



Italie

Il Laboratorio



Rapport d'intervention au congrès de Paris le 13 mars 2007

Cette intervention a été réalisée par **Daniele Gulla**, expert en analyse de voix et visages auprès de la cour de justice en Italie. La traduction a été effectuée par **Nicole Garnier**.



Daniele Gulla et Nicole Garnier

Il s'agit bien sûr de technique, mais il nous a semblé intéressant, peut-être plus spécialement pour ceux qui font de la recherche, de vous communiquer ce texte, d'autant plus que nous l'avons agrémenté de schémas d'appui pour certains aspects pouvant avoir un intérêt général.

Amélioration du signal vocal, traitement, analyses, comparaisons

1. Enregistrement et amélioration de la voix.

Pour avoir une bonne qualité sonore d'écoute des voix, soit normales soit "anormales" et pour en faire l'analyse électro-acoustiques, il faut partir d'un signal d'enregistrement avec un léger rapport Signal/Bruit (S/B) moins de 3/4 dB (décibel).

Si le signal à utiliser est au niveau du bruit de fond, voire inférieur, les possibilités de réussite seront extrêmement basses.

La technique qui permet d'obtenir un enregistrement de départ avec un bon rapport S/B est sans doute la digitale plutôt que l'analogique (enregistrements normaux sur des audiocassettes), obtenue en enregistrant directement sur l'ordinateur (avec le matériel compatible) ou sur les supports magnétiques DAT (Digital Audio Tape) ou Mini Disc.

L'étendue de la dynamique acoustique est de l'ordre de 90 dB et au-delà, pour le système digital contre environ 60 dB pour l'analogique. La différence substantielle entre le digital et l'analogique est que le premier couvre un échantillonnage périodique de la forme d'une onde discontinue au contraire de la seconde.

Avant de passer à la phase de conversion digitale de l'enregistrement audio pour la restauration et le nettoyage du son, il faut avoir pris soin d'enregistrer "les voix" sur un support qualitativement apte afin de réduire au minimum la perte d'informations acoustiques et donc l'intelligibilité du parlé.

Malheureusement, on doit souvent travailler sur des signaux beaucoup trop faibles et brouillés parce qu'ils n'ont pas été enregistrés correctement. Trop souvent on utilise des microphones d'insuffisante qualité ou des mini enregistreurs à cassette avec un haut niveau de distorsion, ce qui rend les analyses et le décodage peu fiables.

Personnellement je conseille d'enregistrer avec un bon microphone (Sennheiser ou AKG ou Neumann) en employant un magnétophone DAT ou l'ordinateur (qui sont maintenant de plus en plus adoptés par la justice et la police scientifique.)

Différemment, pour qui utilise le système analogique normal (magnétophone standard à

cassettes), on conseille l'emploi d'une bonne marque (Sony, Teac, Philips...) et d'un préamplificateur microphonique en entrée (mieux s'il a un pré-égaliseur de qualité, personnellement je conseille les préamplificateurs Focusrite).

Comme *lignes directrices* reprenons les principales caractéristiques pour effectuer un enregistrement de bonne qualité, valable aussi bien pour le domaine judiciaire que pour un enregistrement microphonique optimal en EVP :

- 1 - Acquisition du signal
- 2 - Le microphone.
- 3 - L'enregistreur magnétique
- 4 - L'enregistrement digital.

2. Acquisition du signal.

La précision des analyses spectrales du signal dépend essentiellement de la qualité de l'acquisition et donc du rapport Signal/Bruit déjà cité.

Les moyens qui peuvent influencer la qualité du signal à acquérir sont : Le microphone, l'enregistreur, la bande magnétique, la vitesse d'enregistrement, les appareils automatiques d'acquisition comme le Trevisan ou un activateur vocal (enclenchement par la voix - VAS).

En outre la qualité du signal dépend aussi de la réverbération et des bruits contenus dans l'ambiance, la fréquence d'échantillonnage (en anglais : *Sampling rate*), la qualité des convertisseurs Analogiques et Digitaux (A/D) ou (D/A), le lieu de conservation des enregistrements.

2.1 Le microphone.

Le microphone transforme l'énergie sonore en énergie électrique. La qualité et la fiabilité d'un microphone se mesurent en tenant compte de quelques caractéristiques telles que la courbe de réponse, la directivité, la sensibilité, la distorsion, le bruit intrinsèque et la dynamique.

La courbe de réponse indique la régularité de l'amplitude du signal électrique que le microphone fournit en sortie, en fonction de la fréquence du signal acoustique en entrée. Il indique en fait l'intervalle de fréquence qu'il est possible d'enregistrer.

La directivité : le microphone peut être **omnidirectionnel** et enregistrer le signal quelle que soit sa provenance.



Il peut être **bidirectionnel** et enregistrer ainsi le signal d'au moins deux directions précises.



Enfin il peut être **unidirectionnel** et enregistrer le signal seulement d'une direction.



La sensibilité est la capacité d'un microphone à capter de faibles signaux sonores. Il est évident que plus un microphone est sensible, mieux il réussit à enregistrer des sources sonores de basse intensité mais ceci n'est pas toujours un avantage en TCI.

La distorsion se réfère à la déformation induite au cours de la transformation de l'onde sonore en onde électrique ; c'est la mesure du rapport de similitude entre l'onde originale et celle enregistrée. Un microphone ne doit pas être trop sensible autrement il introduit, ou enregistre, le plus petit bruit.

La dynamique, c'est l'intervalle délimité par le niveau du bruit de fond (c'est-à-dire la sensibilité) et le niveau auquel s'introduit la distorsion ; elle est exprimée en dB.

Enfin il faut tenir compte des influences extérieures comme la pression atmosphérique, la température, le degré d'humidité, paramètres qui peuvent influencer la membrane du micro et donc ses caractéristiques intrinsèques.

2.2 Le magnétophone.

Le magnétophone est l'appareil qui traduit le signal électrique provenant d'un microphone, en variations magnétiques sur la bande (les variations magnétiques sont dues à des particules de fer ou de charbon).

Les paramètres importants pour la mesure de la qualité d'un enregistreur sont multiples. Voici les plus importants :

2.2.1 La vitesse de défilement de la bande.

Plus elle est rapide, plus elle enregistre d'informations. Celles que l'on utilise pour l'usage domestique sont de 2.38 cm/s et 4.75 cm/s. Les semi professionnelles vont de 9.52 cm/s à 19.05 cm/s.

Les professionnelles ont des vitesses de 38.1 cm/s à 76.2 cm/s. En plus d'avoir d'une vitesse de déroulement adaptée, un bon enregistreur doit garantir la stabilité, puisqu'il est notoire que ses oscillations causent des variations dans le ton de la voix.

2.2.2 Réponse globale de fréquence (à travers la bande).

Le niveau d'un signal sinusoïdal se mesure en fonction de la fréquence dans la gamme 20 Hz - 20 KHz.

2.2.3 Dynamique.

La dynamique est normalement mesurée en Décibel (dB). Elle correspond au rapport entre le niveau le plus élevé d'un signal enregistré et le niveau du bruit intrinsèque.

La dynamique est directement liée à la vitesse de défilement du ruban et au nombre de pistes d'enregistrement utilisables. Plus la vitesse d'enregistrement sera importante et meilleure sera la dynamique.

La dynamique des enregistreurs professionnels ou semi professionnels (76.2 cm/s - 38.1 cm/s), avec des bande de $\frac{1}{4}$ de

pouce, est de 65 dB pour une piste et 40 dB pour quatre pistes.

2.2.4 Distorsion.

La distorsion se mesure à une fréquence donnée pour déterminer les niveaux d'enregistrement et de reproduction. Très souvent les EVP sont le résultat d'un signal (bruit fréquemment) complexe avec un certain degré de distorsion.

2.2.5 la bande magnétique.

L'épaisseur de la bande magnétique utilisée pendant un enregistrement détermine les caractéristiques comme la force, la flexibilité et l'élasticité. Sa largeur est habituellement normalisée à 6.25 mm.

Il existe au moins quatre types de bande magnétique : le Standard ou ST environ de 50 micron, le Longue Durée ou LP environ de 35 micron, à Double Durée ou DP environ 25 micron et de celui à Triple Durée ou TD environ de 18 micron. La bande standard a une grande résistance mécanique et est utilisable même en conditions climatiques différentes, donc elle est plus conseillée.

Les caractéristiques électromagnétiques d'une bande sont très importantes. Il faut par exemple connaître le bruit de fond, ou le niveau maximum d'écho magnétique.

L'effet écho, par exemple, se présente après une certaine période de conservation de la bande lorsque celle-ci est enroulée dans sa bobine : le signal enregistré sur une piste peut, si la bobine n'est pas conservée comme il faut, se recopier sur les pistes adjacentes. Un tel effet est causé par la température et aussi par le temps même de conservation.

2.2.6 L'enregistrement digital.

La conversion d'un signal analogique en signal digital (A/D) se produit à travers une transformation de l'onde en une représentation d'un code binaire qui peut être stocké sur support DAT ou un ordinateur pour être exploité.

La fréquence d'échantillonnage est dite *Sampling Rate*. Celle-ci est réalisée sur un enregistreur digital

ou directement sur l'ordinateur à travers la prise d'acquisition audio.

Par fréquence d'échantillonnage il faut voir le nombre d'échantillons considérés dans l'unité de temps. Le nombre de bit (Binary Digit) se réfère par contre au nombre de différences considérées sur l'axe d'amplitude de chaque échantillon.

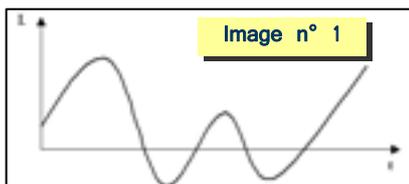
Puisque chaque valeur d'amplitude est représentée par un nombre de bit, le nombre total des échantillons aura une puissance de 2. Une conversion à 16 bit (optimal) sera équivalente par conséquent à un nombre d'échantillons pairs à 216.

Selon le théorème de Nyquist, pour coder digitalement la bande de fréquence désirée, la fréquence d'échantillonnage doit être au moins égale à la puissance deux. On en déduit qu'un signal audio avec une bande passante de 20 KHz demandera une fréquence d'échantillonnage d'au moins 40.000 échantillons à la seconde.

Par exemple, on choisit une fréquence d'échantillonnage de 44.1 KHz pour correctement coder une bande passante effective de 20 KHz.

Voyons maintenant le détail des aspects théoriques et techniques.

L'image n° 1 représente un signal analogique de variation continue en fonction du temps t.



En ordonnée, le niveau L pour les signaux acoustiques est en principe exprimé en décibel (dB).

L'image n° 2 représente la conversion digitale du précédent signal analogique. Il en résulte une séquence de segments et non plus une variation continue. Le nombre de segments représentés dans un temps T équivaut à la fréquence



d'échantillonnage.

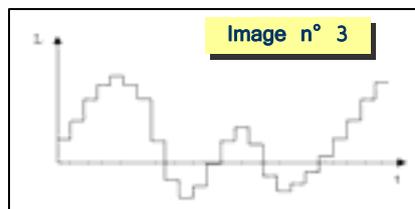
Une fois le signal segmenté dans le temps, ce qui est très simple parce qu'il suffit d'avoir un minuteur qui scande les instants où effectuer la lecture, il faut le segmenter en amplitude.

Pour faire cela, la valeur est envoyée en entrée vers un circuit appelé convertisseur analogique - digital (A/D) qui produit en sortie une séquence de bit. Pour chaque échantillon du convertisseur sont émis un certain nombre de bit qui en représentent la résolution.

Les valeurs typiques de résolution d'un convertisseur A/D sont de 16, 20 ou 24 bit. Par exemple 16 est le nombre de bit utilisé par les lecteurs CD, 20 ceux utilisés par enregistreurs digitaux, 24 ceux utilisés par les lecteurs DVD.

Plus la résolution est précise, plus les segments le seront ; en effet à travers ce dispositif le signal est transformé suivant un certain nombre de niveaux : 16 bit équivalent à 65536 niveaux, 20 bit à 1048576 niveaux et 24 bit à 16777216 niveaux. Une petite erreur inhérente aux représentations du nombre réel sur un nombre fini de chiffres représente la quantification.

Les fréquences d'échantillonnage les plus utilisées sont 44.1 kHz pour les lecteurs CD et pour l'ordinateur, 48 kHz pour les enregistreurs digitaux, 96 kHz pour les lecteurs DVD et 192 kHz pour systèmes professionnels de haute précision.

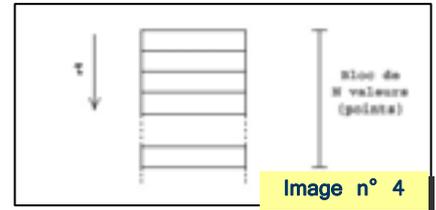


En reprenant la largeur des niveaux du signal échantillonné (image n° 3) on obtient l'intervalle d'échantillonnage dans le temps alors que la hauteur des niveaux est donnée par la résolution du convertisseur A/D.

Il est clair que si on augmente la résolution sur l'amplitude avec un meilleur convertisseur A/D, on doit augmenter la résolution sur la fréquence. Il n'y a par exemple,

pas de sens à avoir une résolution de 24 bit sous échantillonnée à travers une fréquence de 44.1 kHz. Lorsque les niveaux ne sont pas équilibrés entre eux cela crée des "images sonores flottantes".

En sortie du convertisseur le son est représenté par une suite de nombres qui sont mémorisés en cellules comme représenté ci-dessous (image n° 4).

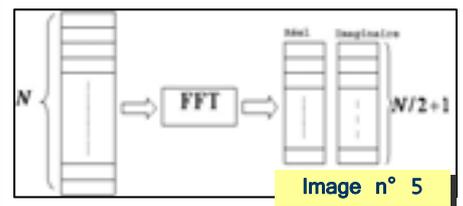


A partir de cette suite de données on pourra élaborer le spectre du signal.

3. Élaboration FFT.

L'algorithme mathématique qui calcule le spectre à partir des échantillons quantitatifs s'appelle FFT (Fast Fourier Transform ou Transformation Rapide de Fourier, mathématicien français. La transformée de Fourier analyse le contenu fréquentiel d'un signal).

Il agit sur N échantillons réels dans le temps en entrée, tout en transformant en N/2+1 les échantillons des complexes en fréquence avec un intervalle constant, chacun étant constitué d'une partie réelle et d'une partie imaginaire (image n° 5).



L'algorithme produit alors une suite linéaire sur l'axe des fréquences. Des représentations de ce type attribuent la même importance à toutes les fréquences aussi bien basses que hautes.

4. Post-traitement

Une fois obtenu le spectre de bande étroite on réalise le post-traitement, c'est-à-dire l'élaboration qui succède au spectre et permet de changer l'axe des fréquences linéaires en tiers

d'octave. On produit en même temps une constante de temps rapide ou lente de façon à ne pas avoir trop de spectres, ni conserver de mémoire précédente, mais de les répartir dans le temps afin de visualiser les bandes qui varient de manière plus souple.

5. Nettoyage et amélioration du son vocal.

Un des modes les plus simples pour améliorer un son vocal ou musical est celui de le compresser en bandes de fréquence déterminée et d'en contrôler la dynamique avec des multiprocesseurs multi bandes et des égaliseurs paramétriques de qualité.

Dans ce domaine, il existe beaucoup de matériels et/ou de logiciels à petit prix sur le marché, même si, évidemment, ceux qui donnent les effets désirés ne sont jamais les moins chers.

Récemment on a réalisé beaucoup d'avancées dans l'élaboration digitale du signal (DSP), qui peuvent réduire le bruit de fond présent dans des séquences sonores préenregistrées, (qui peut comprendre aussi bien des effets collatéraux de type analogique, comme par exemple des bruissement, cliquetis, raps, pops et même quelques types de distorsion présents dans le signal même du processeur).

Quelques algorithmes se sont révélés en mesure d'atténuer les effets de *granulosité* d'un enregistrement effectué en basse résolution (comme par exemple 8 ou 12 bit en digital), voire d'éliminer presque totalement un bruit de fond ambiant envahissant jusqu'au total nettoyage et la restauration de l'enregistrement vocal.

Voici maintenant un exemple de restauration d'un enregistrement vocal avec la *technique mixte de type segmental* avec l'amélioration des harmoniques et des formants (voir **image n° 6** (enregistrement original) et **image n° 7** (restauré et amélioré)).

Les méthodologies utilisables sont nombreuses et diversifiées autant en

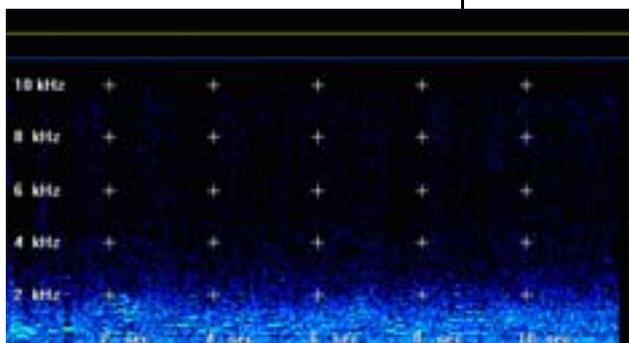


Image n° 6 - Voix originale masquée par des bruits de fond excessifs

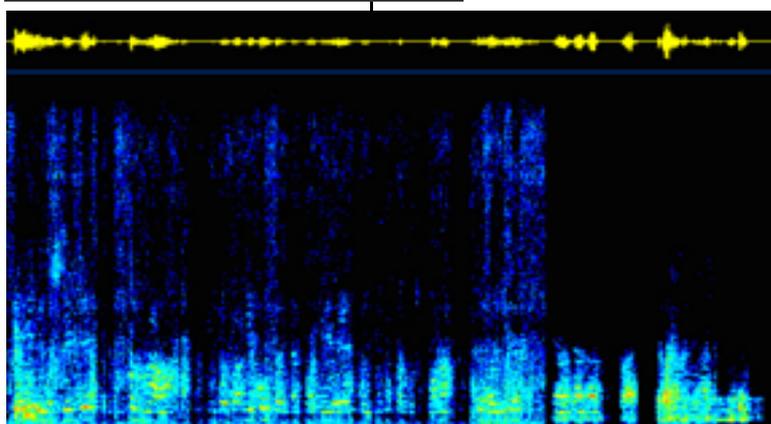


Image n° 7 - Voix après nettoyage et amélioration

Résultat après reconstruction segmentale du signal

termes d'investissements qu'en complexité de calculs mathématiques opérationnels.

Dans les enregistrements radiophoniques ou téléphoniques, la bande passante étant moins de 4 kHz, il est absurde d'utiliser des conversions à 192 kHz ; il suffit de 44.1 kHz et 16 bit, pour avoir un signal de haute qualité.

Dans le cas de travaux sur des voix microphoniques obtenues par enregistrement ambiant par contre, est conseillé d'échantillonner au moins à 48 kHz/96 kHz et 24 bit, et à 192 kHz seulement si les signaux sont *extrêmement faibles* ou presque au niveau du bruit de fond.

Souvent les présumées voix anormales (que nous rebaptisons *voix I.P.S. = Interazione Psi Strumentale* ou Voix résultant d'une interaction psi sur les instruments) sont, probablement par leurs caractéristiques ou par les variations du canal de communication, d'insuffisante qualité spectrale et avec beaucoup de bruit.

Lors de communications reçues à des moments différents, par des

expérimentateurs et appareillages différents, les mêmes entités communicantes produisent des spectres relativement différents, caractérisés par une inhabituelle fluctuation dans la dominance des fréquences et des amplitudes, et surtout dans la dominante du temps.

Dans les observations réalisées, on a noté des variations toujours plus fréquentes sur l'amplitude du signal (modulation d'amplitude) avec un maintien presque constant, voire des fluctuations très légères, de la fréquence. En conséquence, on utilise des algorithmes de calcul et des égaliseurs paramétriques adaptés à ces types de signaux.

6. Analyses et comparaison des caractéristiques des voix paranormales

Une fois que nous avons nettoyé et amélioré le signal audio contenant la voix qui nous intéresse, il faut procéder aux analyses électroacoustiques pour vérifier qu'il s'agit vraiment d'une voix paranormale et non d'un bruit ou d'une voix normale.

Les analyses principales qui sont faites avec l'ordinateur peuvent se

servir de programmes plus ou moins sophistiqués et coûteux et de programmes gratuits (*freeware*) qui sont souvent mis à disposition pour les étudiants dans diverses Universités.

Nous conseillons SFS (Speech Filing System ou le Praat) qui sont complètement gratuits et disposent de divers outils similaires pour traiter les voix paranormales. Tous les deux sont chargeables gratuitement sur Internet.

On peut faire sur les voix des analyses multiples, mais les principales, celles qui peuvent nous fournir des informations utiles sur la nature de la voix sont les suivantes :

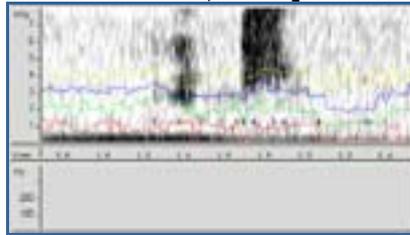
- Vérifier la présence de la Fréquence Fondamentale (F0) ;
- Vérifier la présence de vibration des cordes vocales (analyses sonographiques) ;
- Vérifier la présence et la constitution des Formants (F1, F2, F3, Fn) ;
- Extrapoler toutes les valeurs précédentes avec mesure de la Déviation Standard (SD) ;
- Comparer les valeurs extrapolées avec celles d'un échantillon de population contenu dans une base de données ;
- Vérifier l'évolution temporelle du signal (Modulation en amplitude et fréquence) ;
- Vérifier le Bruit et l'Énergie ;
- Vérifier les perturbations de tonalité ou les tremblements de la voix.

Avec le programme SFS et les divers outils, y compris la base de données MBROLA qui incorpore des voyelles en diverses langues, on peut commencer à exécuter des analyses qui peuvent mettre en évidence si la voix que nous avons enregistrée est classée en tant que voix humaine normale ou comme probable voix paranormale.

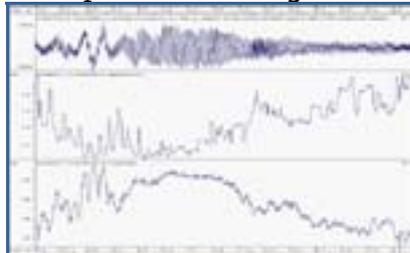
En synthèse, voici un exemple d'analyse sur une voix enregistrée pendant une expérimentation exécutée avec le Projet SFINGE par Le Laboratoire à Bologne.

Analyse de rapport sonore pour le message :
"TU CE L'HAI IL SANGUE"
 Français : "C'est toi qui a le sang"

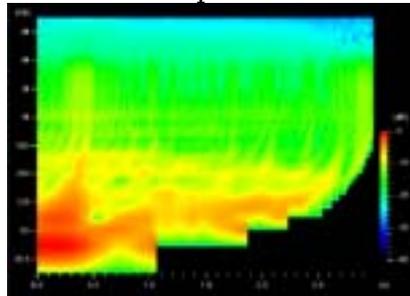
a) Analyse sonore de F0 (fréquence fondamentale) de la phrase.



b) Analyse du bruit et de la périodicité du signal

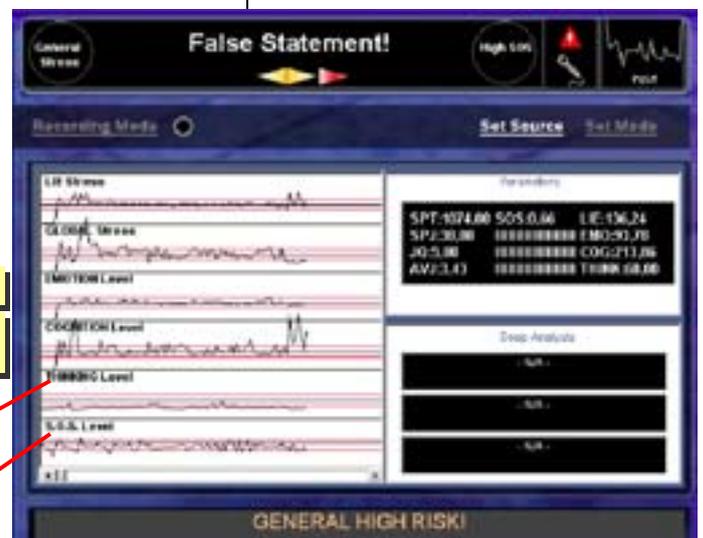


c) Analyse de la réverbération et du spectre



d) Analyse du tremblement vocal

- Niveau de tension
- Niveau d'émotion
- Niveau de reconnaissance
- Niveau de compréhension
- Niveau d'alerte



Conclusion

La brève relation exposée contient une synthèse théorique et méthodologique, avec quelques suggestions, pour qui veut pénétrer dans le difficile secteur judiciaire de l'acoustique. Ces quelques lignes sont un guide complémentaire utile soit pour les analyses concernant les voix humaines, soit pour les essais

acoustiques provenant d'expériences "psychophoniques".

Pour ces derniers, ayant relevé beaucoup de variantes dans mes expériences d'étude et de classement phonétique, je peux ajouter que les méthodes standard ne sont pas toujours applicables, mais que souvent il faut appliquer un filtrage manuel approprié au répertoire acoustique, tant celui-ci peut avoir ses propres caractéristiques uniques.

De ces considérations on déduit que, outre les systèmes automatiques et semi-automatiques d'amélioration acoustique sur le signal, il faut pouvoir disposer de systèmes paramétrables de contrôle de façon à les adapter manuellement au son particulier qu'on examine.

Cette opération ne peut pas avoir lieu sans une préparation technique en électroacoustique de l'opérateur qui souvent est le même que l'expérimentateur en Tci.

Notre Laboratoire de Recherche, en plus d'offrir une série de cours et d'approfondissements en technique et de parapsychologie,

met à disposition des associés, des consultations gratuites d'experts dans les secteurs les plus variés : informatique, électroacoustique, graphique, neurophysiologie, médecine, psychologie, théologie, biologie, physique, parapsychologie, etc., s'offrant comme structure d'utilité publique à disposition d'associations ou du privé.