



HAL
open science

Mesure de l'intelligibilité après cancer oral ou oropharyngé par un système de reconnaissance automatique de la parole

Mathieu Balaguer, Lucile Gelin, Virginie Woisard, Jérôme Farinas, Julien Pinquier

► To cite this version:

Mathieu Balaguer, Lucile Gelin, Virginie Woisard, Jérôme Farinas, Julien Pinquier. Mesure de l'intelligibilité après cancer oral ou oropharyngé par un système de reconnaissance automatique de la parole. 9ème Colloque inter-régional Recherche paramédicale Grand Sud-Ouest (2022), Sep 2022, Poitiers, France. , 2022. hal-03837679

HAL Id: hal-03837679

<https://hal.science/hal-03837679>

Submitted on 3 Nov 2022

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.



Distributed under a Creative Commons Attribution 4.0 International License

Mesure de l'intelligibilité après cancer oral ou oropharyngé par un système de reconnaissance automatique de la parole

Mathieu Balaguer ^(1,2), Lucile Gelin ⁽¹⁾, Virginie Woisard ^(2,3), Jérôme Farinas ⁽¹⁾, Julien Pinquier ⁽¹⁾

(1) IRIT, Université de Toulouse, CNRS, Toulouse INP, UT3, Toulouse, France – (2) CHU Larrey, Toulouse, France – (3) Laboratoire de Neuro-Psycho-Linguistique LNPL, Université Toulouse II, France

Introduction

Cancers de la cavité buccale et de l'oropharynx

- **Dégradation** des capacités de production de parole, dont l'**intelligibilité** [1,2,3]

Évaluation clinique actuelle majoritairement perceptive, mais

- **Manque de standardisation** des outils de mesures [4]
- **Variabilité importante des scores** liée à la subjectivité de l'évaluation [5]

Développement des outils d'analyse automatique de la parole

- Scores souvent issus d'**analyses de sons** de la parole (phonèmes) **isolés** [6]
- **Pas d'étude menée sur la parole spontanée**, pourtant la plus proche des productions quotidiennes des sujets [7]

Objectif

Prédire l'intelligibilité de la parole après traitement d'un cancer oral ou oropharyngé au moyen de scores issus d'un **système de reconnaissance automatique** de la parole.

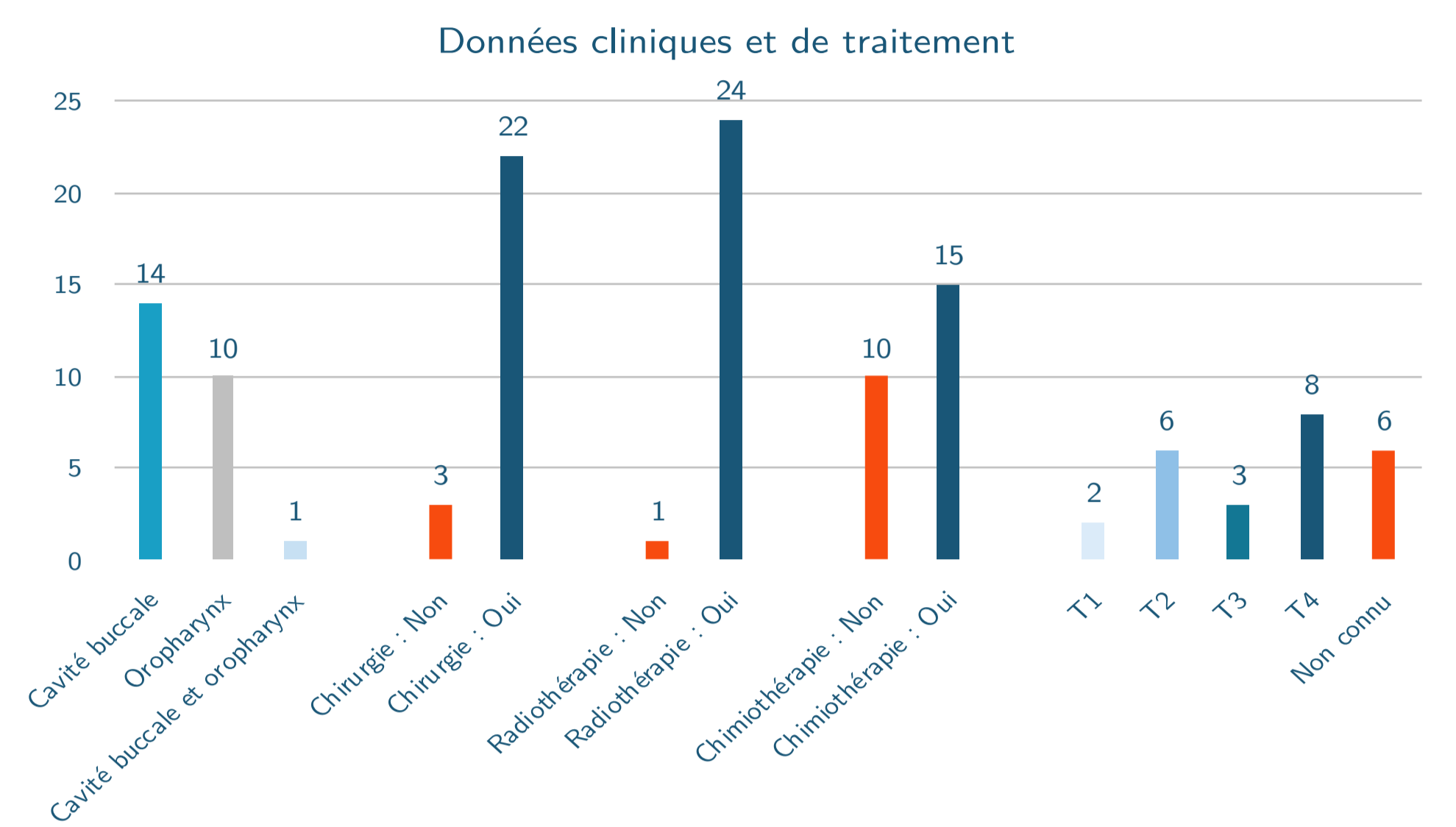
Matériel et méthode

25 patients

- 60 % hommes
- Âge médian : 67 ans, EIQ 12
- Délai moyen depuis la fin du traitement : 87 mois, écart-type : 121

Enregistrement de la parole spontanée des sujets au cours d'un entretien semi-dirigé

- Équivalent à une situation de « véritable » parole spontanée [8]



Analyse perceptive

Détermination du **score de référence perceptif d'intelligibilité**

- 6 auditeurs experts
- Évaluation de 0 (intelligibilité très dégradée) à 10 (absence de trouble)

Analyse automatique

Système de reconnaissance automatique de la parole adapté à la reconnaissance de phonèmes en parole atypique de type **Time-Delay Neural Network** [9,10]

- Entraînement sur près de 150 h d'enregistrements de parole saine (Corpus Common Voice)
- Décodage : détection parmi 33 phonèmes du français, score de confiance associé à chaque phonème reconnu

Résultats

1 Analyse perceptive de l'intelligibilité

Min-Max = 1,17 - 10 // Moyenne (SD) = 6,87 (2,71) // Médiane (EIQ) = 7,17 (4,34)

2 Analyse automatique

Extraction de 22 paramètres automatiques

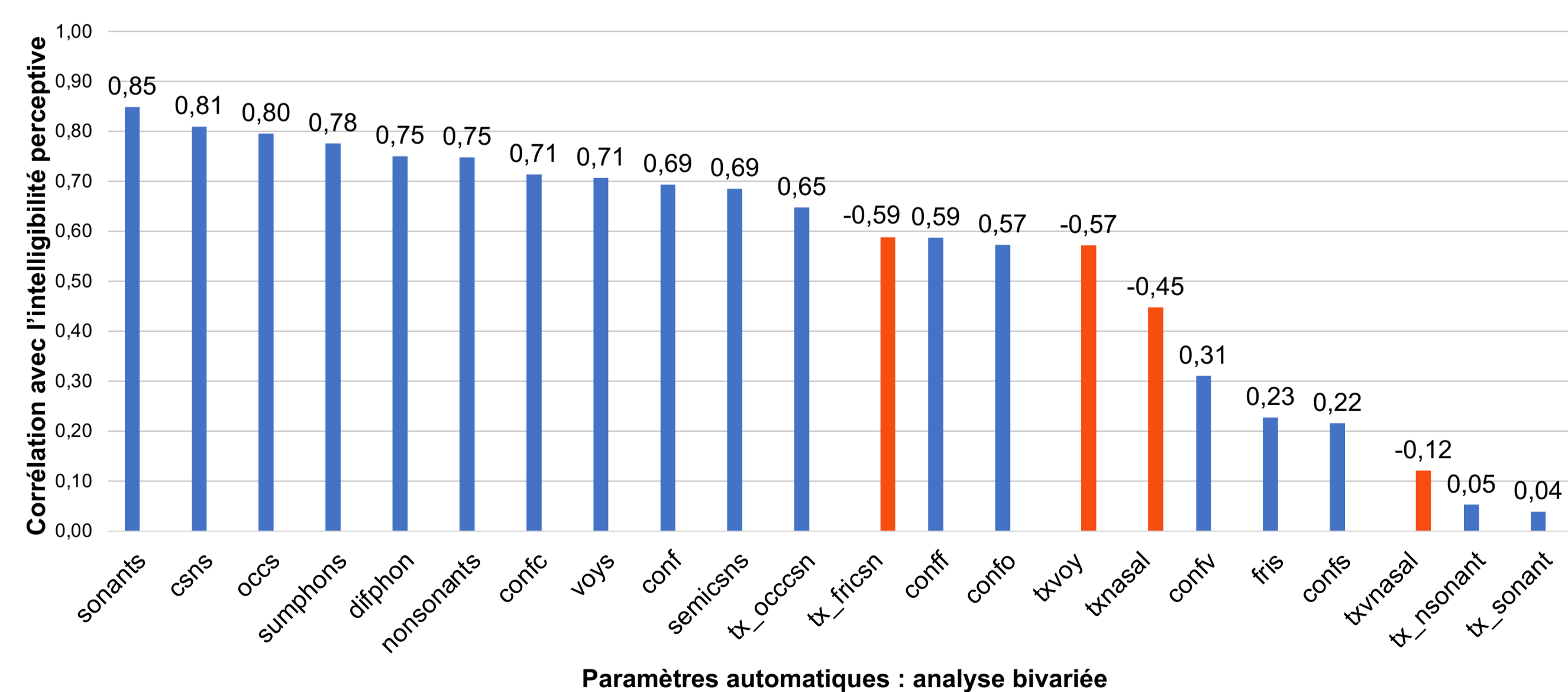
Inventaire phonémique

- Nombre de phonèmes reconnu par seconde (1)
- Nombre total de phonèmes différents reconnus (1)
- Nombre de consonnes, de voyelles, de semi-consonnes, d'occlusives, de fricatives, de sonantes, de non sonantes par seconde (7)
- Proportion de chaque type (7)

Scores de confiance

- Score de confiance moyen global (1)
- Score de confiance moyen sur les consonnes, les voyelles, les semi-consonnes, les occlusives et les fricatives (5)

3 Modélisation



Paramètres automatiques : analyse bivariée

Modélisation par **régression LASSO**

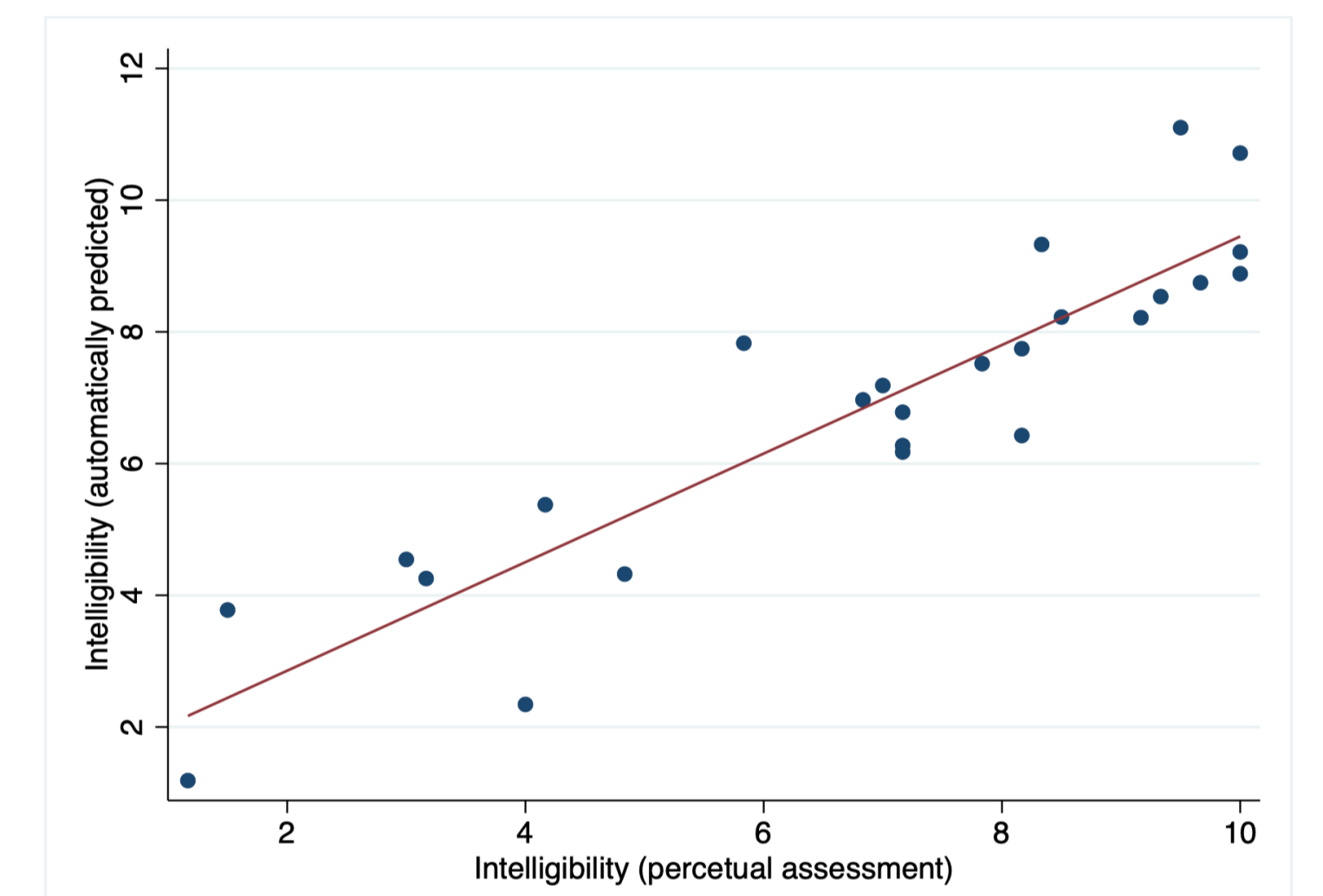
3 variables explicatives retenues :

- ▶ proportion d'occlusives parmi les consonnes (tx_occsn)
- ▶ nombre de sonantes par seconde ($sonants$)
- ▶ score de confiance moyen sur les fricatives ($confff$)

$$\text{Intelligibilité} = -0,073 + (6,188 \times tx_occsn) + (4,982 \times sonants) + (0,851 \times confff)$$

4 Analyse des performances

- **Corrélation** entre intelligibilité perceptive et intelligibilité prédite automatiquement : $r_s = 0,91$ ($p < 0,001$)
- Validation croisée
 - ▶ 5 blocs : $r_s = 0,90$ ($p < 0,001$)
 - ▶ Leave one out : $r_s = 0,90$ ($p < 0,001$)



Discussion

Prédiction automatique de l'intelligibilité correcte et fiable à partir de 3 paramètres issus d'une analyse automatique de la parole spontanée par un système de reconnaissance automatique de la parole.

Perspectives

- **Élargir la taille de l'échantillon** : généralisabilité des résultats
- **Considérer les mesures automatiques** dans la prédiction des altérations de la communication et de la **qualité de vie** [11]
- **Développer l'applicabilité clinique courante** : **développement de dispositifs mobiles** sur tablette en cours d'expérimentation au CHU de Toulouse

Références

- [1] Mlynarek, A., Rieger, J., Harris, J., O'Connell, D., Al-Qahtani, K., Ansari, K., Chau, J., & Seikaly, H. (2008). Methods of functional outcomes assessment following treatment of oral and oropharyngeal cancer: review of the literature. *Journal of Otolaryngology - Head & Neck Surgery*, 37(1), 2–10. <https://doi.org/10.2310/7070.2008.1001>
- [2] Borggren, P. A., Verdonck-De Leeuw, I. M., Muller, M. J., Heijligers, M. L. C. H., De Bree, R., Aaronson, N. K., & Leemans, C. R. (2007). Quality of life and functional status in patients with cancer of the oral cavity and oropharynx: Pretreatment values of a prospective study. *European Archives of Oto-Rhino-Laryngology*, 264(6), 651–657. <https://doi.org/10.1007/s00405-007-0249-5>
- [3] Pommée, T., Balaguer, M., Mauclair, J., Pinquier, J., & Woisard, V. (2021). Intelligibility and comprehensibility : A Delphi consensus study. *International Journal of Language & Communication Disorders*, 57(1), 21–41. <https://doi.org/10.1111/1460-6984.12672>
- [4] Pommée, T., Balaguer, M., Mauclair, J., Pinquier, J., & Woisard, V. (2021a). Assessment of adult speech disorders: current situation and needs in French-speaking clinical practice. *Logopedics Phoniatrics Vocology*, 0(0), 1–15. <https://doi.org/10.1080/14015439.2020.1870245>
- [5] Pommée, T., Balaguer, M., Mauclair, J., Pinquier, J., & Woisard, V. (2021b). Intelligibility and comprehensibility: A Delphi consensus study. *International Journal of Language & Communication Disorders*, 1–44. <https://doi.org/10.1111/1460-6984.12672>
- [6] Fex, S. (1992). Perceptual evaluation. *Journal of Voice*, 6(2), 155–158.
- [7] Balaguer, M., Pommée, T., Farinas, J., Pinquier, J., Woisard, V., & Speyer, R. (2020). Effects of oral and oropharyngeal cancer on speech intelligibility using acoustic analysis: Systematic review. *Head & Neck*, 42(1), 111–130. <https://doi.org/10.1002/hed.25949>
- [8] Knuijt, S., Kalf, J. G., van Engelen, B. G. M., de Swart, B. J. M., & Geurts, A. C. H. (2017). The Radboud Dysarthria Assessment: Development and Clinimetric Evaluation. *Folia Phoniatrica et Logopaedica*, 69(4), 143–153. <https://doi.org/10.1159/000484556>
- [9] Prins, R., & Bastiaanse, R. (2004). Analysing the spontaneous speech of aphasic speakers. *Aphasiology*, 18(12), 1075–1091. <https://doi.org/10.1080/02687030444000534>
- [10] Povey, D., Cheng, G., Wang, Y., Li, K., Xu, H., Yarmohammadi, M., & Khudanpur, S. (2018). Semi-Orthogonal Low-Rank Matrix Factorization for Deep Neural Networks. *Interspeech 2018, 2018-Sept(2)*, 3743–3747. <https://doi.org/10.21437/Interspeech.2018-1417>
- [11] Gelin, L., Daniel, M., Pinquier, J., & Pellegrini, T. (2021). *End-to-end acoustic modelling for phone recognition of young readers*. <https://www.lalilo.com/>
- [12] Balaguer, M. (2021). *Mesure de l'altération de la communication par analyses automatiques de la parole spontanée après traitement d'un cancer oral ou oropharyngé* [Thèse de doctorat, Université Paul Sabatier Toulouse III]. <http://www.theses.fr/2021TOU30109/document>