

Analyse Acoustique des Voix de Patients Laryngectomisés Algériens

Kamel Ferrat, Mhania Guerti

المُلخَص

تقدم هذه الدراسة تحليلا صوتيا للأشخاص الذين تعرّضوا لاستئصال كليّ للحنجرة في المحيط الاستشفائي الجزائري، ولهذا الغرض اعتمدنا مدوّنة صوتية أعدت في الفترة بين أكتوبر 2008 وسبتمبر 2009، شارك فيها 8 من أشخاص مستأصلي الحنجرة كليّا. أجريت التّسجيلات قبل مرحلة إعادة التأهيل، وبعد 3 و6 و11 شهرا أثناء مرحلة إعادة التأهيل. استعملت برمجيات PRAAT و MATLAB التي تناولت التردد الأساسي والانحراف المعياري له وقيمتيه الدنيا والقصوى والبواني الثلاثة الأولى F1، F2، و F3، وكذلك نسبة التوافق للضجيج (HNR) والشدة (Intensité) ودرجة اضطرابها (Shimmer) ودرجة اضطراب التردد الأساسي (Jitter) ونسبة الهمس في الكلام (DUF). تشير النتائج المحصّل عليها إلى اضطراب معتبر للتردد الأساسي وانخفاض في الشدة مقارنة بالصوت السليم، كما يلاحظ ارتفاع محسوس في عاملي Jitter و Shimmer وانخفاض في عامل HNR وارتفاع في عامل DUF عند المدّ. يهدف هذا العمل إلى تشخيص موضوعي للكلام المضطرب واستثماره في الوسط الاستشفائي الجزائري.

الكلمات المفاتيح : استئصال الحنجرة، اضطرابات الكلام، التحليل الفيزيائي، الجيتر (Jitter)، الشيمر (Shimmer).

Analyse Acoustique des Voix de Patients Laryngectomisés Algériens

Kamel Ferrat¹, Mhania Guerti^{1,2}

¹ Centre de Recherche Scientifique et Technique pour le Développement
de la Langue Arabe (CRSTDLA), Alger, Algérie.

² Laboratoire Signal et Communications,
Ecole Nationale Polytechnique Alger Algérie

kamelferrat@yahoo.fr, mhanian.guerti@enp.edu.dz

Résumé

Dans le cadre de cette étude, nous présentons une analyse acoustique de la voix de patients ayant subi une Laryngectomie Totale (LT) en milieu hospitalier algérien. Pour réaliser cet objectif, nous avons utilisé un corpus de sons collectés entre Octobre 2008 à Septembre 2009. Ce dernier a été prononcé par huit patients. Les enregistrements ont été faits avant le début de la rééducation et après trois, six et onze mois de rééducation.

L'analyse acoustique a été réalisée en utilisant le logiciel d'analyse Praat et le logiciel de programmation Matlab. Cette analyse concerne la fréquence fondamentale F_0 (Hz) avec l'écart-type et les valeurs extrêmes (F_{0max} et F_{0min}), les trois premiers formants (F_1 , F_2 , F_3), le Jitter (%), le Shimmer (dB), le HNR (dB), le pourcentage de trames non voisées DUF (%) et l'intensité (dB).

Les résultats de l'analyse nous montrent des troubles assez importants de F_0 et une diminution de l'intensité par rapport à la voix laryngienne normale. De plus, nous notons une augmentation perceptible des valeurs du Jitter et du Shimmer, une diminution des valeurs HNR et une aug-

mentation du DUF, lors de la prononciation des voyelles soutenues.

Ce travail a pour but la caractérisation objective des paroles pathologiques en vue de leur exploitation en réhabilitation de la parole dans le milieu hospitalier algérien.

Mots Clés : Laryngectomie Totale, Pathologie de la voix, Analyse acoustique, sons arabes, Jitter, Shimmer.

1. Introduction

Une Laryngectomie Totale (LT) est une procédure chirurgicale qui consiste en l'ablation complète du larynx pour le traitement d'un cancer à l'état avancé. Ainsi, le patient perd ses cordes vocales qui lui permettent la voix laryngée. De même, l'air provenant des poumons, source de départ de toute parole humaine, ne passe plus par les cavités de l'appareil phonatoire. C'est-à-dire que le malade après ablation du larynx se trouve amputé de la partie essentielle de la voix et de la parole. Une technique de rééducation appropriée permet de lui procurer une nouvelle parole dite "œsophagienne" [1-3].

A quoi correspond cette parole ? Comment sont les valeurs de ces caractéris-

tiques acoustiques par rapport à la voix laryngienne normale ? Comment évoluent ces caractéristiques au cours du temps et par rapport à la technique de rééducation utilisée ?

Répondre à ces questions est l'un des objectifs essentiels de notre étude.

2. Parole Pathologique après LT

Après Laryngectomie Totale, différentes voies pour la restauration de la communication orale sont exploitées. Nous distinguons principalement trois approches de réhabilitation vocale : la parole œsophagienne ES (Esophageal Speech), la parole trachéo-œsophagienne TES (Tracheo-Esophageal Speech), le larynx artificiel AL (Artificial Larynx) [4-7].

La Technique ES est l'approche prédominante de la réhabilitation de la parole de patients ayant subi une LT.

Lors de l'opération chirurgicale, une ouverture permanente, appelée "trachéostomie", est pratiquée à la base du cou. Cette ouverture permet au patient de respirer en empruntant donc cette nouvelle voie (figure 1) [6].

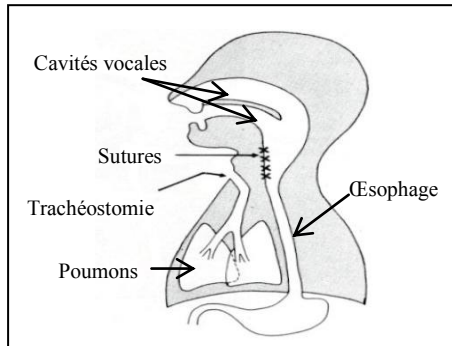


Figure 1: Trachéostomie après ablation du larynx

Comme, le patient ne peut plus utiliser l'air provenant des poumons, l'œsophage jouera désormais le rôle de réserve d'air.

C'est pourquoi, nous parlons de "parole ou voix œsophagienne" (figure 2).

A l'entrée de l'œsophage se trouve un petit muscle qui peut vibrer en produisant un son si une bouffée d'air remonte en le traversant. Ce petit muscle joue le rôle de vibreur de remplacement.

Nous retrouvons bien une analogie avec les trois conditions nécessaires à la parole: un souffle d'air provenant de l'œsophage, un muscle vibrant, souple et élastique appelé segment Néo-Vibreur Pharyngo-œsophagien NVP et les cavités vocales qui n'ont pas subi de changements notables. Il reste donc au patient à apprendre à exploiter ce nouveau mécanisme de la production de la parole [8-9].

La méthode consiste à emmagasiner de l'air dans la partie inférieure de l'œsophage (déglutition de l'air) et provoquer ensuite une éructation volontaire qui fait vibrer le segment NVP, pour donner un semblant de voisement (figure 2).

En Algérie, la technique ES est la plus privilégiée. En effet, elle ne nécessite pas de dispositifs coûteux tels que les prothèses. En outre, cette méthode est généralement commode et permet un degré élevé d'intelligibilité après une période suffisante de rééducation [10-12].

3. Analyse Acoustique de la Parole Pathologique PP

L'étude de la Parole Pathologique (PP) pourra conduire à des diagnostics automatiques et à l'établissement de systèmes experts permettant de caractériser les anomalies vocales.

La recherche de techniques objectives est certes difficile. En premier lieu, parce que nous ne pouvons se passer de notre système auditif pour entendre et analyser la production vocale.

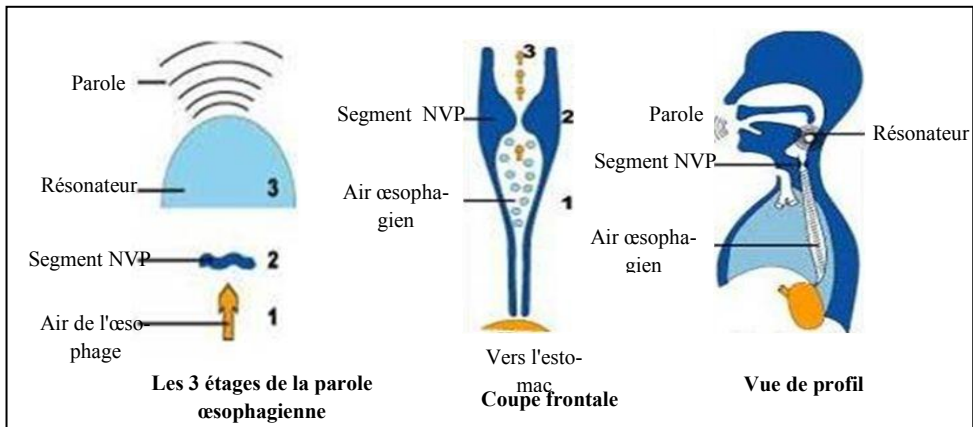


Figure 2 : Principe d'une parole œsophagienne.

L'analyse perceptive reste donc la référence, bien qu'elle soit subjective en évaluant le trouble vocal comme un tout. Aujourd'hui avec le développement d'outils et de logiciels informatiques assez performants, l'analyse acoustique tente d'apporter cette objectivité qui a toujours fait défaut dans la caractérisation de la voix et de la parole humaines. En d'autres termes, elle renforce et appuie le message auditif et rend objectif ce qui échappe parfois à l'oreille de l'orthophoniste rééducateur.

Ces dernières années, l'analyse des PP a connu un vif intérêt aussi bien dans l'étude de leurs physiologies et physiopathologies mais surtout de leurs moyens d'exploration et de mesure. Chaque mesure apporte des informations différentes sur les aspects de la production sonore. Une grande complexité des phénomènes de la production vocale rendent impossible l'élaboration d'une méthode unique de mesure et de quantification. En conséquence, aucune mesure ne suffit à elle seule à caractériser la PP.

Une Parole Normale PN est analysée essentiellement par l'observation des principaux paramètres suivants :

- le pitch ou fréquence fondamentale (F_0), qui permet de mesurer les vibrations des cordes vocales et de distinguer entre un son aigu et un son grave;
- les formants qui nous permettent d'étudier les transformations que subit le signal de parole lors de son passage à travers les cavités de l'appareil phonatoire;
- l'énergie du signal de parole qui nous permet de distinguer entre un signal fort et un signal faible.

Cependant, l'analyse d'une PP fait appel en addition, à d'autres paramètres aussi importants basés sur la technique d'analyse cepstrale du signal vocal. Parmi ces paramètres, nous avons utilisé :

- le Jitter qui correspond au degré de perturbation de la fréquence fondamentale (F_0), donc aux troubles de vibrations des cordes vocales de l'appareil phonatoire;
- le Shimmer qui correspond au degré de perturbation de l'intensité, donc aux

passages brusques et anormales d'une voix forte à une voix faible lors d'un acte de parole et vice-versa;

- le pourcentage de trames non voisées à l'intérieur du signal de parole ou Degree of Unvoiced Frames (DUF);
- le rapport Harmoniques/bruit HNR (Harmonics to Noise Ratio) qui permet d'évaluer l'influence du bruit sur les harmoniques du signal, afin d'évaluer le timbre de la parole, c'est-à-dire sa qualité [13].

Les deux paramètres Jitter et Shimmer sont largement utilisés pour la caractérisation de la qualité de la PP [14-16]. Le Jitter augmente en cas de pathologie vocale et une valeur supérieure à 1.04 % correspond à une PP. Le Shimmer augmente également en cas de pathologie vocale, et une valeur supérieure à 3.81% correspond à une PP. En décibel, ce seuil correspond à 0.35 dB [17].

3.1. Prétraitement du signal de parole

Les conditions d'acquisition du signal de parole telles que le bruit environnant et la fidélité des équipements du microphone (distorsions et bruits du filtrage du canal de transmission) jouent un très grand rôle sur la qualité d'acquisition, et donc sur la qualité du Traitement Automatique de la Parole (TAP). Ainsi, nous devons faire un prétraitement pour le signal avant d'extraire les paramètres acoustiques, de manière à éliminer les facteurs extérieurs qui pourraient influencer sur les résultats à obtenir :

- une préaccentuation dont l'objectif est d'augmenter la quantité d'énergie dans les hautes fréquences et d'avoir une compensation de filtrage des effets de l'acquisition du signal. Pour cela, le signal de parole enregistré est appliqué à l'entrée d'un filtre de premier ordre FIR (Finite Impulse Response) (Figure 3).

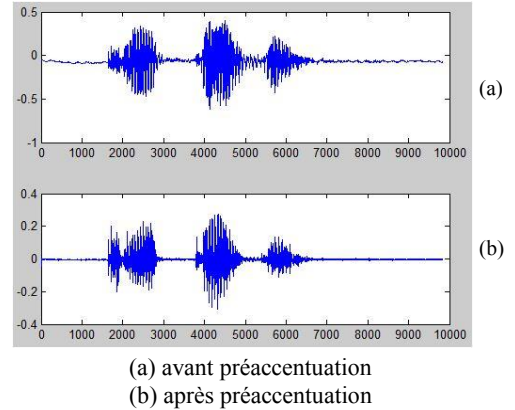


Figure 3 : Représentation du mot [kataba].

une élimination des échantillons inutiles du signal de parole enregistré, sans perdre les informations pertinentes véhiculées par ce signal (Figure 4). Une fonction procédure sous Matlab utilise un seuil minimal d'énergie moyenne, calculée sur la base des enregistrements de différents bruits d'environnement.

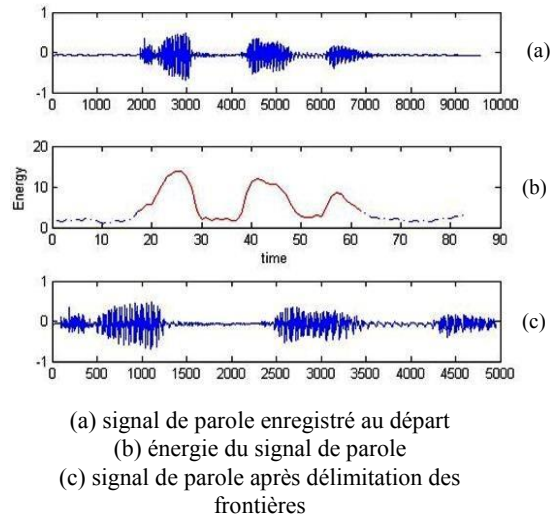


Figure 4 : Elimination des trames inutiles de début et fin du mot [kataba].

Dès que l'énergie dépasse un seuil minimal dans une trame (Figure 4.b), nous considérons que le début de parole commence à partir de cette trame. Toutes les autres trames précédentes sont éliminées. La même procédure est appliquée à la fin du signal de parole.

- une extraction des paramètres acoustiques sur des segments de signal de parole supposés stables, du fait que celui-ci est connu comme non stationnaire. Pour cela, nous choisissons des fenêtres de taille de 30 ms, car l'observation du signal de parole montre qu'il n'évolue pas ou peu sur des durées de cette taille. Les paramètres sont extraits avec un pas de 10 ms sur toute la fenêtre.

3.2. Analyse acoustique de la Parole Œsophagienne

Nous avons fait une analyse acoustique à l'aide du logiciel de programmation Matlab et du logiciel d'analyse acoustique Praat. Cette analyse concerne la fréquence fondamentale F_0 (Hz) avec l'écart-type et les valeurs extrêmes (F_{0max} et F_{0min}), les trois premiers formants (F_1 , F_2 , F_3), le Jitter (%), le Shimmer (dB), le HNR (dB), le pourcentage de trames non voisées DUF (%) et enfin l'énergie (dB). Nous avons utilisé le logiciel Praat pour extraire les paramètres suivants : F_0 , F_1 , F_2 , F_3 , DUF, intensité et HNR. Le Jitter, le Shimmer et l'énergie ont été extraits à partir du logiciel Matlab.

3.2.1. Enregistrements du corpus

Nous avons exploité un corpus de sons collectés entre Octobre 2008 et Septembre 2009. Les enregistrements ont été effectués avant le début de la rééducation, et après trois, six et onze mois de rééducation. Le corpus se compose de voyelles

soutenues, de paires [CV] (C: Consonne, V: Voyelle), de mots isolés et de textes lus. Pour l'extraction des paramètres acoustiques, nous avons exploité les voyelles soutenues [ā, ī, ū]. Les paires [CV] et les mots isolés ont été exploités pour détecter d'éventuelles confusions entre les consonnes ayant un même lieu d'articulation, mais diffèrent par le trait Voisé/Non Voisé. Pour cette raison, nous avons choisi les consonnes correspondant à des paires Voisé/Non Voisé: pharyngales [ɛ] (ع) / [H] (ح); uvulaires [ɣ] (غ) / [χ] (خ), vélaires [g] / [k] et dentales [d] / [t]. Les textes lus nous ont permis d'estimer la fluidité et l'intelligibilité de la parole prononcée par les patients.

3.2.2. Présentation des patients laryngectomisés

Les sons ont été prononcés par huit patients ayant subi une Laryngectomie Totale et s'exprimant en Arabe dialectal Algérois. L'âge de ces patients varie entre 47 à 59 ans, l'âge moyen est de 54 ans. Tous les patients ont bénéficié d'une rééducation vocale par un orthophoniste (séances périodiques durant 11 mois). Ils ont appris la nouvelle parole ES par la technique dite "d'inhalation" [10]. Le même corpus a été prononcé par cinq locuteurs normaux, ne présentant aucune pathologie de la voix, ni de la parole.

3.2.3. Résultats et interprétations

Un exemple de paramètres extraits en utilisant la voyelle soutenue [ā] est donné dans les Tables 1et 2.

Toutes les valeurs des paramètres acoustiques sont extraites à partir de sons prononcés par les patients avant et après 3, 6 et 11 mois de rééducation.

Voyelle [ā]	Jitter (%)	Shimmer (%)	Shimmer (dB)	HNR (dB)	DUF (%)
Normal	00.247	03.410	00.297	20.748	00.000
Avant rééduc.	-	-	-	-	88.570
3 mois	12.271	07.891	01.119	01.710	46.930
6 mois	06.013	07.910	01.114	02.154	41.370
11 mois	01.745	10.588	00.681	03.414	13.115

Table 1. Valeurs du Jitter, Shimmer, HNR et DUF de la voyelle [ā].

Voyelle [ā]	Pitch F ₀ (Hz)	F ₀ Moy. (Hz)	Ecart- type F ₀ (Hz)	F _{0min} (Hz)	F _{0max} (Hz)	F ₁ (Hz)	F ₂ (Hz)	F ₃ (Hz)	Energie (dB)
Normal	129.36	129.08	0.911	127.92	131.41	646	1038	2806	73.13
Avant Réédu.	-	-	-	-	-	1089	1590	3100	57.90
3 mois	061.15	060.42	02.144	059.26	064.73	1099	1769	3311	59.30
6 mois	062.39	068.74	15.140	049.97	102.73	768	1870	3216	56.09
11 mois	104.01	099.03	09.955	094.41	113.46	970	1680	3051	58.62

Table 2. Valeurs de F₀, formants et énergie de la voyelle [ā].

Nous avons complété l'analyse acoustique par un jugement global de perception de la parole afin de mesurer l'acceptabilité de la parole après onze mois de rééducation. Ce jugement global est basé sur la perception subjective de l'auditeur en ce qui concerne la maîtrise et l'intelligibilité de la parole. Pour cela, nous avons utilisé un jury d'experts composé de trois locuteurs natifs algériens, sans aucune connaissance

préalable des mots et des phrases du corpus enregistré. Enfin, un regard sur les représentations sonographiques des sons prononcés par les patients permet de confirmer certains troubles de la prononciation de certaines consonnes (Figures 5 et 6).

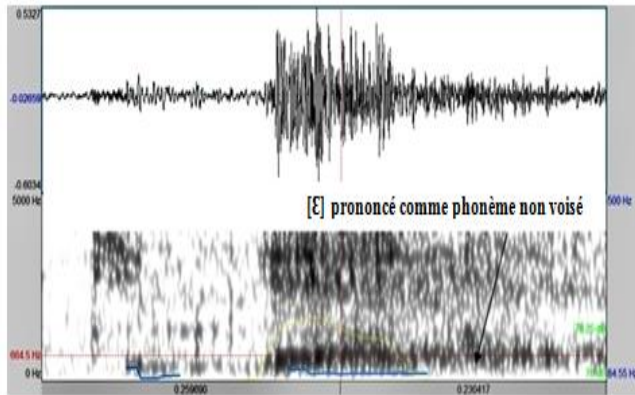


Figure 5: Prononciation du [ɛ] dans le mot [tbiɛ] (Elle vend).

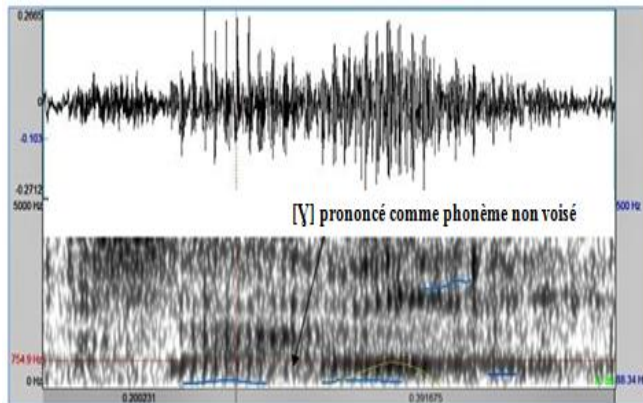


Figure 6 : Prononciation du [ʏ] dans le mot [saʏr] (Petit).

L'interprétation des résultats obtenus nous montre :

- un voisement qui commence à apparaître après trois mois de rééducation (> 60 Hz). Il reste que le pitch est relativement faible, même après onze mois de rééducation. La faible capacité de voisement ($F_0 \leq 100$ Hz) peut être expliquée par la forme, le volume et l'élasticité du segment néo-vibrateur NVP, qui diffèrent totalement de ceux des cordes vocales. Le segment NVP semble toujours non périodique, en

raison des caractéristiques anatomiques de ses structures vibrantes. Nous notons que ces résultats sont en accord avec quelques études précédentes rapportées dans d'autres langues [18] ;

- des valeurs du Jitter et du Shimmer tendant vers des valeurs de seuil normal. Ceci est probablement dû au fait qu'après une période de rééducation, le patient est en mesure de mieux contrôler ses efforts par une meilleure con-

naissance des divers organes de la nouvelle phonation ;

- une voix pathologique tellement bruitée au début, que le patient aura probablement besoin de plus de temps pour acquérir des valeurs HNR proches du seuil de normalité ;

- un nombre de trames non voisées (DUF) à l'intérieur d'une voyelle diminuant de façon significative au cours de la rééducation. Ceci montre que le timbre de la voix s'améliore progressivement au cours du temps ;

- une légère augmentation des valeurs des formants F_1 , F_2 et F_3 . Ceci peut être probablement expliqué par le fait que la distance entre le segment NVP et la première cavité de l'appareil phonatoire (cavité pharyngale) est modifiée. Ces résultats sont en accord avec ceux rapportés antérieurement dans d'autres langues [19, 20] ;

- une énergie moins importante après LT, car la quantité d'air obtenue par érucation reste insuffisante (moins de 70 ml) par rapport à celle qui résulte des poumons dans la parole normale laryngée (environ 5000 ml) ;

- une confusion entre certaines consonnes, en accord avec des études précédentes rapportées dans la littérature concernant la laryngectomie [21,22], contrairement à d'autres études qui rapportent un taux de confusion négligeable [23]. Dans notre étude, nous notons une confusion perceptible des consonnes postérieures ([ɛ] (ع), le [H] (ح), le [ɣ] (غ) et le [χ] (خ)) contrairement aux consonnes antérieures ([g], [k], [d], [t]) ;

- enfin, les patients sont unanimes à constater que la voyelle [i] est très difficile à prononcer, en particulier lors d'une parole continue. La voyelle antérieure [i] est articulée avec le devant de la langue très rapproché de la ré-

gion alvéolaire, entraînant une poussée vers l'avant de la masse de la langue.

Le pharynx inférieur se trouve ainsi élargi au cours de la prononciation du [i]. Avec les changements subis par cette partie du conduit vocal après LT, il est possible que ce soient les contraintes articulatoires qui font que le [i] est plus difficile à prononcer que les autres voyelles. Ainsi, la technique de rééducation doit faire un effort plus important vers une meilleure prononciation de cette voyelle.

4. Conclusion

En conclusion, l'analyse acoustique de la voix du patient avec TL montre que la majorité des paramètres acoustiques ont tendance à se rapprocher de la normalité de seuil après une période suffisante de rééducation. Ainsi, une prise en charge adéquate d'une LT permet une parole œsophagienne ES assez intelligible, pour permettre au patient de dialoguer avec son entourage, en dépit de l'ablation totale des cordes vocales. Il reste qu'au cours de notre étude, des lacunes ont été observées dans la prise en charge des patients algériens :

- absence de formation des orthophonistes dans la manipulation des logiciels d'analyse acoustique. Selon les résultats trouvés, il est important de renforcer les compétences et les connaissances de l'orthophoniste, ce qui lui permet d'adapter sa technique de réhabilitation en donnant plus d'importance au phénomène de voisement et aux consonnes postérieures de l'arabe. Dans notre étude, nous avons noté une confusion perceptible des consonnes postérieures (pharyngales [ɛ] (ع) / [H] (ح) et uvulaires [ɣ] (غ)/[χ] (خ)) contrairement aux consonnes antérieures (vélares [g]/[k] et dentales [d]/[t]). Ceci est très important à noter car les tests

de rééducation orthophoniques exploités dans les hôpitaux algériens sont importés de France et donc plus adaptés aux consonnes antérieures car la langue française présente peu de consonnes postérieures. Une coordination concrète des efforts entre le chercheur au laboratoire de recherche, l'enseignant au milieu universitaire et l'orthophoniste rééducateur au milieu hospitalier, pourra sûrement donner une meilleure approche de rééducation, donc une technique plus appropriée, adaptée au milieu socioculturel algérien ;

- l'utilisation exclusive et abusive de l'ouïe pour évaluer l'effet de la rééducation vocale dans le milieu hospitalier algérien. L'évaluation de la parole pathologique est essentiellement basée sur la perception subjective des rééducateurs orthophonistes sans aucune analyse acoustique ou articulatoire. Peu d'attention est accordée à l'analyse acoustique dans le cursus universitaire de l'étudiant, future orthophoniste à l'hôpital ;

- l'analyse acoustique ne pourra remplacer la voie classique de la rééducation mise en place par l'orthophoniste dans un milieu hospitalier, mais elle pourra l'aider dans l'amélioration de sa technique de rééducation et surtout l'évaluation périodique et objective de l'acquisition du nouveau mécanisme de phonation par le patient. De même, elle renforce et appuie le message auditif et rend objectif ce qui échappe parfois à l'oreille de l'orthophoniste rééducateur.

5. Références

- [1] Singer, M.I., Blom, E.D. and Hamaker, R.C., "Voice Rehabilitation after Total Laryngectomy", *J. Otolaryngol*, 12:329-334, 1983.
- [2] Koike, M., Kobayashi, N., Hirose, H. and Hara, Y., "Speech Rehabilitation after Total Laryngectomy", *Acta Otolaryngol*, 122: 107-112, 2002.
- [3] Stajner-Katusic, S., Horga, D., Musura, M. and Globek, D., "Voice and speech after laryngectomy", *Clin. Linguist. Phon.*, 20 (2-3):195-203, 2006.
- [4] Diedrich, W.M., "The mechanism of esophageal speech", *Annals of the New York Academy of Sciences*, 155: 303-317, 1968.
- [5] Duguay, M.J., "Esophageal voice: an historical review", *J. Voice*, 3: 264-268, 1989.
- [6] Hocevar-Boltezar, I. and Zargi, M., "Communication after laryngectomy", *Radiol. Oncol.*, 35: 249-254, 2001.
- [7] Elmiyeh, B., Dwivedi, R.C., Jallali, N., Chisholm, E.J., Kazi, R., Clarke, P.M. and Rhys-Evans, P.H., "Surgical voice restoration after total laryngectomy: An overview", *Indian J Cancer*, 47(3): 239-247, 2010.
- [8] Sato, T., "Oesophageal Speech and Rehabilitation of the Laryngectomized", Kanehara & Co., Ltd., Tokyo, 1993.
- [9] Kearney, A., "Esophageal Speech", *Oto laryngol. Clin. N. Am.*, 37:613-625, 2004.
- [10] Ferrat, K., "Evaluation acoustique d'une technique de rééducation de laryngectomie totale en milieu hospitalier algérien", *Cahiers du CRSTDLA*, ISSN: 1112-735x, ISBN: 978-9961-9660-1-5, 2, 2006.
- [11] Ferrat, K. and Guerti, M., "An Acoustical Study of sounds Produced by Algerian Esophageal Speakers", *African Health Sciences*, ISSN:1680-6905, EISSN: 1729-0503, 2012.
- [12] Ferrat, K. and Guerti, M., "Analyse acoustique et évaluation de la technique de rééducation du cas de laryngectomie totale en milieu clinique algérien", *Séminaire international sur les pathologies du langage. Evaluation et prise en charge en milieu clinique algérien*, CRSTDLA, 2012.

- [13] Yumoto, E., Gould, W. and Bear, T., "Harmonic-to-noise ratio as index of the degree of hoarseness", *J. Acoust. Soc. Amer.*, 71(3):1544-1550, 1982.
- [14] Heiberger, V.L. and Horii, Y., "Jitter and Shimmer in sustained phonation", *Speech and Language : Advances in Basic Research and Practice*, 7:299-332, 1982.
- [15] Munoz, J., Mendoza, E., Fresneda, M.D., Carballo, G. and Lopez, P., "Acoustic and perceptual indicators of normal and pathological voice", *Folia Phoniatr Logop*, 55:102-114, 2003.
- [16] Kreiman, J. and Gerratt, B.R., "Perception of aperiodicity in pathological voice", *J. Acoust. Soc. Am.*, 117:2201-2211, 2005.
- [17] Kay Elemetrics, Multi Dimensional Voice Program, Model 5105 Lincoln Park, NJ: Kay Elemetrics Corporation, 2008.
- [18] Liu, H., Wan, M., Wang, S., Wang, X. and Lu, C., "Acoustic characteristics of Mandarin esophageal speech", *J. Acoust. Soc. Amer.*, 118:1016-1025, 2005.
- [19] Sisty, N.L. and Weinberg, B., "Formant frequency characteristics of esophageal speech", *J. Speech Hear. Research*, 15: 439-448, 1972.
- [20] Cervera, T., Miralles, J.L. and González-Alvarez, J., "Acoustical analysis of Spanish vowels produced by laryngectomized subjects", *J. Speech Lang. Hear. Research*, 44:988-996, 2001.
- [21] Ng, M.L. and Chu, R., "An Acoustical and Perceptual Study of Vowels Produced by Alaryngeal Speakers of Cantonese", *Folia Phoniatr. Logop.*, 61:97-104, 2009.
- [22] Crevier-Buchman, L., Vaissière, J., Maeda, S. and Brasnu, D., "Etude de l'intelligibilité des consonnes du français après laryngectomie partielle supra cricoïdienne", *Rev. Laryngol. Otol. Rhinol.*, 123:307-310, 2002.
- [23] Hirose, H., "Voicing Distinction in esophageal Speech", *Acta Otolaryngol. suppl.*, 524:56-63, 1996.