

# Étude des erreurs de transcription automatique pour la détection de la somnolence à long terme de patients hypersomniaques

Vincent P. MARTIN<sup>1</sup>

Jean-Luc ROUAS<sup>1</sup>

Pierre PHILIP<sup>2</sup>

<sup>1</sup>LaBRI,  
UMR 5800 CNRS,  
Bordeaux INP,  
Université de Bordeaux  
<sup>2</sup>SANPSY,  
USR 3413 CNRS  
CHU de Bordeaux  
Université de Bordeaux

**Contexte** : Lors de précédents travaux basés sur le corpus TILE enregistré au CHU de Bordeaux contenant les enregistrements de 93 patients atteints de diverses formes d'hypersomnolence (Martin et al., 2020b), nous avons montré qu'à la fois des marqueurs acoustiques de la voix (Martin et al., 2019) et les erreurs de lecture faites lors de la lecture de textes (Martin et al., 2020a) permettent d'estimer la somnolence subjective à court terme (mesurée par la Karolinska Sleepiness Scale - KSS), la somnolence diurne excessive (mesurée par l'Epworth Sleepiness Scale - ESS) et la somnolence objective (mesurée par la latence d'endormissement au test itératif de latence d'endormissement - TILE) du lecteur.

**Méthode** : Notre hypothèse est qu'avec la somnolence, l'articulation et la qualité de la diction diminuent, tandis que le nombre d'hésitations et de reprises augmente. Dans cette première approche, nous proposons de mesurer le pouvoir discriminant du nombre d'erreurs (insertions, délétions et substitutions) faites par un

système de transcription automatique (STA) pour estimer la somnolence. Le système utilisé, décrit dans (Boyer and Rouas, 2019), est un système end-to-end basé sur des caractères utilisant des transducteurs RNN. Afin d'assurer la validité externe des marqueurs sélectionnés, nous ne conservons que ceux ne corrélant pas ( $\rho$  de Spearman) avec l'âge, l'IMC, le tour de cou, le niveau social (années d'études après le brevet), l'anxiété ou la dépression (HAD) du patient, et ne discriminant pas son sexe (U de Mann-Whitney) ou sa pathologie (ANOVA univariée).

**Résultats :** Le pourcentage d'insertions à chaque session corrèle avec le KSS moyen (Figure 1) et discrimine les classes somnolent - S - (KSS moyen  $< 5$ ) et non-somnolent - NS - de manière significative (Tableau 1, Figure 2A). Le nombre moyen d'insertion par session est efficace pour distinguer les patients souffrant de somnolence excessive diurne (ESS  $> 10$ ) des autres : le nombre d'insertions semble lié à la somnolence subjective du patient. Par ailleurs, le nombre moyen de substitutions par session discrimine les patients atteints de somnolence diurne excessive objective (latence moyenne du TILE  $\leq 8$  minutes), et semble donc refléter l'impact de la somnolence sur l'activité cérébrale, indépendamment du ressenti du patient. Ces marqueurs ne sont cependant pas suffisants pour classifier à eux seuls les classes Somnolent et Non-Somnolent (Figure 2B), mais sont prometteurs pour compléter et améliorer les approches existantes.

Tableau 1. Erreurs du système de transcription automatique utilisés comme biomarqueurs de la somnolence

RÉGRESSION			
Mesure de la somnolence	Biomarqueur	Statistique ( $\rho$ de Sperman ou U de Mann-Whitney)	Sig.
KSS moy	Nombre d'insertions / Nombre de mots détectés	$\rho = 0.21$ $p=0.037$	*
DISCRIMINATION SOMNOLENT / NON SOMNOLENT			
TILE moy	Nombre de substitutions	$U = 547.5$ $p=0.044$	*
KSS moy	Nombre d'insertions	$U = 658$ , $p=0.02$	*
ESS	Nombre d'insertions	$U = 330$ , $p=0.008$	**

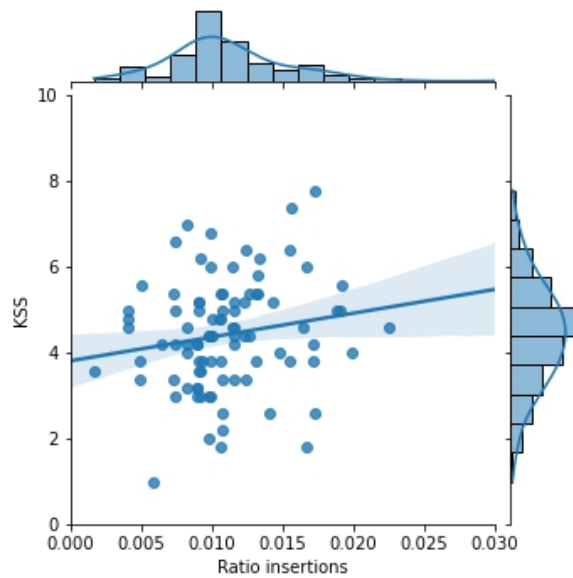


Figure 1. Corrélation entre le nombre moyen d'insertions par locuteur et le KSS moyen

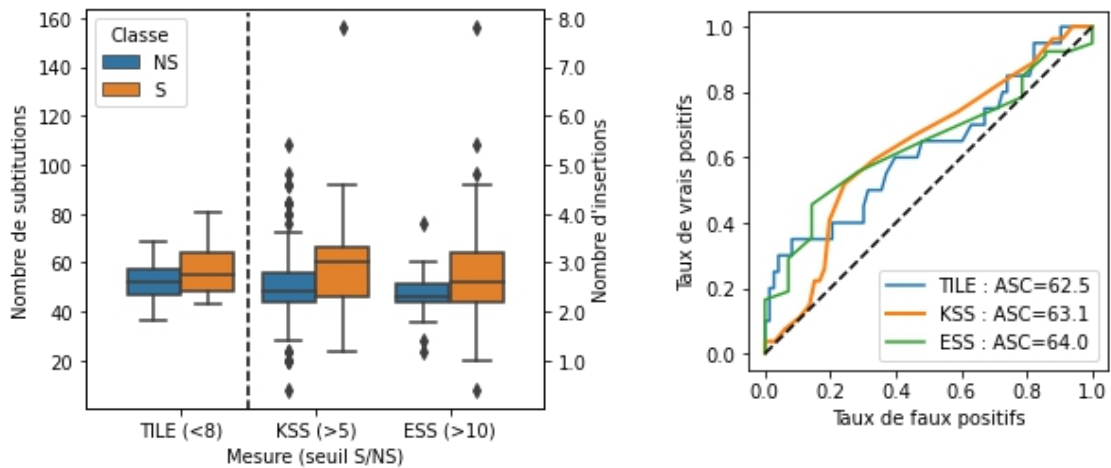


Figure 2. (A) Moyenne et écart-type des biomarqueurs discriminant les deux classes somnolent (S) et non-somnolent (NS). (B) Courbes ROC associées aux marqueurs discriminant les classes de somnolence. ASC : aire sous la courbe

## Références bibliographiques

BOYER, F., ROUAS, J.-L., End-to-End Speech Recognition: A review for the French Language. ArXiv191008502 Cs Eess, 2019.

<http://arxiv.org/abs/1910.08502>

MARTIN, V.P., CHAPOUTHIER, G., RIEANT, M., ROUAS, J.-L., PHILIP, P., Using reading mistakes as features for sleepiness detection in speech, in 10TH INTERNATIONAL CONFERENCE ON SPEECH PROSODY, 2020a, Tokyo, 985–989. DOI : [10.21437/SpeechProsody.2020-201](https://doi.org/10.21437/SpeechProsody.2020-201)

MARTIN, V.P., ROUAS, J.-L., MICOULAUD-FRANCHI, J.-A., PHILIP, P., The Objective and Subjective Sleepiness Voice Corpora, in 12TH LANGUAGE RESOURCES AND EVALUATION CONFERENCE, Marseille, 2020b, 6525-6533. <https://www.aclweb.org/anthology/2020.lrec-1.803.pdf>

MARTIN, V.P., ROUAS, J.-L., THIVEL, P., KRAJEWSKI, J., Sleepiness detection on read speech using simple features, in 10TH CONFERENCE ON SPEECH TECHNOLOGY AND HUMAN-COMPUTER DIALOGUE. Timisoara, 2019. DOI : [10.1109/SPED.2019.8906577](https://doi.org/10.1109/SPED.2019.8906577)