

## RÉSUMÉ:

*Cet article décrit l'utilisation en orthophonie de VOCALAB, logiciel de visualisation spectrale de la parole et de la voix. Outil simple et convivial, Vocalab permet de réaliser des mesures objectives nécessaires au bilan vocal et de visualiser en temps réel la structure harmonique de la voix et de la parole en rééducation. Après un rappel des objectifs du projet et un bref historique du développement du logiciel, nous décrivons les principes de fonctionnement de la chaîne d'acquisition et de traitement du son, puis nous détaillons les modules d'évaluation et de rééducation, ainsi que la médiathèque. Quelques éléments prospectifs sur les futurs développements et applications de cet outil sont donnés en conclusion.*

## MOTS-CLÉS:

Voix — Parole — Informatique — Rééducation — Enfant.

# EVALUATION ET RÉÉDUCATION DE LA VOIX ET DE LA PAROLE AVEC VOCALAB

par Anne MENIN-SICARD, Etienne SICARD

Anne MENIN-SICARD  
Orthophoniste à Toulouse

Etienne SICARD  
Professeur d'électronique, INSA  
DGEI à Toulouse

Adresse: 8 rue de Nîmes,  
31400 Toulouse  
Email: anne-sicard@wanadoo.fr

## SUMMARY: Evaluation and therapy of speech and voice using Vocalab

*This paper described the practical use of Vocalab in speech therapy, for the visualization of voice spectrum. Simple and user-friendly, Vocalab realizes measurements for the vocal evaluation, and displays in real time the harmonic contents of the voice and speech. After a brief description of the project objectives, the sound acquisition and processing are detailed, followed by an in-depth study of the two main modules: evaluation and therapy. Some prospective aspects of the tool development are provided as a conclusion.*

## KEY-WORDS:

Voice — Speech — Computing — Therapy — Children.

L'objectif du projet Vocalab est de mettre à la disposition des orthophonistes un logiciel accessible et fiable rassemblant les outils nécessaires à la visualisation des paramètres de la voix et de la parole. Permettre à l'orthophoniste d'accéder à plus d'objectivité, donc plus de crédibilité aux yeux des prescripteurs et de nos instances supérieures, et homogénéiser nos pratiques en disposant de référents communs rapidement accessibles sont aussi des buts qui s'inscrivent parfaitement dans la mouvance actuelle.

Voir le son sous toutes ses formes est une expérience riche et passionnante. En manipulant l'outil, l'orthophoniste va peu à peu conforter ses connaissances en acoustique et en physiologie, bases théoriques indispensables à la compréhension et l'analyse des données spectrales affichées. Vocalab est conçu pour les assister dans cette démarche objective.

Découvrir sa voix sous la forme d'une trace colorée est aussi une expérience humaine fascinante et ce particulièrement pour les personnes dont la perception auditive est déficiente voire même parfois inexistante. La visualisation joue le rôle de guide-langue visuel et permet au patient, sous les conseils de l'orthophoniste d'ajuster la position de sa langue en fonction du modèle visuel.

Dans le cadre de l'évaluation et de la prise en charge des pathologies vocales, les orthophonistes ressentent de plus en plus l'utilité de visualiser la voix et la parole afin de corréler une analyse auditive subjective avec des données objectives. De nombreux auteurs ont travaillé sur ce sujet et plus particulièrement l'équipe du Pr. Giovanni\* au CHU de la Timone à Marseille en collaboration avec le laboratoire "Parole et langage" du Pr. Teston d'Aix-en-Provence. Pour Buchman\* les mesures des paramètres acoustiques permettent d'établir une sémiologie objective en vue d'un diagnostic mais autorisent aussi des appréciations pronostiques et surtout offrent des possibilités de mesure des effets thérapeutiques. Dalléas\* a étudié l'intérêt de la méthode du spectre moyenné de longue durée (L.T.A.S) et l'analyse du signal temporel qui, corrélé à un jury d'écoute, améliore l'estimation objective des troubles de la voix.

Actuellement, selon une enquête réalisée par Coudière\* sur la région Midi-Pyrénées, une minorité d'orthophonistes sont équipés en logiciel de visualisation spectrale -7 % - pour un taux d'utilisation de l'informatique en rééducation de 68 %. Cependant 80 % des orthophonistes interrogés se disent intéressés par un logiciel permettant de visualiser les paramètres vocaux en temps réel. En effet, l'utilisation des logiciels d'évaluation de la voix est encore trop souvent réservée à des professionnels hospitaliers ou à des orthophonistes chevronnés et spécialisés dans la voix à cause de leur prix élevé et de leur complexité d'utilisation.

Vocalab est né d'une collaboration entre une orthophoniste et un professeur d'infotronique (informatique, électronique). Version informatisée de l'ancien phonaudioscope, il permet de visualiser en 3 dimensions et en temps réel tous les paramètres de la voix et de la parole. L'augmentation constante de la puissance des microprocesseurs a permis de passer d'un traitement du son différé à un traitement en temps réel. Cette possibilité constitue à elle toute seule un progrès considérable pour notre profession et ouvre des chemins d'exploration insoupçonnés. De multiples expérimentations cliniques parallèlement à des efforts continus de programmation ont permis d'enrichir et de faire évoluer le logiciel, l'adaptant ainsi à une pratique libérale efficace. La phase expérimentale a duré huit années pendant lesquelles le logiciel a progressé, pour devenir plus ergonomique et plus fiable. Le passage de l'environnement DOS à Windows a nécessité un travail considérable mais l'ergonomie en a été améliorée jusqu'à ce que les éditions GERIP s'intéressent au produit, donnent un important support pour les animations, le graphisme, et décident d'en assurer la diffusion.

\*1994

\*1991

\*1998

\*2003

Une année a été nécessaire pour adapter le logiciel à l'environnement propre de la société. En 2001, la phase expérimentale a laissé la place à la phase de tests et de diffusion. Le logiciel a été confié à une vingtaine d'experts chargés d'évaluer l'outil et de remplir un questionnaire précis permettant de juger le logiciel en terme de présentation, ergonomie, intérêt et fiabilité des différentes fonctionnalités. L'analyse des résultats s'est faite sur une dizaine de questionnaires retournés. Cette première évaluation a été particulièrement intéressante car elle a permis d'améliorer encore la fiabilité et la cohérence du logiciel.

L'architecture du logiciel comporte 3 grands modules : **Evaluation, Rééducation, Médiathèque**. Le module **Evaluation** rassemble les outils nécessaires au bilan vocal, le module **Rééducation** offre quatre outils utiles à la prise en charge de la voix mais aussi des troubles d'articulation, retard de parole et troubles de type perceptif auditif. Enfin, le module **Médiathèque** comprend un ensemble de fichiers de sons et de vidéos. Après une description du calcul du spectre, nous décrivons successivement ces trois modules.

## CALCUL DU SPECTRE DU SON

Le fonctionnement du logiciel est basé sur le principe de l'acquisition et la décomposition du son en séries de Fourier. Le son est capté par le microphone sous la forme d'un signal analogique. Ce signal est une tension variable en fonction du temps. Une portion de signal de quelques millisecondes est représentée *figure 1 à gauche*. Le signal est converti en une série de valeurs, qui sont stockées dans la mémoire de l'ordinateur. La conversion du format analogique au format numérique (informations de type 0/1) est effectuée grâce à une carte de conversion analogique/numérique existant dans tous les ordinateurs. Le programme Vocalab découpe le signal à la cadence de 11000 échantillons par seconde.

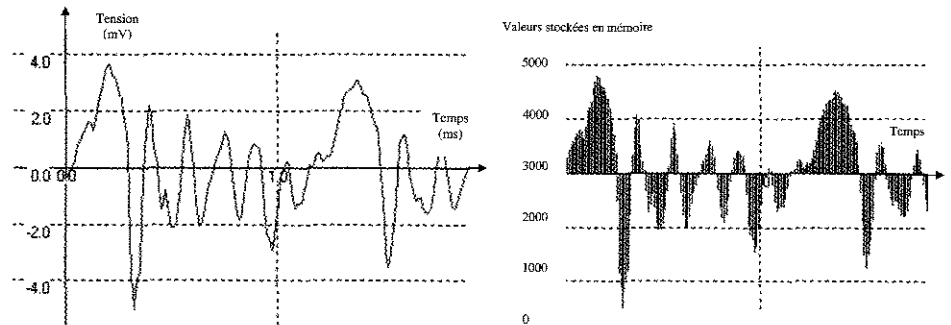


Figure 1 : Allure temporelle du son [a] format analogique à gauche ; échantillonnage du son et stockage en mémoire.

Le logiciel réalise alors le développement en série de Fourier des échantillons sonores, pour en extraire les composantes fréquentielles. Cet outil mathématique permet de décrire toute fonction  $x(t)$  dans un intervalle de temps  $T$ , comme une suite infinie de cosinus et sinus, sous la forme suivante :

$$(1) \quad x(t) = \sum_{n=0}^{\infty} \left( a_n \cos 2\pi n \frac{t}{T} + b_n \sin 2\pi n \frac{t}{T} \right)$$

où

$$(2) \quad a_n = \frac{2}{T} \int_0^T x(t) \cos(2\pi n \frac{t}{T}) dt$$

$$(3) \quad b_n = \frac{2}{T} \int_0^T x(t) \sin(2\pi n \frac{t}{T}) dt$$

Le calcul de  $a_n$  et  $b_n$ , appelés coefficients de la série de Fourier (Equations 2 et 3), impose des calculs assez lourds, à base d'intégrale et de multiplications de cosinus et sinus. Nous nous intéressons principalement au module  $C_n$  exprimé par l'équation (4). Un algorithme appelé "Transformée de Fourier Rapide" permet de réaliser le passage d'un échantillon de son en coefficients  $C_n$ ,  $n$  variant de 1 à 500, plusieurs dizaines de fois par seconde. Il existe une correspondance simple entre  $n$  et la fréquence : à  $n=1$  correspond  $f=1/T$ ,  $n=2$ ,  $f=2/T$ , etc...

$$(4) C_n = \sqrt{a_n^2 + b_n^2}$$

La décomposition spectrale du son de la figure 1 est affichée à la figure 2. L'axe horizontal est cette fois la fréquence, l'axe vertical représente la valeur de  $C_n$  tel qu'il est exprimé dans l'équation (5). Un pic d'amplitude élevé équivaut donc à une forte valeur de  $C_n$ , ou encore une forte composante cosinus ou sinus à la fréquence correspondante. On note une structure spectrale en forme de pics à intervalles réguliers. Ces pics sont appelés harmoniques. Ils sont caractéristiques du spectre de la voix et de la musique.

Le fondamental de la voix correspond au pic le plus à gauche, c'est-à-dire de la fréquence la plus basse. Tous les pics harmoniques ont une fréquence multiple du fondamental. La fréquence du fondamental est appelée  $F_0$ . Elle vaut 115 Hz dans cet exemple. Ainsi, le second pic est à la fréquence  $2 \times F_0$ , le troisième à  $3 \times F_0$ , etc.. On peut ainsi voir jusqu'à 15 harmoniques sur la figure. Le pic d'amplitude maximale ne correspond pas toujours au fondamental  $F_0$ . Ici, le pic d'énergie maximale est à 231 Hz, soit  $2 \times F_0$ .

On désigne par formant un groupe d'harmoniques\*. Sur la figure 2, trois groupes d'harmoniques peuvent être distingués : un formant le plus énergétique regroupant les harmoniques à basse fréquence (115-344 Hz), un formant secondaire englobant les harmoniques intermédiaires (451-816 Hz), enfin un troisième formant pour les harmoniques de haute fréquence (1031 à 1504 Hz). La propriété remarquable de ces formants est la similitude de leur positionnement en fréquence même si l'on modifie la fréquence fondamentale où si l'on teste différents individus.

\*Miller, 1986

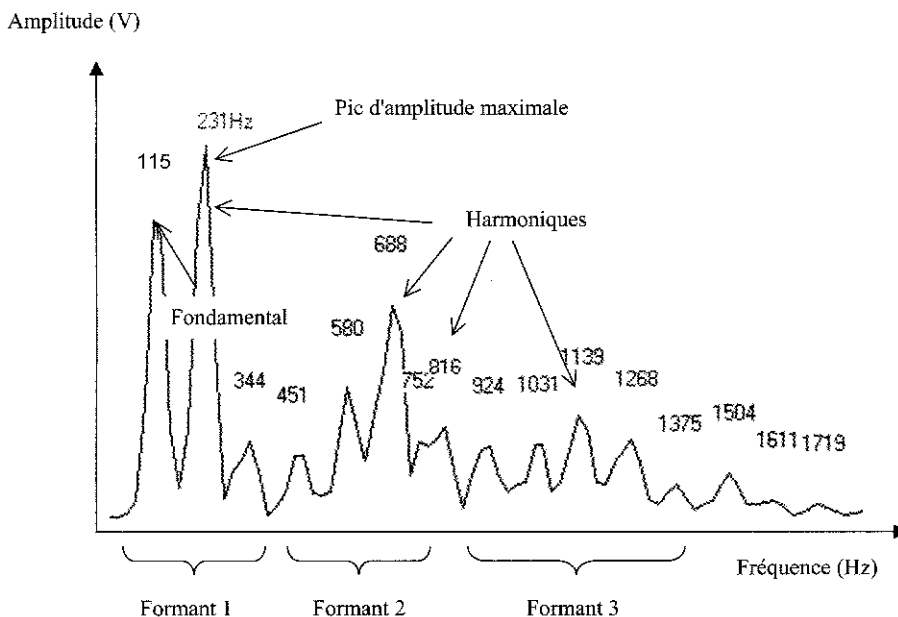


Figure 2 : Spectre du son de la figure 1 montrant le fondamental et ses harmoniques.

Le spectre peut être représenté sous une forme plus ludique : sur la figure 3, chaque harmonique est dessinée sous forme d'un rond. La position du rond sur l'axe vertical correspond à sa fréquence, la taille du rond, corrélée à sa couleur, est liée à son amplitude.

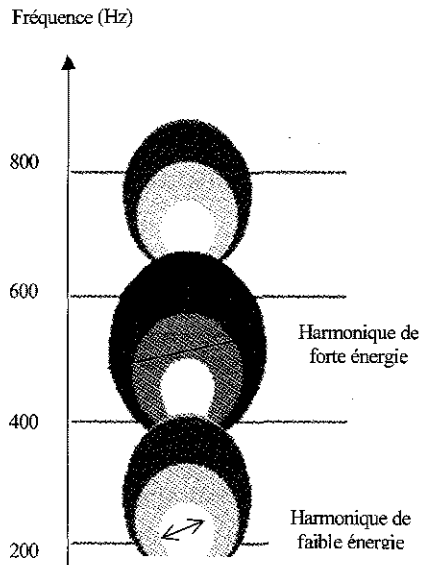


Figure 3: Une portion du spectre est redessinée sous forme de ronds dont le diamètre est lié à l'amplitude de l'harmonique.

### SONAGRAMME

Le sonagramme est une représentation permettant de voir l'ensemble de la décomposition spectrale de la voix sur une même figure. Plusieurs fois par seconde le spectre du son est calculé et les pics harmoniques sont transformés en un jeu de couleurs du plus sombre au plus clair en fonction de l'énergie (Figure 4). Chaque spectre élémentaire est inséré à la droite des spectres existants, permettant de donner une vue de l'évolution des composantes harmoniques (Axe vertical) en fonction du temps (Axe horizontal).

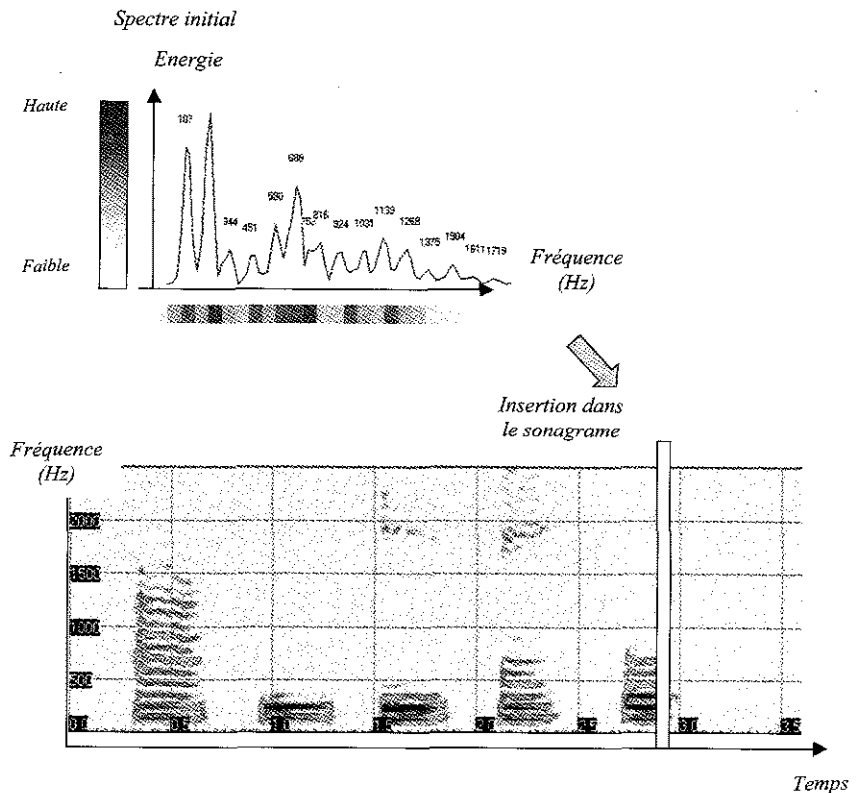


Figure 4: Principes de la construction du sonagramme.

L'analyse spectrale présente un intérêt particulier pour l'évaluation et la rééducation de la voix et de la parole, mais requiert un peu de pratique et d'habitude. Nous avons précédemment conduit une étude de différents sons musicaux et de la parole utilisant le sonagramme, dont les résultats ont été présentés dans Sicard\*. Le spectre de la figure 4 illustre l'enchaînement des voyelles du français standard /a, i, y, e, o/. On distingue bien les harmoniques qui sont représentées par les "traits" horizontaux et certaines zones de renforcement de l'intensité d'harmoniques du son en rapport avec la résonance du tractus vocal.

\*1995

## UTILISATION EN ÉVALUATION

Le module *Evaluation* de Vocalab propose à l'orthophoniste plusieurs outils permettant de réaliser une évaluation objective de la voix, c'est-à-dire une évaluation instrumentale permettant d'obtenir de façon fiable des données chiffrées sur les paramètres de la voix : hauteur, intensité, durée, timbre. Cette évaluation fait suite ou précède l'évaluation subjective, à savoir celle de l'oreille et de l'œil de l'orthophoniste. L'évaluation est dite subjective lorsqu'elle est soumise et donc dépendante de l'expérience du sujet. Tous les auteurs la considèrent comme indispensable et complémentaire de l'approche objective. Cependant, même si la première évaluation "à l'oreille" demeure la principale utilisée actuellement par bon nombre de professionnels comme l'écrit Coulombeau\* (dans Cornut), elle est limitée par son aspect aléatoire et son imprécision. Il existe actuellement des outils professionnels de haute précision tels que EVA\* permettant de réaliser des mesures objectives fiables telles que le débit d'air buccal et nasal, l'électro-glottographie, les mesures de *Shimmer* et de *Jitter*. Cependant ces outils ont l'inconvénient d'être peu accessibles à l'orthophoniste exerçant en cabinet.

\*2002

\*Giovanny, 1994

Dans le module *Evaluation*, la visualisation des composantes spectrales en couleur du type sonagramme est particulièrement intéressante pour l'orthophoniste car elle permet de disposer en temps réel ou en différé d'une véritable photographie de la voix. En effet, distinguer une instabilité en fréquence de celle en intensité est très difficile à l'oreille.

\*2002

Dans Cornut, Coulombeau\* résume parfaitement les éléments à prendre en compte dans l'analyse spectrale : le rapport entre la quantité d'harmoniques et le bruit dans le signal acoustique, les renforcements harmoniques, la stabilité, les modalités d'attaque et d'extinction du son ainsi que le rapport entre l'intensité des harmoniques situés au-dessus de 1 000 Hz et ceux situés en dessous auxquels il faut rajouter la "propreté" du spectre, c'est-à-dire la présence d'éléments aperiodiques tels que le souffle, les éraillures.

Un enregistreur intégré permet de réaliser ses propres échantillons, de les entendre à tout moment, de les stocker et de les analyser. Il est utile de réaliser une dizaine d'échantillons sonores tels que la voix conversationnelle, des vocalises à plusieurs hauteurs, la voix d'appel, modalités d'examen décrites de façon exhaustive par Roubeau\*.

\*1999

Sur le sonagramme de la figure 5, une voix soufflée (en haut) est comparée avec une voix normale (en bas). Pour le spectre du haut, on note l'instabilité en énergie et en fréquence, l'attaque très soufflée, et les traces de souffle en fin de production (zones grises aléatoires). La fréquence fondamentale, c'est-à-dire la raie de fréquence la plus basse, est située vers 330 Hz. Pour la voix normale (en bas), le spectre contient essentiellement les harmoniques, dont la fréquence et l'énergie sont relativement stables. On ne note pas de présence de souffle. L'énergie est plus importante sur le deuxième harmonique que sur la fréquence fondamentale, située vers 450 Hz.

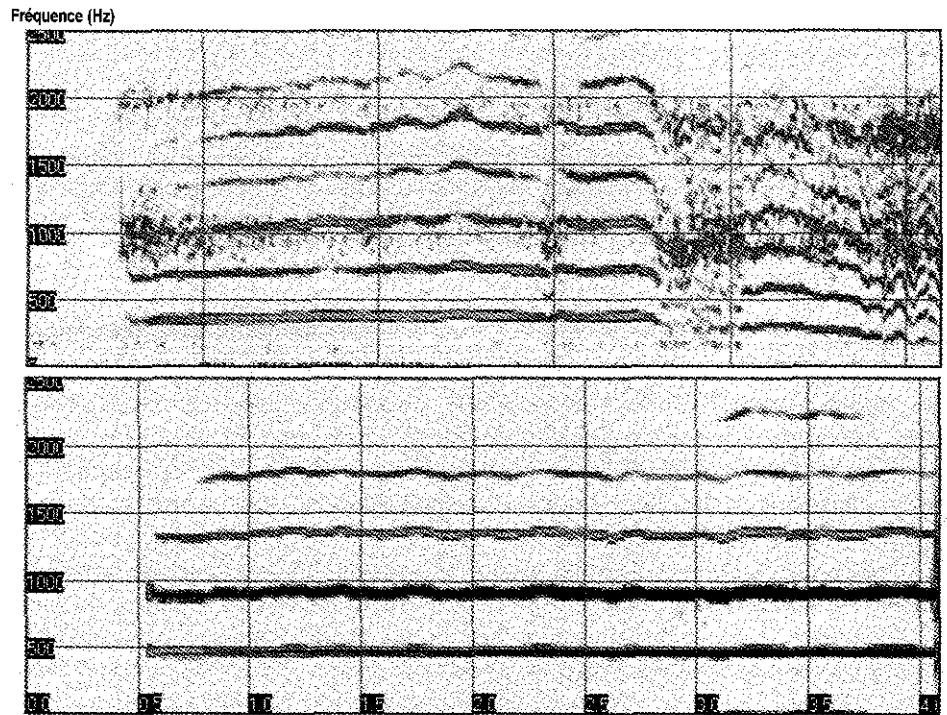


Figure 5 : Visualisation en temps réel de la qualité harmonique de la voix :  
Vocalise [a] (en haut) voix dysphonique ; vocalise [a] en bas, voix normale.

L'analyse spectrale permet de voir les transitions entre le mécanisme I et le mécanisme II de la voix décrits par Roubeau\*. Cette transition lorsqu'elle est mal réalisée se traduit par plusieurs phénomènes visibles sur le spectre : appauvrissement en harmonique, pertes d'énergie sur ces harmoniques, ruptures ou interruption. La "sirène" sur un [a] reportée figure 6 est intéressante. En effet dans la démarche thérapeutique, il est conseillé de partir des "zones de confort" c'est-à-dire les hauteurs où le patient est le plus à l'aise (zone où il y a le plus d'énergies harmoniques) avant d'élargir à des hauteurs où l'aisance vocale est moindre.

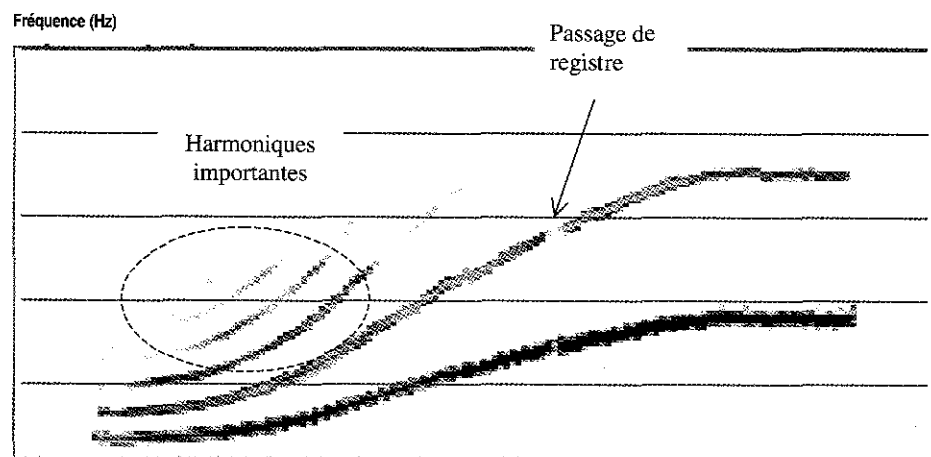


Figure 6 : Sirène sur un [a] : on note la zone résonnante qui apparaît en foncé et l'interruption du son qui correspond au "couac".

Selon le mode de communication, les exigences propres à la voix ne sont pas les mêmes. Dans la voix parlée, c'est essentiellement de la structure des formants et de la précision de l'articulation dont dépendra l'intelligibilité du discours. Dans la voix dite portée ou

projetée utilisée quotidiennement par des professionnels tels que les enseignants, les avocats ou les comédiens, la résonance de la voix, sa richesse harmonique, éléments dits extra-vocaliques représentent un critère de qualité déterminant.

L'aspect phonétique est négligé au profit de la qualité du timbre, de son homogénéité, de sa richesse dans la voix chantée. Selon Miller\* un bon son chanté fait entendre le "vrai son de voix". Les exigences spécifiques de la voix chantée en matière de positionnement et de couplage des résonateurs dépassent largement celles de la voix parlée et requièrent une technique bien particulière.

Sur les spectres, il est utile d'observer la proportion d'énergie relative sur les différents formants. La recherche d'un équilibre entre une pression sous-glottique élevée et une position avantageuse du larynx permet de favoriser et de faire émerger des résonances bien visibles sur le spectre de la figure 7 dans la zone 2 500-3500 Hz. C'est ce que l'on appelle le formant du chanteur.

En résonance optimale, le larynx fonctionne en équilibre et consomme peu d'air. Ce mécanisme de protection du larynx existerait lorsque la pression sus-glottique s'équilibre avec la pression sous-glottique. Cette notion, décrite initialement par Husson\* puis reprise par Cornut\*, et De La Breteque\*\* et Heuillet-Martin\*\*\* dans leurs études, se nomme impédence ramenée sur le larynx (IRL) ou encore couverture du son, terme préférentiellement utilisé par les chanteurs. Nous donnons en figure 7 l'exemple d'une voix normale mais pauvre en énergie extra-vocalique (à gauche), puis une voix plus timbrée où apparaissent des énergies dans la zone 2500-3500 Hz.

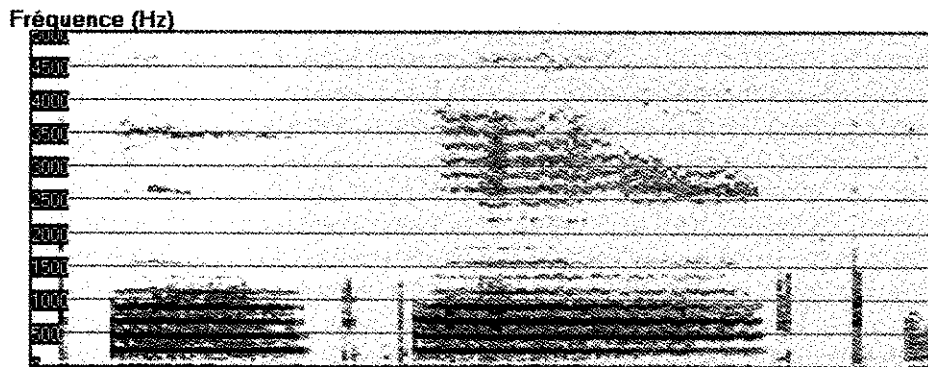


Figure 7 : Vocalise sur le [a], notez l'apparition du formant extra-vocalique sur le son de droite.

Le Temps Phonatoire Maximum (noté TPM) est "le temps de tenue maximale d'une émission vocalique après une inspiration profonde recherchée sur une hauteur et une intensité confortable" selon Heuillet-Martin\*. Le TPM porte sur les sons [s], [z] et [a]. Le logiciel calcule automatiquement le rapport des durées de production des sons [s]/[z] et [a]/[z] ainsi que la moyenne des 3 essais successifs pour les 3 voyelles. Vocalab donne aussi les repères normatifs en fonction du sexe du patient. Si le temps sur [z] est inférieur au temps obtenu sur [s], (rapport [s]/[z] supérieur à 1), cela peut être le témoin d'un couplage phono-respiratoire défectueux sans avoir cependant de valeur de diagnostic. Si le TPM du [z] est très inférieur à celui du [s], Mueller\* suggère que cela pourrait témoigner d'un trouble du fonctionnement laryngé. En effet le [s] peut avoir une durée variable selon la force de constriction. D'autres auteurs pensent que cette valeur permet d'avoir une bonne idée du rendement laryngé du patient et de la pression sous-glottique. Cette épreuve permet d'observer le geste de respiration et de prendre en considération la gestion du souffle. Il constitue un témoin du rendement mécanique du larynx. La figure 8 présente l'écran de Vocalab consacré à l'évaluation du temps phonatoire maximum pour les sons [a], [s] et [z]. Les différents rapports sont calculés puis affichés par le logiciel.

\*1986

\*2002

\*2002

\*\*1997

\*\*\*1995

\*1992

\*1992



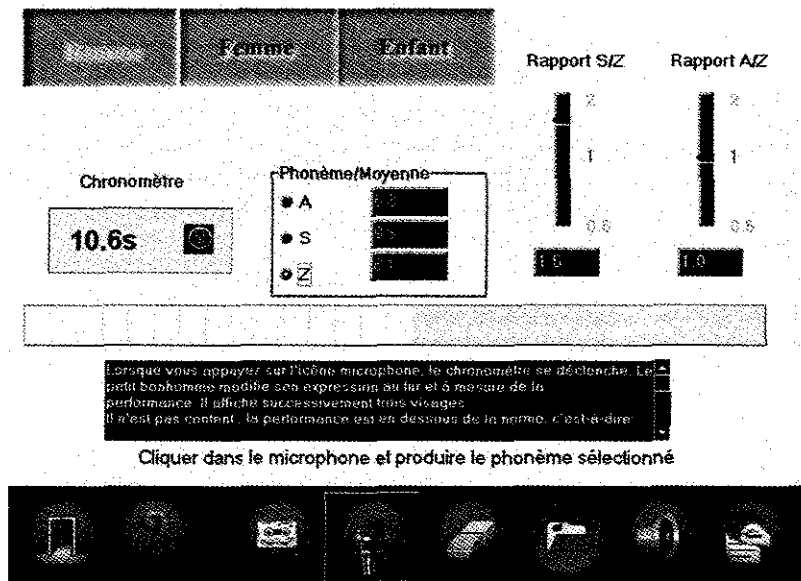


Figure 8: Calcul du TPM et du rapport S/Z chez un homme.

Toutes ces valeurs ne sont pertinentes que corrélées avec d'autres données objectives et avec l'analyse du contexte personnel du patient. Par contre pour un même patient la comparaison entre les valeurs obtenues en bilan et les valeurs après un certain nombre de séances de rééducation permettent de suivre son évolution.

Dans le cadre de l'évaluation de la voix, l'orthophoniste doit pouvoir aussi mesurer de façon précise et rapide le fondamental usuel moyen ainsi que l'étendue vocale de son patient. Ces évaluations sont assez délicates sans outil adapté. Une minorité d'orthophonistes, rompus à l'exercice parviennent sans trop de difficultés à trouver le fondamental usuel de leur patient à l'aide d'un clavier mais ce procédé peu fiable rebute d'emblée les non-initiés. Vocalab calcule de façon précise et fiable le fondamental usuel moyen par un calcul statistique et ce pour n'importe quelle modalité d'examen. Une étude de reproductibilité des résultats est en cours. On obtient ainsi un histogramme dont la base correspond aux hauteurs émises par le patient. Un calcul moyenné permet d'en extraire le Fondamental Usuel Moyen. Celui-ci se trouvant en général dans le tiers inférieur de l'étendue vocale. Le logiciel calcule aussi l'étendue vocale c'est-à-dire les différentes hauteurs que le patient est capable de produire sans forçage et il affiche les valeurs extrêmes calculées instantanément. Le clavier du piano situé juste au-dessus permet de vérifier les données (figure 9). Un phonétogramme permet d'obtenir la dynamique vocale sur ces différentes hauteurs.

La validation scientifique de ces modes de calcul est en cours, en collaboration avec le laboratoire d'acoustique musicale du Pr. Legros à Toulouse. Cette étude est l'un des objectifs scientifiques d'une nouvelle équipe de recherche Unadreo (ERU15).

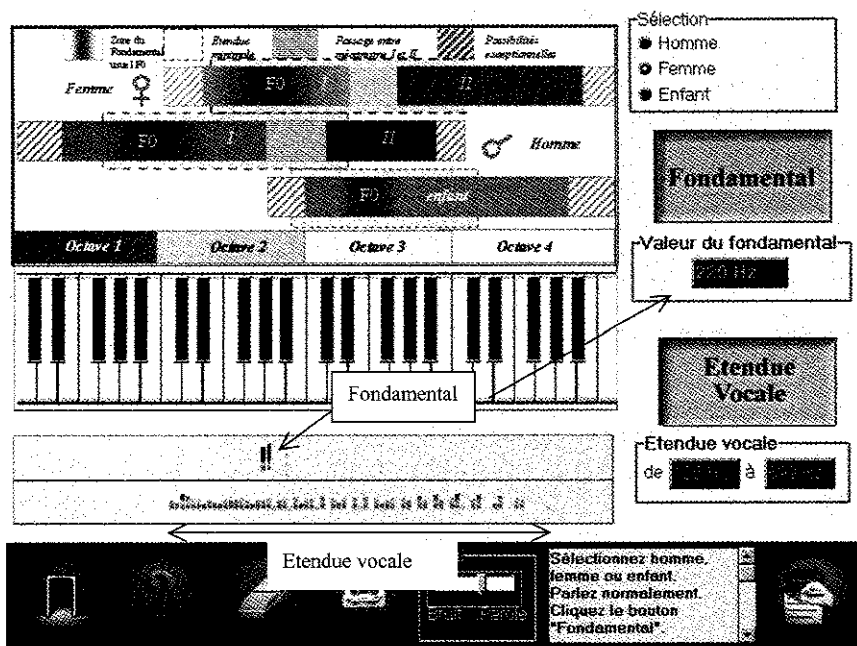


Figure 9: Calcul du Fo et de l'étendue vocale d'une voix de femme.  
Le Fo de 220 Hz correspond à la norme admise, ainsi que l'étendue vocale.

Des repères normatifs de base pour les hommes, les femmes et les enfants sont donnés à l'écran, apparaissant en haut de la figure 9. La zone ombrée correspond au fondamental usuel de voix conversationnelle, les zones hachurées horizontales, aux zones de passage entre le mécanisme I et le mécanisme II qui peuvent se superposer et enfin les hachures obliques aux possibilités exceptionnelles. En pointillés apparaissent pour chaque tessiture les étendues minimales. Le clavier du piano est utile pour jouer des vocalises et connaître la valeur des notes en Hz, évitant ainsi les erreurs d'octave. L'utilisation d'algorithmes spécifiques (tels que le *cepstrum* qui repose sur le principe d'une double transformée de Fourier) a permis de fiabiliser le système de détection du fondamental et de minimiser les effets du parasitage. Il est recommandé toutefois de travailler dans une ambiance la plus calme possible, d'avoir un micro de qualité et de bien régler la carte son et le logiciel pour obtenir de bons résultats.

## UTILISATION EN RÉÉDUCATION

Le module *Rééducation* propose différentes fonctionnalités permettant de travailler spécifiquement les paramètres de la voix et de la parole en rééducation. Nous détaillons dans ce chapitre les outils utiles à la rééducation et illustrons leur fonctionnement par des exemples.

La fonctionnalité *Spectre en temps réel* est sans aucun doute la plus riche en possibilités. Elle permet de travailler chaque paramètre de la voix de façon dissociée. Les patients apprécient particulièrement "*de voir leur voix*" et comprennent mieux les notions abstraites de timbre et de résonance souvent traduites par le thérapeute de façon métaphorique. La visualisation stimule l'intérêt des patients et augmente souvent leur motivation car ils perçoivent mieux la finalité de la prise en charge et ce particulièrement chez l'enfant.

Visualiser le spectre en temps réel est aussi très utile pour travailler les paramètres de la parole avec des enfants ayant un retard de parole ou une déficience perceptive auditive. Le travail des oppositions phonologiques nécessite une perception temporelle d'une finesse extrême que les enfants ayant un retard de parole et/ou plus tard une dyslexie de type phonologique n'ont pas. La plupart des auteurs notamment Dumont\* s'accordent sur l'hypothèse d'un trouble du traitement temporel auditif qui serait responsable des confusions pho-

\* 2001

némiques sourdes/sonores. Prenons l'exemple de la confusion auditive [k/g] particulièrement tenace. L'orthophoniste commence par montrer à l'enfant la différence en choisissant un réglage adéquat (bande de fréquence 0-1 000 Hz et vitesse maximale). Il produit le [k] puis le [g] en prenant bien garde de ne pas produire le [e] derrière. Il commente avec lui les différences entre les deux dessins (figure 10). L'orthophoniste encourage l'enfant à l'imiter. Il est rare qu'il y parvienne au premier essai et produit très souvent le "e" derrière le "g" ou ne produit pas la très brève sonorisation juste avant l'occlusion. Celle-ci est bien visible sur le spectre modèle mais ne sera pas sur celui de l'enfant tant qu'il n'aura pas produit correctement cette vibration initiale. Lorsque l'enfant est capable de produire correctement les occlusives sonores isolément et en contexte, on peut penser qu'il sera capable de les repérer, de les identifier et de faire le bon choix lors du passage à l'écrit.

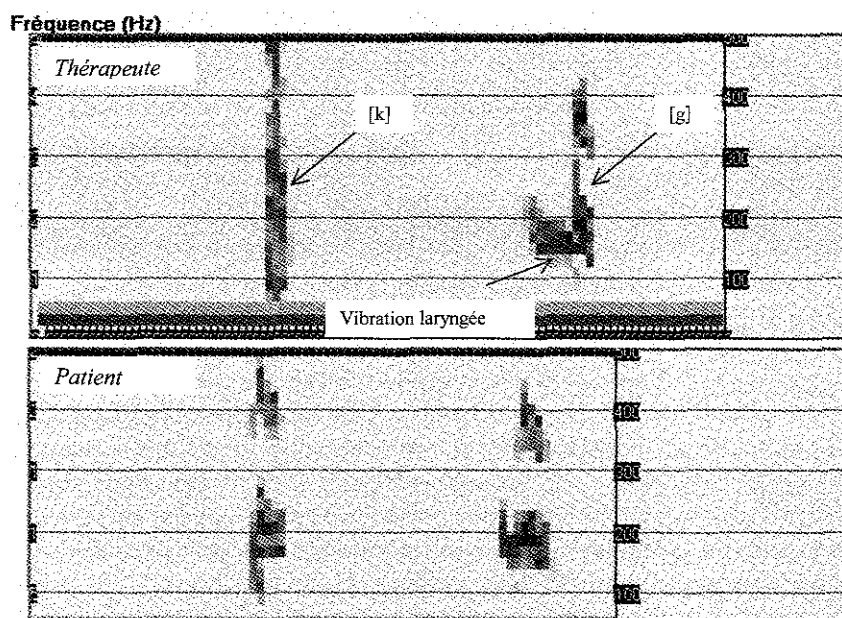


Figure 10: En haut : modèle du thérapeute, en premier le "c" occlusive très brève suivie du "g" plus longue car vibration laryngée. En bas l'enfant produit le "keu" suivie de "gheu", la différence visuelle est peu marquée.

De manière similaire, le trait distinctif entre les constrictives sourdes et sonores (ex : [f] [v]) est bien visible à l'écran figure 11. L'orthophoniste propose à l'enfant de répéter des syllabes puis des mots contenant le phonème travaillé en initiale, médiane et finale. En utilisant les réglages adaptés, tous les phonèmes peuvent être travaillés de cette façon-là à l'exception des phonèmes occlusifs [p] [t] ; [b] [d] qui ne peuvent se distinguer sur le spectre.

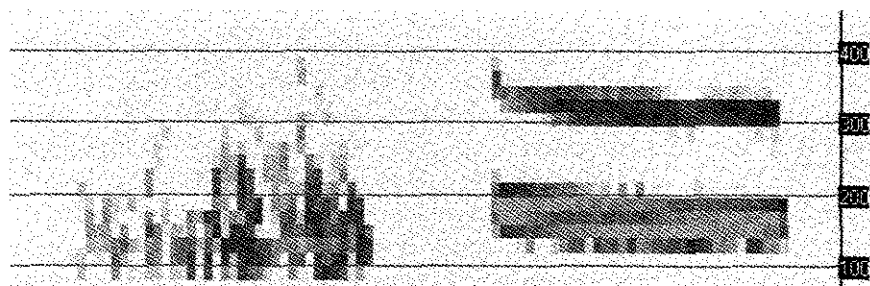


Figure 11 : à gauche le phonème [f] bruit apériodique, suivi du phonème [v] constrictive sonore, périodique. La vibration laryngée apparaît sous forme de deux traits continus.

La fonctionnalité *Timbre* est très intéressante pour le travail de la résonance dans le cas d'un patient dysphonique ou présentant des difficultés de discrimination phonologique. L'axe des fréquences est vertical et la dimension du temps n'est plus représentée. L'écran est divisé en deux parties. La partie gauche destinée au thérapeute et la partie droite destinée au patient.

Pour la représentation du timbre, le code de couleur est le même que pour la fonctionnalité spectre en temps réel. La couleur et la grosseur des cercles sont proportionnelles à leur intensité. Plus les ronds apparaissent en jaune ou blanc plus l'énergie sur ces fréquences est importante.

Pour les enfants déficients auditifs ou même pour des enfants présentant des retards de parole, il est utile de visualiser les paramètres du son et notamment le paramètre d'intensité. L'enfant peut apprendre à contrôler l'intensité de la voix en imitant le modèle du thérapeute.

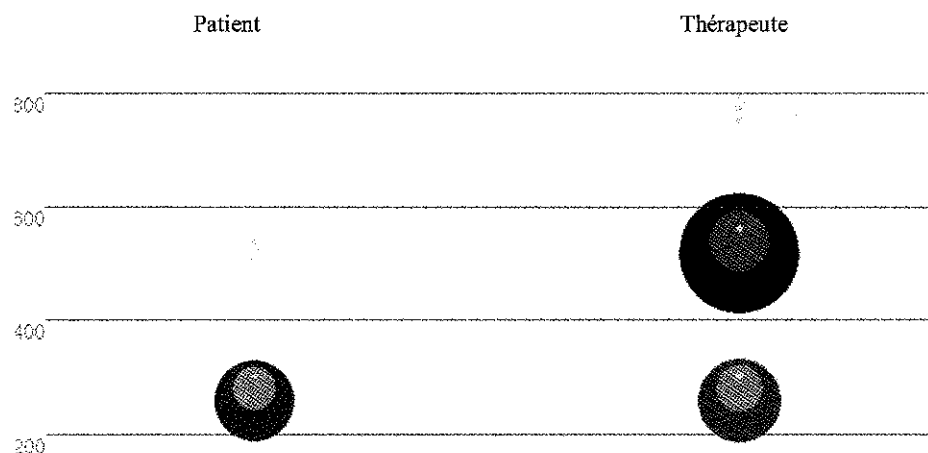


Figure 12: En haut, phonème [i] avec très peu de ronds donc d'harmoniques ; en bas le [é] comporte plus de ronds car la voyelle est plus ouverte.

L'écran de l'outil *Timbre* se révèle particulièrement utile pour visualiser et donc contrôler l'intensité de sa voix, en imitant la grosseur et la couleur des ronds du modèle ou de maîtriser ainsi la hauteur de sa voix en lien avec la hauteur du rond. Les ronds de couleur se déplacent vers le haut quand la voix monte et vice-versa. Cette fonctionnalité permet d'illustrer les différences phonémiques telles que [é]/[i] (figure 12), orale/nasale [a]/[an] ou encore le trait voisé/non voisé. Le thérapeute produit le son isolément et fige le résultat visuel. L'enfant doit ajuster ses organes phonateurs pour obtenir le même résultat.

La fonctionnalité *Hauteur Tonale* permet de visualiser les variations de hauteur du fondamental laryngé. Avec l'enfant ou l'adulte, ce module sert à aider le patient à contrôler la prosodie de la voix ou de la parole. Il est particulièrement utile pour les patients qui ont des difficultés à percevoir si leur voix monte ou descend. Le "dessin" mélodique du thérapeute est reproduit par le patient, comme l'illustre la figure 13. La fonctionnalité hauteur tonale est aussi utile pour les enfants sourds, pour les aider à comprendre et donc à maîtriser l'intonation et la mélodie de leur voix.

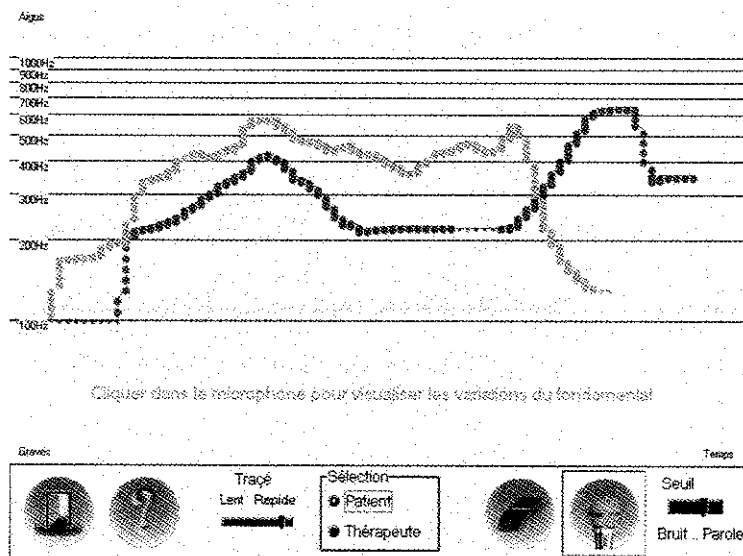


Figure 13 : Deux schémas mélodiques différents sur la vocalise "o"

Cette fonctionnalité peut être aussi utilisée pour travailler la stabilité du fondamental chez les patients dysphoniques et les chanteurs ou à visualiser les pauses inspiratoires. Il est aussi possible de visualiser en temps réel les variations intonatives de la parole mais l'outil perd en fiabilité à cause des bruits d'explosives et de fricatives qui parasitent et altèrent la détection du fondamental.

Enfin, la fonctionnalité *Articulation* a été élaborée pour faciliter le travail spécifique d'articulation des consonnes constrictives [ch], [j], [s], [z], [f], [v]. L'écran se compose d'une visualisation temporelle du son, telle que captée par le microphone, et un accès à une série d'animations. L'allure temporelle du son se modifie selon le phonème produit (figure 14).

L'enfant voit sa production et compare sa courbe au modèle du thérapeute. Pendant que l'enfant tente de produire le phonème demandé, l'orthophoniste renvoie un modèle correctif pour amener progressivement l'enfant à modifier la position de sa langue.

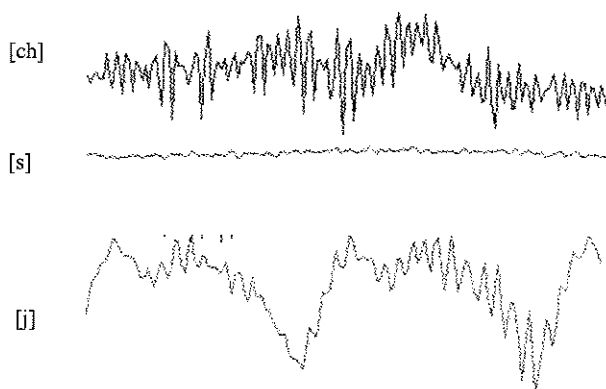


Figure 14 : En haut le phonème [ch], au milieu [s], en bas [j].  
Notez la périodicité de ce dernier liée à la vibration laryngée.

L'animation est déclenchée par l'orthophoniste. Elle suscite une motivation étonnante chez l'enfant et contribue à rendre moins lassantes les séances de répétition nécessaires à l'ancrage du bon geste. Même si l'enfant ne produit pas encore correctement le son demandé, ses efforts pour y parvenir sont toujours gratifiés par l'orthophoniste qui déclenche lui-même l'animation (figure 15). En outre, l'ordinateur joue ici pleinement son rôle de médiateur dans la relation.

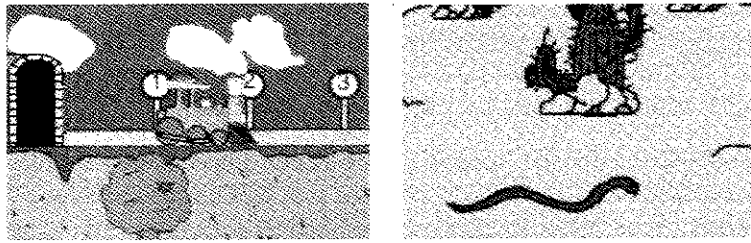


Figure 15 : Deux exemples d'animations : la locomotive pour le [ch] ;  
le serpent pour le [s].

Une bibliothèque de sons enregistrés et des vidéos illustrant les cordes vocales en phonation sont accessibles depuis la médiathèque de Vocalab. Ces vidéos sont utiles pour expliquer aux patients dès le début de la prise en charge, à quoi ressemblent leurs cordes vocales et comment elles fonctionnent. Différentes voix pathologiques sont aussi disponibles, comme le montre la figure 16, grâce aux données fournies par le Dr. Woisard\* du service ORL de l'hôpital de Rangueil, Toulouse. Les échantillons sonores ont pour objectif de mettre à disposition des orthophonistes différentes illustrations des timbres de voix, pathologiques ou non, des voix de femmes, d'hommes et d'enfants, ainsi que des sons d'instruments de musique et de bruits variés.

\*1998

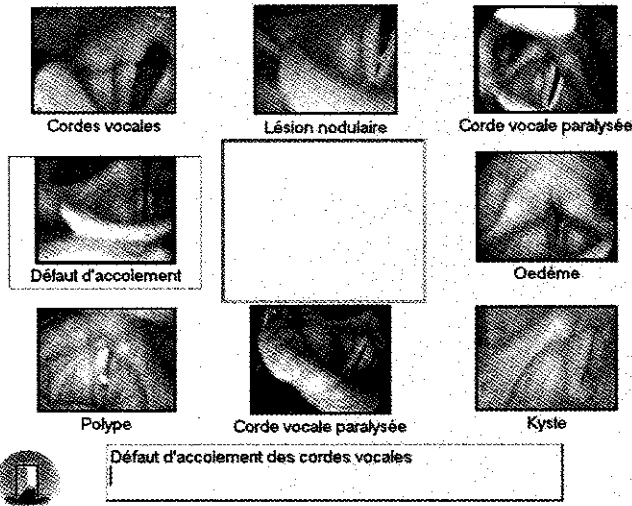


Figure 16 : Médiathèque : vidéos de cordes vocales en phonation,  
normales et pathologiques

En 2001, le logiciel Vocalab s'est vu récompensé par le prix du meilleur CD-ROM au festival audiovisuel de Nancy. Le logiciel est diffusé en France depuis septembre 2002, où environ 700 sites ont été équipés. Une version en Anglais a été présentée au congrès CPLOP à Edinburg, Ecosse\*, en 2003, et à l'ACOA à Ottawa, Canada en 2004. Le logiciel est diffusé au Canada depuis 2004 par "la société Magie Mots" en tenant compte des spécificités de la rééducation orthophonique canadienne.

\*Menin, 2003

## DISCUSSION

L'écran joue le rôle de médiateur objectif dans la relation lorsque celle-ci chargée d'affect semble parfois faire obstacle à la progression du patient et ce malgré nos efforts. L'ordinateur contribue à diversifier le travail, à lever les inhibitions notamment lorsque le travail porte sur la voix à plus forte intensité. Il renforce le message auditif et rend objectif ce qui parfois échappe à l'oreille de l'orthophoniste pourtant entraînée. A l'inverse, une utilisation abusive et non adéquate peut avoir des effets pervers qu'il faut ten-

ter d'éviter. Le patient peut devenir trop dépendant de cette image ou bien focaliser son attention sur l'écran au détriment de son écoute et de ses perceptions kinesthésiques. L'intérêt du feed-back visuel est de rendre plus concret un message complexe que nos patients petits et grands perçoivent, contrôlent et analysent mal.

Bien que simple et convivial, Vocalab nécessite des manipulations préalables et une formation spécifique pour une utilisation optimale. Une documentation en ligne symbolisée dans chaque écran par un point d'interrogation permet néanmoins à tout moment d'être guidé dans son utilisation. Différentes formations ont été organisées à Toulouse, Paris, Lyon et Montréal, où il s'est avéré que la plupart des orthophonistes étaient très peu familiarisés avec la forme et le spectre des sons. Pouvoir interpréter les informations fournies par l'outil *Spectre en Temps Réel* par exemple nécessite une bonne compréhension des harmoniques, une habitude à localiser le fondamental, à distinguer les fluctuations de hauteur, le souffle, les vibrations laryngées, etc... Ces concepts ont été illustrés dans cet article, comme ils l'ont été en formation, mais seule une pratique personnelle régulière permet d'appréhender par étapes successives le monde étrange mais non moins passionnant des spectres.

## CONCLUSION

Dans cet article, nous avons décrit l'utilisation en orthophonie du logiciel Vocalab, spécifiquement conçu pour les orthophonistes pour réaliser des mesures objectives et visualiser en temps réel la structure harmonique de la voix. Nous avons détaillé la chaîne d'acquisition et de traitement du son, en insistant notamment sur la construction du sonagramme. Nous avons ensuite décrit l'utilisation de Vocalab pour l'évaluation et la rééducation de la voix.

Prochainement, une version en anglais de ce logiciel devrait voir le jour. Dans le futur, le logiciel devrait bénéficier d'une bibliothèque de sons et de vidéos étoffée, grâce à l'aimable participation du service ORL de l'hôpital de Rangueil à Toulouse.

## BIBLIOGRAPHIE

- AMY DE LA BRETEQUE, B. (1997). *L'équilibre et le rayonnement de la voix*. Marseille: Solal, 122 p.
- BUCHMAN, L., ARABIA, C., RIGARD, M.T., CHEVRIE-MULLER, C. (1991). Méthodes pour le bilan phoniatric. *Glossa*, 26, 30-36.
- COUDIERE, C. (2003). *De l'utilisation des logiciels pour la voix en rééducation des dysphonies dysfonctionnelles*. Toulouse: Mémoire d'orthophonie, 52-53.
- CORNUT, G. (2002). *Moyens d'investigation et pédagogie de la voix chantée*. Lyon: Symétrie 200 p.
- DALLEAS, B.R., DALLEAS, A. (1988). Evaluation acoustique objective de la qualité du timbre de la voix parlée contribution au bilan des dysphonies. *Bulletin d'Audiophonologie* 4, 77-78.
- DUMONT, A., LORENZI, C., GERARD, CL., VAN DEN ABEELE, T. (2001). Acuité temporelle auditive et troubles du langage. *Glossa*, 76, 28-37.
- GIOVANNI, A., TESTON, B., ROBERT, D., GUARELLA, M.D., ZANARET, M., CANNONI, M. (1994). Evaluation vocale assistée: présentation du matériel et de la méthode. *Bulletin d'Audiophonologie*, 5, 529-542.
- HEUILLET-MARTIN, G., DANOY, M.C. (1992). Le Protocole du bilan vocal, *Bulletin d'audiophonologie*, 4, 63-69.
- HEUILLET-MARTIN, G. GARSON-BAVARD, H. LEGRE, A. (1995). *Une voix pour tous*. Marseille: Solal. Tome 1 p 68-69.
- MENIN, A., SICARD, E. (2003). Vocelab, a software for voice and speech control in Actes du CPLOL, Edimburgh, 2003.
- MUELLER, P.B., LARSON, G. W. SUMMERS, P.A. (1992). The viability of the S/Z ratio as a clinical voice evaluation tool. *Bulletin d'audiophonologie*, 8, 379-386
- MILLER, R., (1986) La structure du chant. Paris: Cité de la musique. p 52-59.
- ROUBEAU, B. (1999) Déterminisme laryngé des registres vocaux. *Journal de l'AFPC*, p 47-52.
- ROUBEAU, B. (2002) Le bilan vocal, *Rééducation Orthophonique*, p 54-59.
- SICARD, E., MENIN, A. (1995). Analyse Spectrale des sons musicaux et de la parole. *Bulletin de l'Union des Physiciens*, 776, p 70-76.
- WOISARD, V. (1998). Dysfonctionnement des cordes vocales. *La revue du praticien*, Tome 12-441, p15-18.