

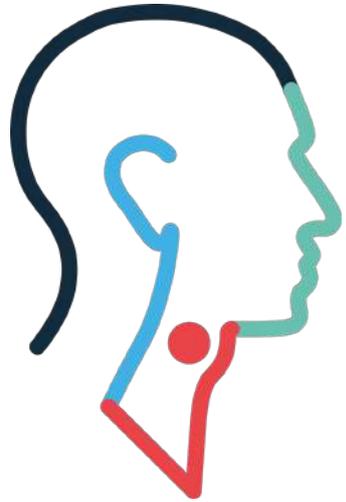
55^e
congrès
SFCCF | Société
Française de
Carcinologie
Cervico-Faciale
2023

16 & 17
novembre 2023

NANTES
Halle 6

**Mesure du trouble
de parole dans le suivi
cancérologique
grâce à la reconnaissance
automatique de la parole**

●
**Mathieu Balaguer,
Julien Piquier, Jérôme Farinas,
Virginie Woisard**



55^e
congrès
SFCCF Société Française de
2023 Carcinologie Cervico-Faciale

● CONTEXTE

- **Troubles de la parole** après cancer de la cavité buccale et de l'oropharynx

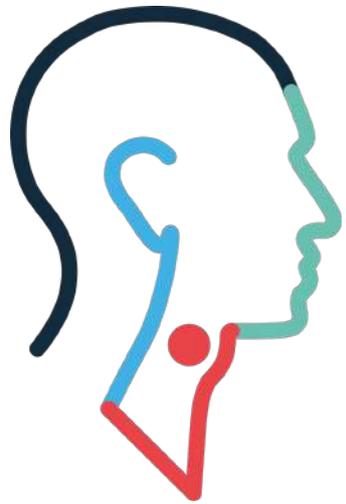
- Symptôme **le plus fréquent** (*Borggreven et al., 2007*)
- **Enjeu** thérapeutique majeur

- **Évaluation clinique** du trouble de parole

Sévérité du trouble

Altération
de la communication

- Majoritairement **perceptive**
(*Middag et al., 2014 ; Pommée et al., 2022*)
- **Variable**
(*Kim et al., 2011 ; Kuo & Tjaden, 2016*)
- Score holistique de communication (SHC)
(*Balaguer, 2021*)
- **Auto-questionnaires**

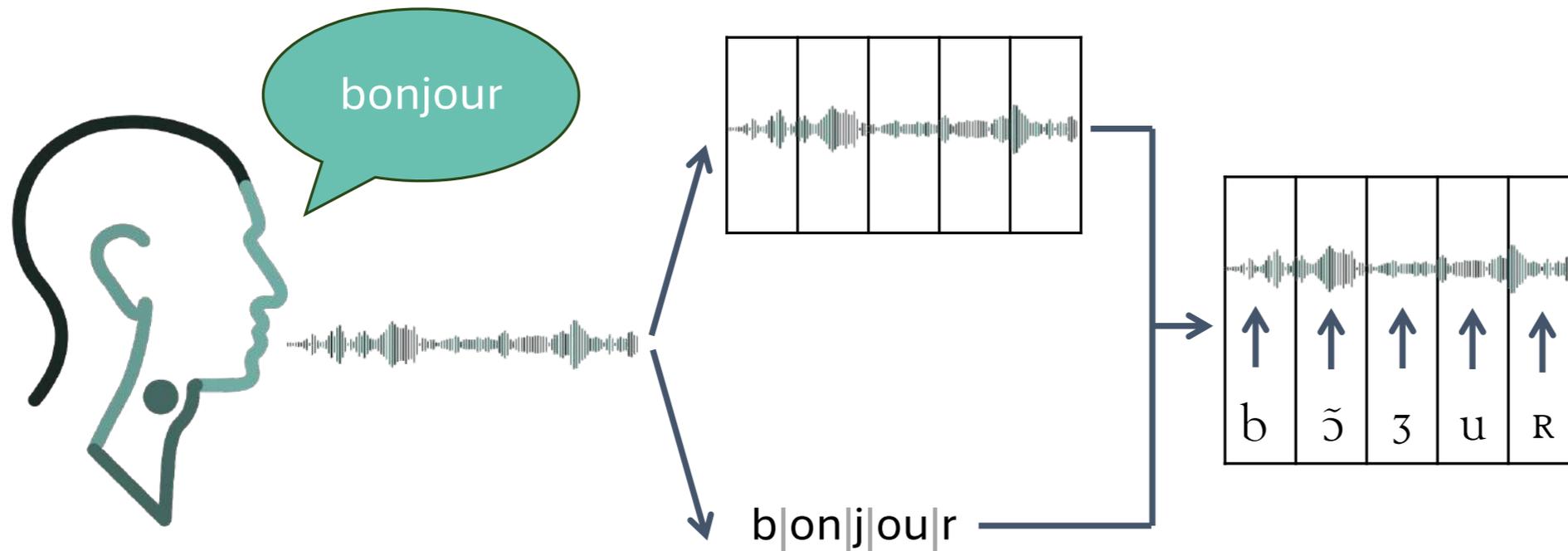


● CONTEXTE

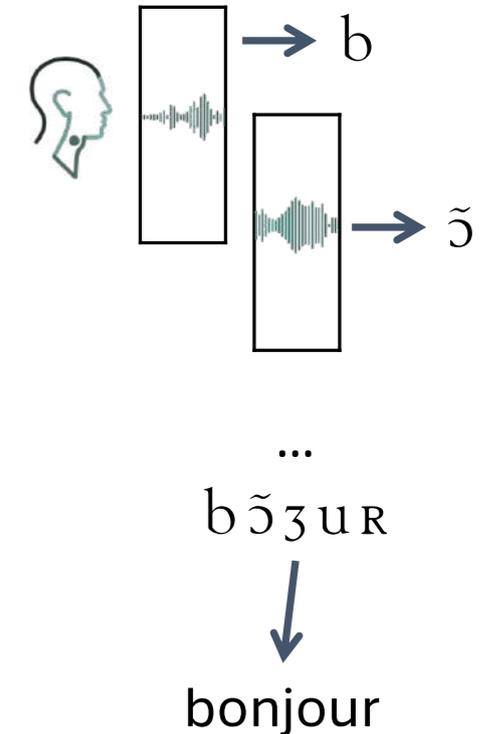
- Développement des outils d'**analyse automatique**
 - Notamment des systèmes de RAP : **reconnaissance automatique de la parole** (Li, 2021)
 - Avantages : simples, fiables, déterministes
 - Fonctionnement schématique

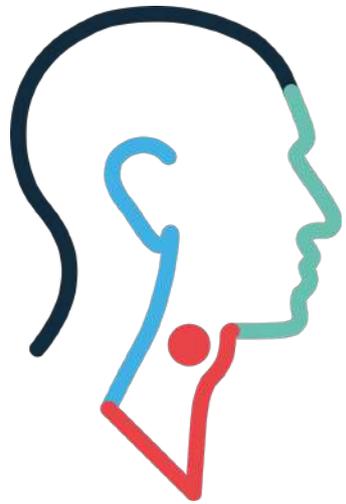
55^e
congrès
SFCCF Société Française de
Carcinologie
Cervico-Faciale
2023

Entraînement



Décodage

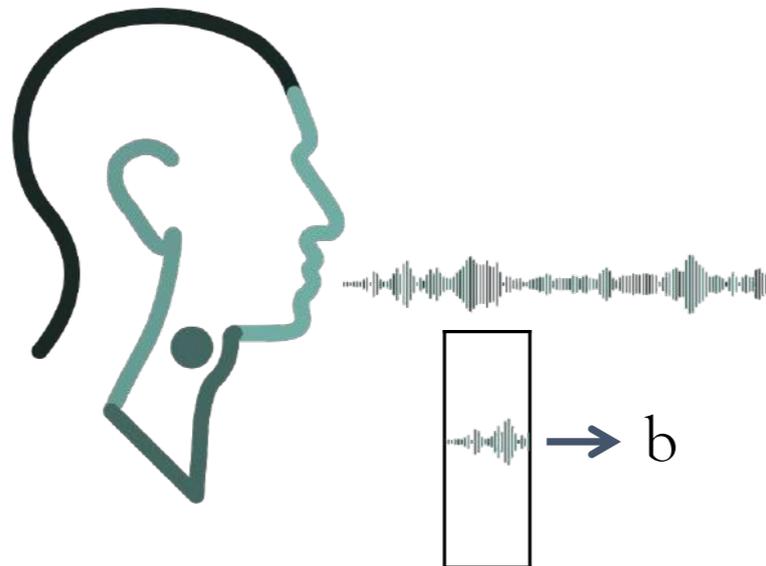




● CONTEXTE

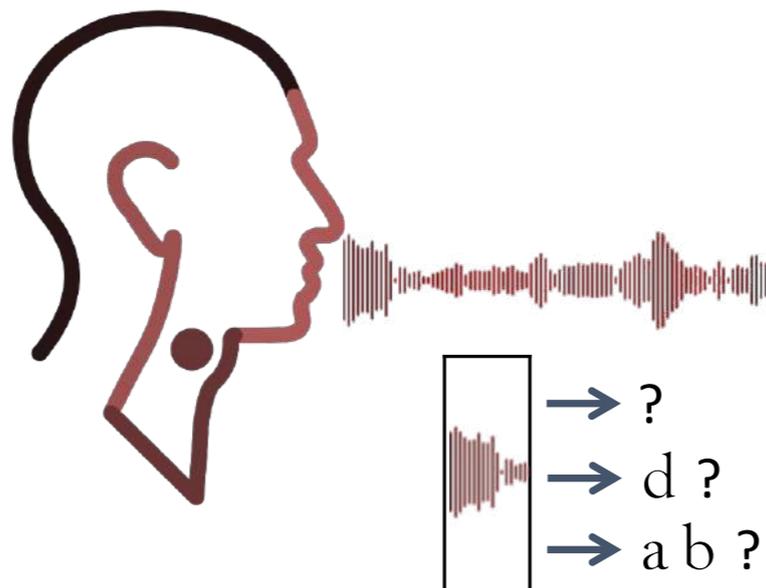
- **Taux d'erreur mots**
(Word Error Rate, *Errattahi et al., 2018*)

$$WER = \frac{\text{somme des erreurs (substitutions, insertions, délétions)}}{\text{nombre de mots cibles}}$$



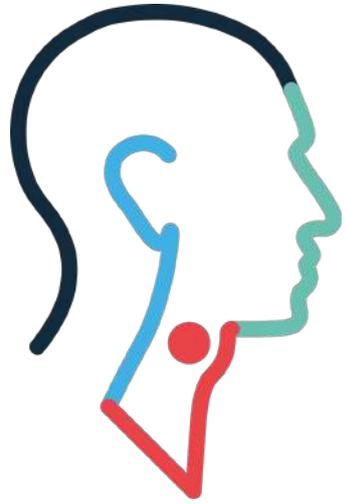
- **Parole typique**

- 8 % - 18 % en français
(*Radford et al., 2022*)



- **Parole cancérologique**

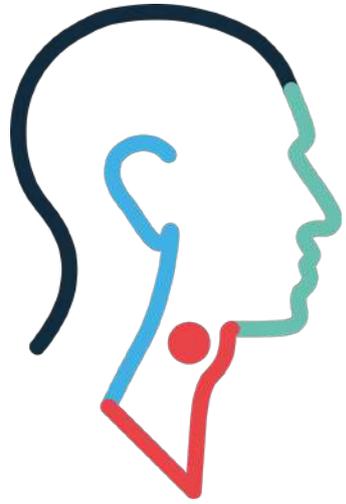
- > 84 % (*Balaguer, 2021*)
- Plus la forme acoustique sera éloignée de la forme typique, plus la parole sera dégradée et plus les erreurs de reconnaissance seront élevées



● OBJECTIF

55^e
congrès
SFCCF Société
Française de
Carcinologie
Cervico-Faciale
2023

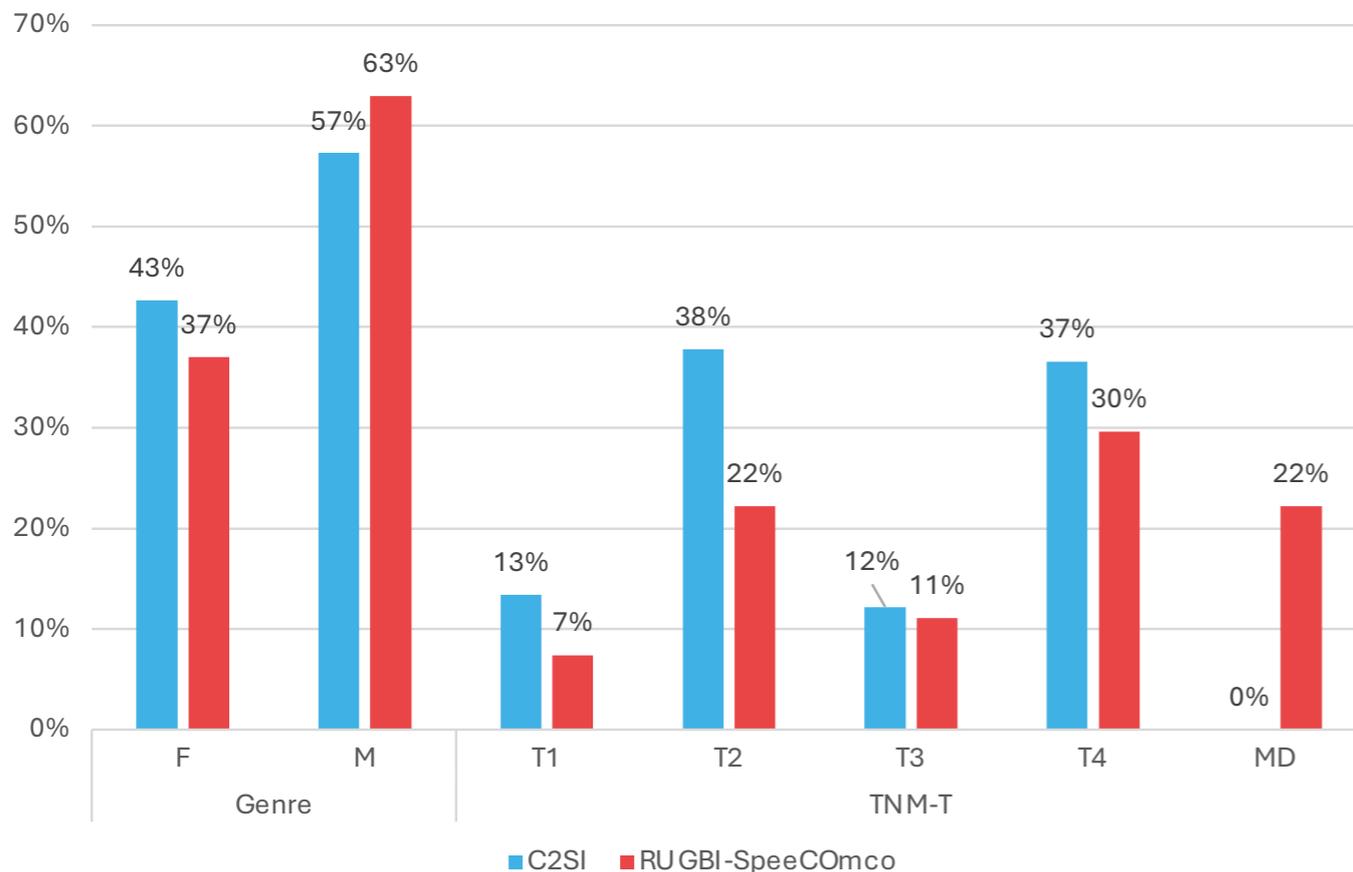
- Analyser les **performances** de systèmes ASR (taux d'erreurs mots), appliqués à la parole après cancers de la cavité buccale ou de l'oropharynx, dans le diagnostic
 - (1) de la **sévérité** du trouble de parole
 - (2) de **l'altération de la communication**



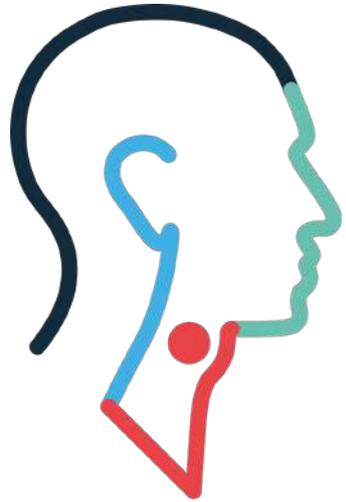
● MATÉRIEL ET MÉTHODES

- Données issues de deux projets
 - Projet **C2SI** (Woisard et al., 2021) : 82 sujets
 - Projet **RUGBI**, corpus **SpeeCOmco** (Balaguer et al., 2023) : 27 sujets

55^e
congrès
SFCCF Société Française de
Carcinologie
Cervico-Faciale
2023



		C2SI	RUGBI-SpeeCOmco
Âge	Moy (sd)	65,52 (9,58)	66,26 (9,63)
	Med (EIQ)	66 (13)	67 (13)
Délai depuis la fin du traitement	Moy (sd)	63,73 (57,15)	89,37 (119,79)
	Med (EIQ)	38,5 (66)	40 (142)
Sévérité du trouble de parole	Moy (sd)	6,18 (2,28)	5,90 (2,67)
	Med (EIQ)	6,33 (3,33)	6,33 (4,5)
Score holistique de communication (SHC)	Moy (sd)		100 (10)
	Med (EIQ)		99,26 (18,46)



55^e
congrès
SFCCF Société
Française de
Carcinologie
Cervico-Faciale
2023

● MATÉRIEL ET MÉTHODES

- Tâche de référence
 - **Lecture de texte** « La chèvre de Monsieur Seguin »

- **Scores de référence**

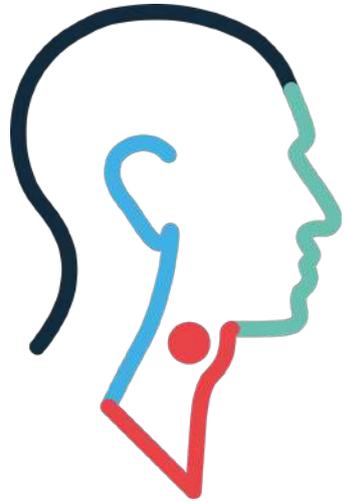
Sévérité du trouble

- 6 experts
- Score de **0** (trouble majeur) à **10** (pas de trouble)

Altération
de la communication

- Score SHC basé sur des autoquestionnaires
(Balaguer et al., 2023)
- Score centré sur **100**
 - Plus le score est élevé, meilleur est le ressenti
de communication

M. Seguin n'avait jamais eu de bonheur avec ses chèvres. Il les perdait toutes de la même façon : un beau matin, elles cassaient leur corde, s'en allaient dans la montagne, et là-haut le loup les mangeait. Ni les caresses de leur maître, ni la peur du loup, rien ne les retenait. C'était, paraît-il, des chèvres indépendantes, voulant à tout prix le grand air et la liberté.



● MATÉRIEL ET MÉTHODES

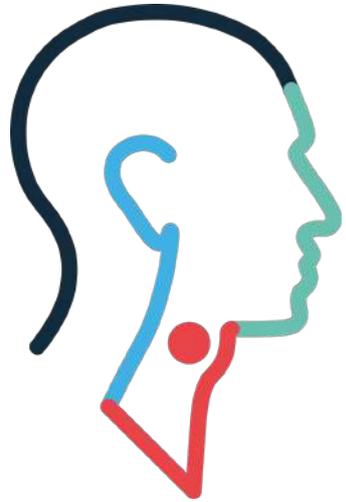
- Reconnaissance de la parole
 - Deux systèmes de RAP utilisés

Systeme	Architecture	Corpus d'entraînement (en heures)	WER sur parole typique (set de test)
OpenAI-Whisper (OAW)¹	Transformer Seq2seq	Multilingue (117 113 h)	8,3 %
Wav2vec2.0 avec SpeechBrain (SBW)²	Wav2vec avec CTC / attention	Multilingue + Français (7 600 h)	9,96 %

- Utilisation de la boîte à outils HTK (*Young et al., 2015*) pour le calcul du WER
 - Alignement entre texte de référence et texte reconnu par algorithme de Viterbi

¹ <https://github.com/openai/whisper>

² <https://huggingface.co/speechbrain/asr-wav2vec2-commonvoice-fr>



● RÉSULTATS

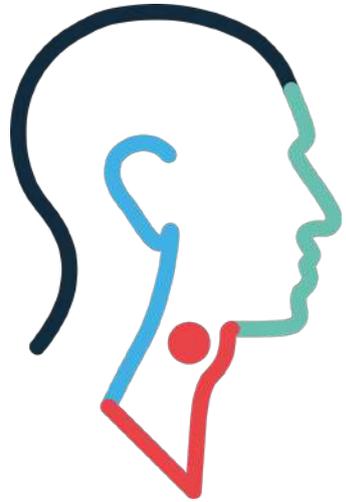
55^e
congrès
SFCCF Société
Française de
Carcinologie
Cervico-Faciale
2023

- Taux d'erreur mots (WER)

Corpus	n	OAW	SBW	p (Wilcoxon)
C2SI	82	45,75 %	47,31 %	0,09
RUGBI-SpeeCOmco	27	40,17 %	40,55 %	0,65
Global (C2SI + RUGBI-SpeeCOmco)	109	41,56 %	42,23 %	0,12

- Différences liées

- À la sévérité moyenne du trouble de parole ?
 - C2SI : 6,18/10 (sd = 2,28)
 - RUGBI-SpeeCOmco : 5,90/10 (sd = 2,67)
- Aux conditions d'enregistrement ?

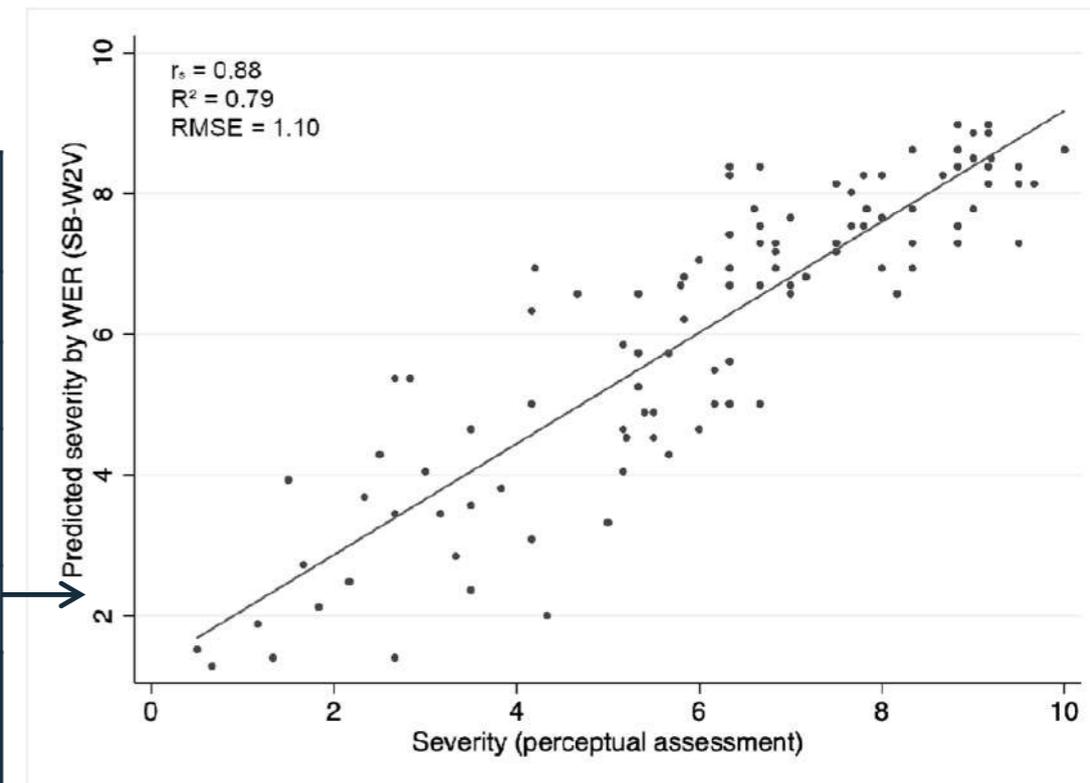


● RÉSULTATS

- Prédiction de la **sévérité** du trouble de parole et de l'**altération de la communication**

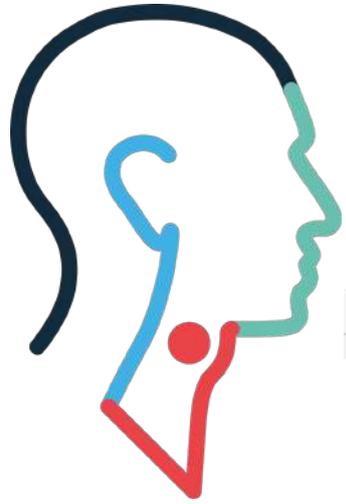
55^e
congrès
SFCCF Société Française de
Carcinologie Cervico-Faciale
2023

		WER	
		(r _s et * : p-value < 0,001)	
		OAW	SBW
Sévérité	C2SI	-0,76 *	-0,87 *
	RUGBI-SpeeCOmco	-0,87 *	-0,95 *
	Global	-0,79 *	-0,88 *
Altération de la communication	RUGBI-SpeeCOmco	-0,77 *	-0,84 *



$$\text{Sévérité} = 9,706 - (0,085 \times WER_{SBW})$$

$$\text{SHC (Altération de la communication)} = 113,847 - (0,286 \times WER_{SBW})$$



● RÉSULTATS

55^e
congrès
SFCCF Société Française de Carcinologie Cervico-Faciale
2023

```
SpeechBrain -- -bash -- 85x36
Last login: Thu Oct 26 08:54:12 on ttys000

The default interactive shell is now zsh.
To update your account to use zsh, please run `chsh -s /bin/zsh`.
For more details, please visit https://support.apple.com/kb/HT208050.
[lenovo-mathieu:~ mathieu$ source /Users/mathieu/venv-sb/bin/activate ]
[(venv-sb) lenovo-mathieu:~ mathieu$ cd /Users/mathieu/Documents/SpeechBrain/ ]
[(venv-sb) lenovo-mathieu:SpeechBrain mathieu$ bash wer-global_script.sh ]

Définir le contexte pathologique
 1 : Cancer de la cavité buccale ou de l'oropharynx
 2 : Maladie de Parkinson

1

torchvision is not available - cannot save figures
speechbrain.lobes.models.huggingface_wav2vec - wav2vec 2.0 is frozen.
CFS10-TXT MONSIEUR SEGA N AVAIT JAMAIS EU DE BONHEUR AVEC SES CHÈVRES IL LES PERDAINT
TOUTES DE LA MÊME FAÇON UN BEAU MATIN ELLE CASSAIENT LEUR CODE S EN ALLAIT DANS LA M
ONTAGNE ET LAEU LELUT LESMAINVUE NI LES CARESSES DE LEUR MAÎTRE NI LA PEUR DU LOUP RI
EN NE LES RETENAIT C ÉTAIT PARAÎT IL DES CHÈVRES INDÉPENDANTE VOULANT À TOUT PRIX LE
GRANTAIR ET LA LIBERTÉ

HTK Results Analysis at Thu Oct 26 08:55:29 2023
Ref: labels/
Rec: recFiles/CFS10-TXT.rec
=====
# Snt | Corr  Sub  Del  Ins  Err  S. Err
-----
Sum/Avg | 1 | 80.28 14.08 5.63 0.00 19.72 100.00

Indiquer le WER : 19.72
Score de sévérité : 8.02
Score d'altération de la communication : 108.2
(venv-sb) lenovo-mathieu:SpeechBrain mathieu$
```

```
SpeechBrain -- -bash -- 85x36
The default interactive shell is now zsh.
To update your account to use zsh, please run `chsh -s /bin/zsh`.
For more details, please visit https://support.apple.com/kb/HT208050.
[lenovo-mathieu:~ mathieu$ source /Users/mathieu/venv-sb/bin/activate ]
[(venv-sb) lenovo-mathieu:~ mathieu$ cd /Users/mathieu/Documents/SpeechBrain/ ]
[(venv-sb) lenovo-mathieu:SpeechBrain mathieu$ bash wer-global_script.sh ]

Définir le contexte pathologique
 1 : Cancer de la cavité buccale ou de l'oropharynx
 2 : Maladie de Parkinson

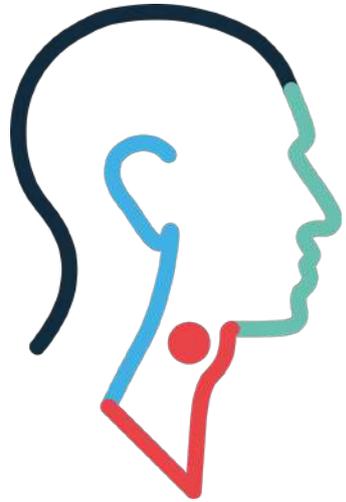
1

torchvision is not available - cannot save figures
speechbrain.lobes.models.huggingface_wav2vec - wav2vec 2.0 is frozen.
CMP15-TXT IL SE FOIN MADE VAMA EU DE BONHEUR ADOLPFAL ILE PERDE TOUT DE LA MÊME FAÇON
UN BEAU MAFIN ELE FAFEHAUD FANAE VAAMTAV E LAOLELOUEE MENVI MIRACA DEERMAT MEU A PEU
R DEOU MELEURDEMAN FAISAIT PARASBL T DE FRABRE INV BENDANC BOUA AGOUP U GRAND FRÈRE
À LA ILDAFEL

HTK Results Analysis at Thu Oct 26 09:07:18 2023
Ref: labels/
Rec: recFiles/CMP15-TXT.rec
=====
# Snt | Corr  Sub  Del  Ins  Err  S. Err
-----
Sum/Avg | 1 | 15.49 50.70 33.80 0.00 84.51 100.00

Indiquer le WER : 84.51
Score de sévérité : 2.48
Score d'altération de la communication : 89.67
(venv-sb) lenovo-mathieu:SpeechBrain mathieu$
```

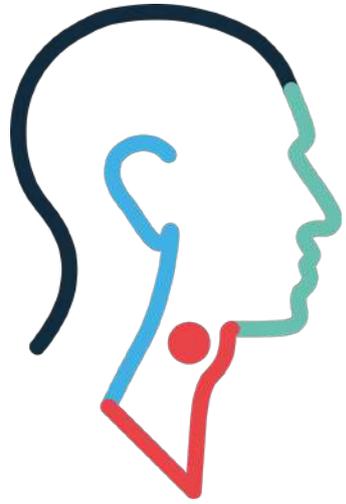
Clinique	F, 70 ans Base de langue G et sillon amygdalogue - T2N0M0 Oropharyngectomie avec lambeau libre G + RTE - 63 mois	H, 62 ans Plancher buccal - T2N0M1 Pelviglossectomie D avec lambeau antébrachial + RTE + Chimio - 6 mois
WER	19,72	84,51
Sévérité	Référence : 7,67 Prédite : 8,02	Référence : 2,17 Prédite : 2,48
Altération de la communication	Référence : 104,97 Prédite : 108,2	Référence : 89,32 Prédite : 89,67



55^e
congrès
SFCCF Société
Française de
Carcinologie
Cervico-Faciale
2023

● CONCLUSION

- **Prédiction du trouble de parole en contexte cancérologique possible** grâce à la métrique du taux d'erreurs mots issue de systèmes de reconnaissance automatique de la parole
- **Résultats comparables** à des études antérieures (*Quintas et al., 2023 ; Woisard et al., 2022 ; Windrich et al., 2008*)
 - Mais **moins de ressources** utilisées
- Outil **prometteur en clinique courante** : automatique, fiable, simple d'utilisation et open source
- **Perspectives**
 - Nouveaux systèmes de reconnaissance automatique de la parole : quelles performances en parole pathologique ? (*Open AI Whisper X, 2023*)
 - Si amélioration de la reconnaissance de la parole pathologique : analyses de contenu, outil de compensation...
 - Implémentation clinique : applications, outils en ligne



55^e
congrès
SFCCF Société Française de
Carcinologie Cervico-Faciale
2023

● RÉFÉRENCES

- Balaguer, M., Pinquier, J., Farinas, J., & Woisard, V. (2023). Development of a holistic communication score (HoCoS) in patients treated for oral or oropharyngeal cancer: Preliminary validation. *International Journal of Language & Communication Disorders*, 58(1), 39–51. <https://doi.org/10.1111/1460-6984.12766>
- 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* [Université Toulouse III Paul Sabatier]. <http://www.theses.fr/2021TOU30109>
- Borggreven, P. A., Verdonck-de Leeuw, I. M., Muller, M. J., Heiligers, 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>
- Errattahi, R., El Hannani, A., & Ouahmane, H. (2018). Automatic Speech Recognition Errors Detection and Correction: A Review. *Procedia Computer Science*, 128, 32–37. <https://doi.org/10.1016/j.procs.2018.03.005>
- Kim, Y., Kent, R. D., & Weismer, G. (2011). An Acoustic Study of the Relationships Among Neurologic Disease, Dysarthria Type, and Severity of Dysarthria. *Journal of Speech, Language, and Hearing Research*, 54(2), 417–429. [https://doi.org/10.1044/1092-4388\(2010/10-0020\)](https://doi.org/10.1044/1092-4388(2010/10-0020))
- Kuo, C., & Tjaden, K. (2016). Acoustic variation during passage reading for speakers with dysarthria and healthy controls. *Journal of Communication Disorders*, 62, 30–44. <https://doi.org/10.1016/j.jcomdis.2016.05.003>
- Li, J. (2021). Recent Advances in End-to-End Automatic Speech Recognition. *Signal Processing*, 1(2), 95–123. <https://doi.org/https://doi.org/10.48550/arXiv.2111.01690>
- Middag, C., Clapham, R., van Son, R., & Martens, J.-P. (2014). Robust automatic intelligibility assessment techniques evaluated on speakers treated for head and neck cancer. *Computer Speech & Language*, 28(2), 467–482. <https://doi.org/10.1016/j.csl.2012.10.007>
- Pommée, T., Balaguer, M., Mauclair, J., Pinquier, J., & Woisard, V. (2022). Assessment of adult speech disorders: current situation and needs in French-speaking clinical practice. *Logopedics Phoniatrics Vocology*, 47(2), 92–108. <https://doi.org/10.1080/14015439.2020.1870245>
- Quintas, S., Mauclair, J., Woisard, V., & Pinquier, J. (2020). Automatic Prediction of Speech Intelligibility Based on X-Vectors in the Context of Head and Neck Cancer. *Interspeech 2020, 2020-October*, 4976–4980. <https://doi.org/10.21437/Interspeech.2020-1431>
- Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2022). Robust Speech Recognition via Large-Scale Weak Supervision. *Openai.Com*. <https://doi.org/https://doi.org/10.48550/arXiv.2212.04356>
- Windrich, M., Maier, A., Kohler, R., Nöth, E., Nkenke, E., Eysholdt, U., & Schuster, M. (2008). Automatic Quantification of Speech Intelligibility of Adults with Oral Squamous Cell Carcinoma. *Folia Phoniatrica et Logopaedica*, 60(3), 151–156. <https://doi.org/10.1159/000121004>
- Woisard, V., Astésano, C., Balaguer, M., Farinas, J., Fredouille, C., Gaillard, P., Ghio, A., Giusti, L., Laaridh, I., Lalain, M., Lepage, B., Mauclair, J., Nocaudie, O., Pinquier, J., Pouchoulin, G., Puech, M., Robert, D., & Roger, V. (2021). C2SI corpus: a database of speech disorder productions to assess intelligibility and quality of life in head and neck cancers. *Language Resources and Evaluation*, 55(1), 173–190. <https://doi.org/10.1007/s10579-020-09496-3>
- Woisard, V., Balaguer, M., Fredouille, C., Farinas, J., Ghio, A., Lalain, M., Puech, M., Astesano, C., Pinquier, J., & Lepage, B. (2022). Construction of an automatic score for the evaluation of speech disorders among patients treated for a cancer of the oral cavity or the oropharynx: The Carcinologic Speech Severity Index. *Head & Neck*, 44(1), 71–88. <https://doi.org/10.1002/hed.26903>
- Young, S., Evermann, G., Gales, M., Hain, T., Kershaw, D., Liu, X. (Andrew), Moore, G., Odell, J., Ollason, D., Pover, D., Ragni, A., Valtchev, V., Woodland, P., & Zhang, C. (2015). *The HTK Book (for HTK Version 3.5, documentation alpha version)* (Issue December). Cambridge University Engineering Department.