

Proposition aux 7èmes Journées de Phonétique Clinique, Paris

Titre : Etude de paramètres acoustiques des voix de patients traités pour un cancer ORL dans le cadre du projet C2SI

Title: Study of acoustic parameters of the voice of patients treated for Oropharyngeal Cancer within the C2SI project

Auteurs : E. Sicard(1), J. Mauclair(2), V. Woisard(3)

Affiliation :

(1) INSA de Toulouse, Université de Toulouse, France

(2) IRIT-Samova, Université de Toulouse, France

(3) IUCT - Oncopole Toulouse, Octogone-Lordat, Université Jean-Jaurès, France

Résumé :

Dans le cadre du projet “*Carcinologic Speech Severity Index*” (C2SI), les voix de 35 patients ayant été traités pour un cancer oral et pharyngé ont été étudiées selon différents paramètres acoustiques, couplée à la reconnaissance vocale automatisée. L’étude révèle, par comparaison avec un groupe témoin, une pauvreté harmonique significative sur un /a/ tenu ainsi que sur du texte lu, ainsi qu’une dégradation spectrale des voyelles et des consonnes, notamment dans les zones des formants F2 à F4, corrélée avec l’intelligibilité et les erreurs d’alignement sur un ensemble de pseudo-mots. L’étude ne révèle cependant pas de différence significative de l’histogramme du fondamental de la voix pour un texte lu, entre le groupe témoin et le groupe patients.

Mots clé : cancer, voix, indicateurs, LTASS, segmentation, pauvreté harmonique

Abstract:

As part of the Carcinologic Speech Severity Index project (C2SI), the voices of 35 patients who were treated for oral and pharyngeal cancer were studied using different acoustic parameters, coupled with automated speech recognition. The study revealed, by comparison with a control group, significant harmonic poverty on sustained /a/ as well as on read text, as well as a spectral degradation of vowels and consonants, in particular in the areas of formants F2 to F4, correlated with intelligibility and misalignment on a set of pseudo-words. However, the study does not reveal any significant difference in the histogram of the fundamental of the voice between the control and patient groups.

Keywords : Cancer, voice, indicators, LTASS, segmentation, harmonic poverty

Introduction

Dans le cadre du projet C2SI [1], la voix de différents patients ayant été traités pour un cancer oral et pharyngé de stade T2 à T4 (selon la classification TNM) par chirurgie, radiothérapie et/ou chimiothérapie, a été enregistrée au moins 6 mois après la fin du traitement. Un groupe de locuteurs sains a aussi été constitué afin de servir de témoin. Dans le but de déterminer des modèles prédictifs d’intelligibilité et de compréhension sur la base de l’analyse automatique de la parole, nous avons étudié les caractéristiques des /a/ tenus, du texte lu pour chaque patient et témoin, afin de tenter de dégager des tendances.

Méthode

L'enregistrement de la voix a été conduit selon un protocole et avec un matériel identique, comprenant notamment un /a/ tenu, la lecture d'un texte extrait de la nouvelle « La chèvre de M. Seguin », d'Alphonse Daudet, la reproduction d'un modèle sonore de 50 pseudo-mots aléatoires, et différentes autres tâches telles que la description de scènes. Les fichiers sonores issus des enregistrements ont été homogénéisés et regroupés dans une base de données en ligne, partagée par les membres du projet. Différentes données patients et scores d'intelligibilité ont aussi été mis à disposition. 45 enregistrements ont été utilisés dans cette étude, répartis en 35 patients et 10 témoins. L'analyse des /a/ tenus et du texte lu a été conduite en utilisant deux types d'outils : l'un ciblé sur le bilan et la rééducation orthophonique VOCALAB [2], l'autre en utilisant la librairie OpenSmile [3].

Résultats

Analyse du /a/

La limite normal/pathologique étant à 1 pour les 5 indicateurs, on observe une altération significative de la pauvreté harmonique (figure 1), avec une moyenne à 1.76. Les données témoins sont dans la norme.

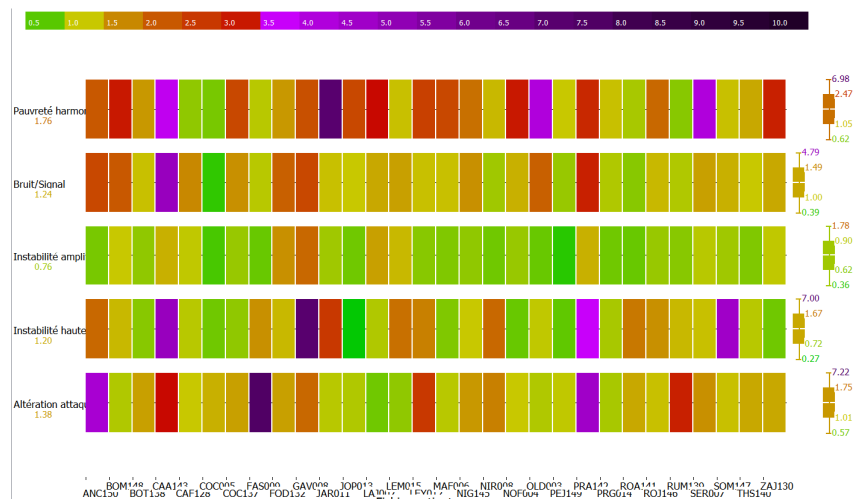


Figure 1 : Indicateurs d'altération de la voix des /a/ tenus [2]

Analyse du texte lu

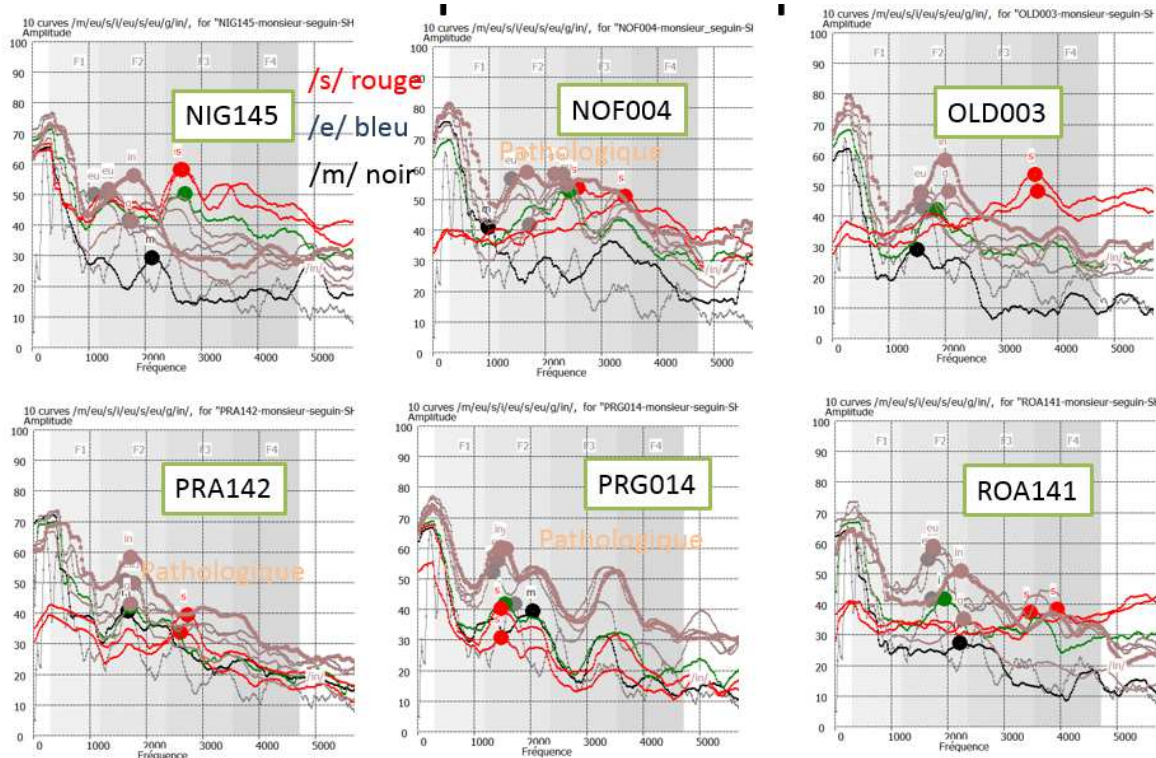
L'histogramme du fondamental sur un texte lu permet d'extraire l'excursion typique du F0 exprimé en notes. La comparaison entre le groupe patient et témoin ne révèle pas de différence significative, les moyennes étant autour de 8 notes, au-dessus de la norme (6 notes).

Analyse long terme du spectre

Le profil typique du spectre moyenné sur chaque groupe semble confirmer l'observation sur le /a/, à savoir un appauvrissement harmonique dans la zone F3-F4. Le nombre réduit de fichiers témoins disponibles ne permet toutefois pas de conclure quant à la pertinence de la comparaison.

Analyse des voyelles et consonnes

L'alignement manuel d'une portion du texte lu a permis le calcul du spectre moyen d'un ensemble de consonnes et voyelles. Celui-ci est normalement bien différenciés en fréquence et en énergie, avec des extrêmes telles que le /m/ et le /s/ (noir et rouge sur la Figure 2). Les voix pathologiques se distinguent non seulement par la pauvreté harmonique déjà observée sur le /a/, mais aussi une incapacité à différencier les formants F3-F4 (NOF004, PRA142..) tandis que 30% des patients ont des spectrogrammes proches des voix normales (NIG145, OLD003..).



Discussion

L'analyse des voix de patients ayant été atteints et traités pour un cancer ORL a permis de dégager quelques tendances : une pauvreté harmonique marquée associée à une instabilité en hauteur et intensité un peu élevée, un histogramme du fondamental sur un texte lu normal, et une faible différenciation fréquentielle entre voyelles et consonnes pour plus de 50% des cas, notamment dans la zone F3-F4, laissant craindre des confusions dans la perception des pseudo-mots et une perte d'intelligibilité.

Références :

- [1] <https://www.irit.fr/recherches/SAMOVA/pagec2si.html>
- [2] Menin-Sicard A., Sicard E., (2016) "Evaluation et réhabilitation de la voix - Approche clinique et objective", De Boeck Supérieur, ISBN 9782353273188, 288 pp.
- [3] Gloinec B., (2016). "Development of Methods to Characterize the Intelligibility of Patients Recovering from Throat Surgery", Master 2 report, INP-ENSEEIH Toulouse, Sept. 2016
- [4] Leino, T. (2009), "Long-Term Average Spectrum in Screening of Voice Quality in Speech", Journal of Voice, Volume 23, Issue 6, pp. 671–676