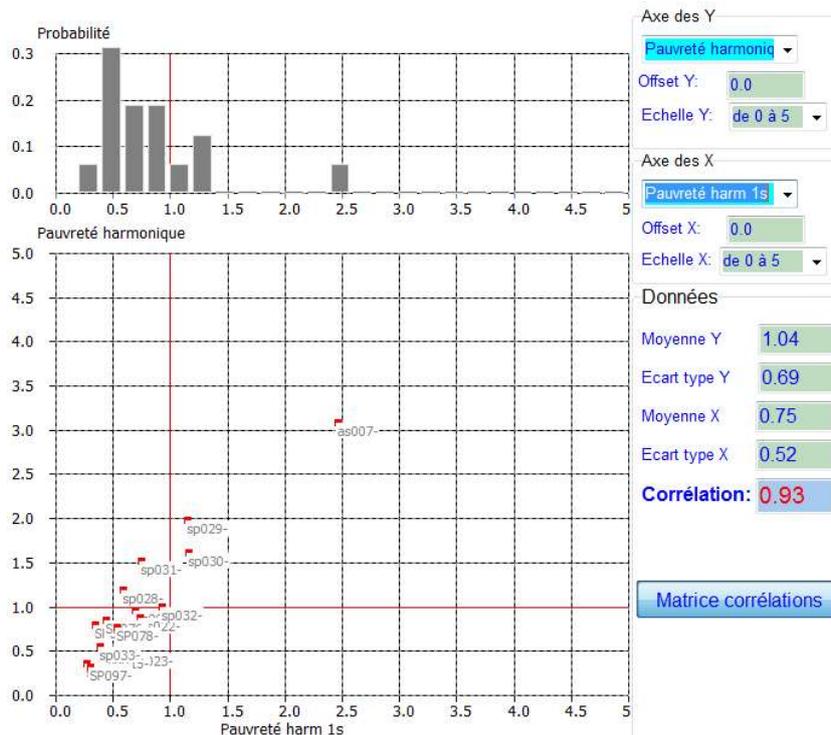


Figure 24: Un coefficient de corrélation proche de 1 ou -1 indique que les deux grandeurs sélectionnées évaluent des phénomènes similaires (ici la pauvreté harmonique globale, et celle observée sur 1 seconde après l'attaque).

	Attaque	Instabilité hauteur	Instabilité Amplitude	Bruit/Signal	Pauvreté harmonique
Attaque	1.00	-0.11	0.44	0.24	0.32
Instabilité hauteur		1.00	-0.22	0.54	-0.09
Instabilité Amplitude			1.00	0.35	0.43
Bruit/Signal				1.00	0.44
Pauvreté harmonique					1.00

Figure 21 : Matrice de corrélation entre indicateurs

## Analyse long terme

L'analyse long terme correspond au calcul de la moyenne statistique du spectrogramme. C'est une représentation amplitude/fréquence qui permet d'identifier les zones d'énergie importantes, et par différence, identifier les caractéristiques fréquentielles d'un groupe de voix données.

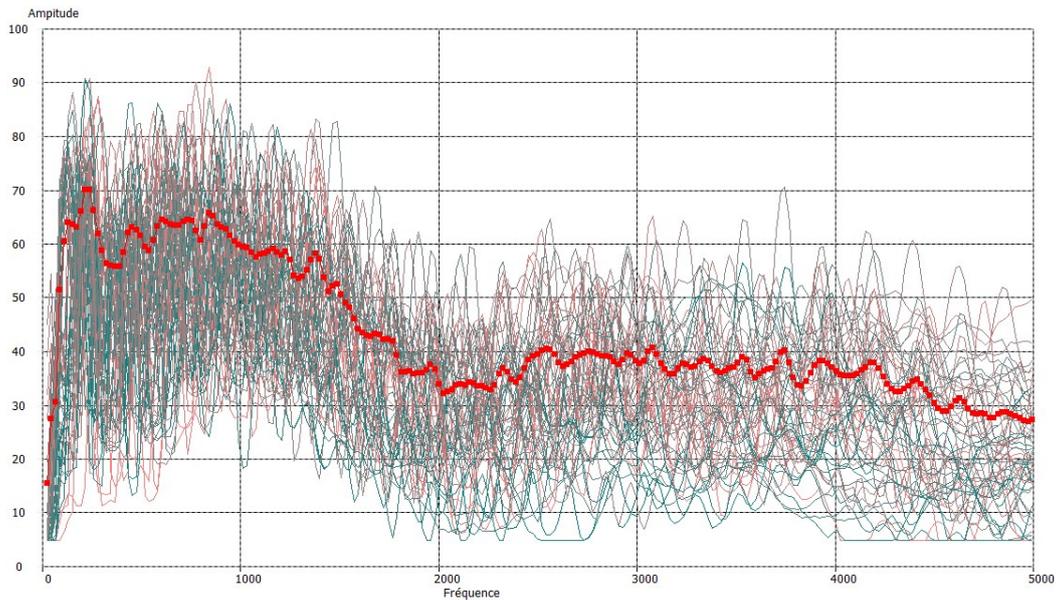


Figure 25 : Analyse « Long Terme » du /a/ de 60 cas de voix normales

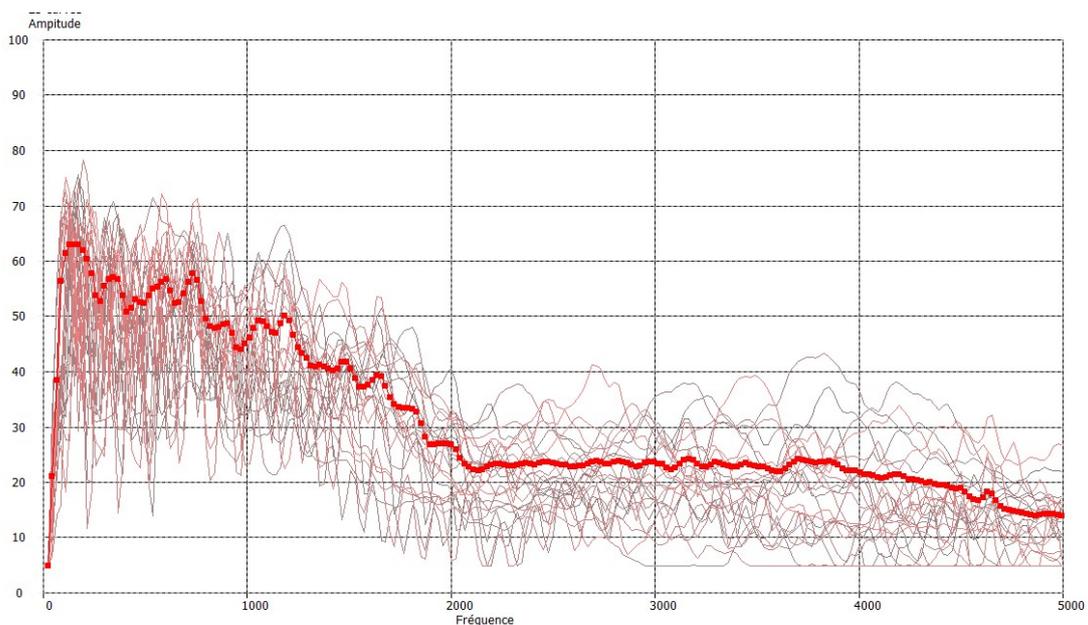


Figure 26 : Analyse « Long Terme » du /a/ de 25 cas de voix de patients atteints de cancers ORL

Chaque voix de patient est moyennée sur l'ensemble du spectrogramme et donne lieu à une courbe grise. La moyenne de ces courbes est dessinée en rouge. Elle est accessible par l'outil **Médiathèque > Sons > Statistiques**, onglet « **Long Terme** ». On peut noter une énergie significative de 100 à 1000Hz, une décroissance rapide jusqu'à 2000Hz, puis un plateau jusqu'à 4500 Hz environ.

En comparant les analyses long terme des témoins (Figure 25) et des patients ayant été traités pour des cancers ORL (figure 26), nous pouvons dégager une tendance à la décroissance rapide

de l'énergie dès 500 Hz, et un niveau très bas de toute la zone 2000-4500Hz, ce qui s'apparente à une pauvreté harmonique significative.

## Conclusion

---

Nous avons donné dans ce chapitre quelques rappels théoriques de calcul statistique, que nous avons appliqué à l'analyse de la voix. Nous avons aussi présenté l'outil de calcul statistique implémenté dans VOCALAB pour dégager les tendances principales des voix normales et pathologiques, à des fins de comparaison. Ces outils ont été utilisés dans le cadre de mémoires d'orthophonie dont quelques résultats ont été présentés. Ces outils implémentés dans Vocalab permettent de mieux comprendre l'influence des pathologies organiques ou fonctionnelles sur les paramètres de la voix et d'aider les orthophonistes à déterminer précisément des axes thérapeutiques spécifiques. Les outils objectifs favorisent une démarche scientifique sans pour autant négliger l'aspect clinique et la prise en compte de paramètres subjectifs tels que l'aspect émotionnel et psychique du patient.

A l'heure où la reconnaissance de la profession, de son niveau de qualification est remis en question, il est crucial de se doter d'outils scientifiques, objectifs utilisés par un nombre significatif de cliniciens afin d'en valider les fonctionnalités et les résultats.