



**HAL**  
open science

## L'identification automatique des différents bruits de gorge chez le sujet sain : une étude pilote

Yohan Gallois, Jérôme Farinas, Mari-Wenn Paugam, Linda Nicolini, Virginie Woisard

► **To cite this version:**

Yohan Gallois, Jérôme Farinas, Mari-Wenn Paugam, Linda Nicolini, Virginie Woisard. L'identification automatique des différents bruits de gorge chez le sujet sain : une étude pilote. Séminaire AFCP 2021 – Phonétique Clinique, May 2021, Toulouse (virtuel), France. . hal-03269247

**HAL Id: hal-03269247**

**<https://ut3-toulouseinp.hal.science/hal-03269247>**

Submitted on 23 Jun 2021

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

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

# L'identification automatique des différents bruits de gorge chez le sujet sain : une étude pilote.

Yohan GALLOIS<sup>1,2</sup>  
Jérôme FARINAS<sup>3</sup>  
Mari-Wenn PAUGAM<sup>4</sup>  
Linda NICOLINI<sup>5</sup>  
Virginie WOISARD<sup>2,4,6</sup>

<sup>1</sup>*Service ORL Hôpital Pierre Paul Riquet, Hôpitaux de Toulouse, France*

<sup>2</sup>*Laboratoire Octogone-Lordat, Université Toulouse II, France*

<sup>3</sup>*Institut de Recherche en Informatique de Toulouse, CNRS, Université Toulouse III, France*

<sup>4</sup>*Centre de formation Universitaire en Orthophonie, Faculté de Médecine, Université Toulouse III, France*

<sup>5</sup>*Swallis Medical, Toulouse France*

<sup>6</sup>*Service ORL Hôpital Larrey, Hôpitaux de Toulouse, France*

Le carrefour aérodigestif coordonne plusieurs fonctions volontaires, spontanées et réflexes, rendant particulièrement complexe la compréhension de son contrôle neurologique et l'interprétation des explorations des troubles fonctionnels impliquant cette région [1]. Peu de travaux ont une approche holistique des fonctions du carrefour, leur intrication devant être compatible avec l'efficacité de chacune d'entre elles [2].

Un dispositif non-invasif permettant d'identifier, quantifier et décrire automatiquement les différents événements survenant au niveau de la gorge s'avèrerait intéressant pour étudier les interactions entre ses fonctions vitales et la gestion de la parole [3].

L'objectif de ce travail est de savoir si un tel système automatique peut identifier les différents bruits de gorge chez un sujet sain.

Onze sujets sains (sex-ratio H/F=1,2) ont participé à l'étude, avec un âge moyen de 42,8ans +/-15,7. Nous avons mesuré leur activité motrice

cervicale sous forme de fichiers audios, et de capteurs vibro-acoustiques montés sur collier cervical, lors de 14 tâches différentes de phonation (voyelle tenue, mot bisyllabique, phrase courte), déglutition (salive, liquide, yaourt), mécanismes de protection (toux, hémage) et respiration volontaires (respirations forcées), sous contrôle vidéo anonymisé. Les fichiers ont été découpés selon l'ordre des tâches par deux investigateurs différents.

Ces segments audios ont permis de modéliser les différents événements par Modèle de Markov Cachés (toolbox HTK [4]). Après une phase d'initiation par Viterbi, une réestimation par Baum-Welch est réalisée. La reconnaissance Viterbi a été effectuée avec une grammaire libre ou contrainte.

Les résultats de détection des algorithmes ont été interprétés avec les matrices de confusion, le taux de détections correctes, et la précision qui prend en compte les erreurs d'insertions.

Avec une grammaire libre, l'algorithme actuel présente un taux de détections correctes de 83,44% mais une précision moins bonne (-220,86% du fait du grand nombre d'insertions). Avec une grammaire contrainte, le taux de réponses correctes est affaibli (77,15%) mais la précision améliorée (56,62%).

Les événements les mieux détectés sont les déglutitions d'eau, les productions du mot ou de la phrase. La mauvaise précision est liée surtout aux insertions attribuées aux événements spontanés (respiration, déglutition de salive), favorisé par la variabilité inter-investigateur du séquençage des fichiers audios, notamment des événements spontanés. Ces premiers résultats nous encouragent dans le développement de tests standardisés en utilisant une grammaire contrainte. La prise en compte des capteurs de mouvements devrait permettre de renforcer la robustesse de l'algorithme. Disposer de plus d'enregistrements permettrait également d'envisager d'autres méthodes d'apprentissage, par exemple des réseaux profonds récurrents.

[1] LUDLOW, CL. Central Nervous System Control of Voice and Swallowing (2015). *Journal of Clinical Neurophysiology*, vol 32, p. 294-303.

[2] LEE, K., NI, X., LEE, JY, ARAFA H., PE, DJ., XU, S., et al (2020). Mechano-acoustic sensing of physiological processes and body motions

via a soft wireless device placed at the suprasternal notch. *Nat Biomed Eng.* vol. 4, p.148-58.

[3] MAO, S., ZHANG, Z., KHALIFA, Y., DONOHUE, C., J.L. COYLE, SEJDIC, E. Neck sensor-supported hyoid bone movement tracking during swallowing. Royal Society Open Science. *Royal Society* vol. 6, no 7, p. 181982.

[4] YOUNG, S.J., YOUNG, S.. The HTK hidden Markov model toolkit: Design and philosophy. University of Cambridge, Department of Engineering Cambridge, 1993