

Interprétation d'un score d'intelligibilité dans le cadre de l'évaluation de troubles de la parole au travers d'une représentation profonde

Sondes Abderrazek, Corinne Fredouille, Alain Ghio, Muriel Lalain, Christine Meunier, Virginie Woisard



Contexte

Les troubles de la parole affectent la capacité d'une personne à former des sons corrects.



Cancers de la tête et du cou



Maladies neurologiques



Anomalies physiques



- Isolement social
- Dépression
- Anxiété
- ...

Affecter la qualité de vie de la personne

Contexte



L'évaluation des troubles de la parole est une mesure cruciale pour :

- Suivre l'évolution d'une pathologie
- Évaluer l'efficacité d'une rééducation orthophonique
- Comprendre les effets d'un traitement particulier

Contexte

Évaluation perceptive : un expert clinicien attribue différents scores pour évaluer la qualité de la parole d'un patient.

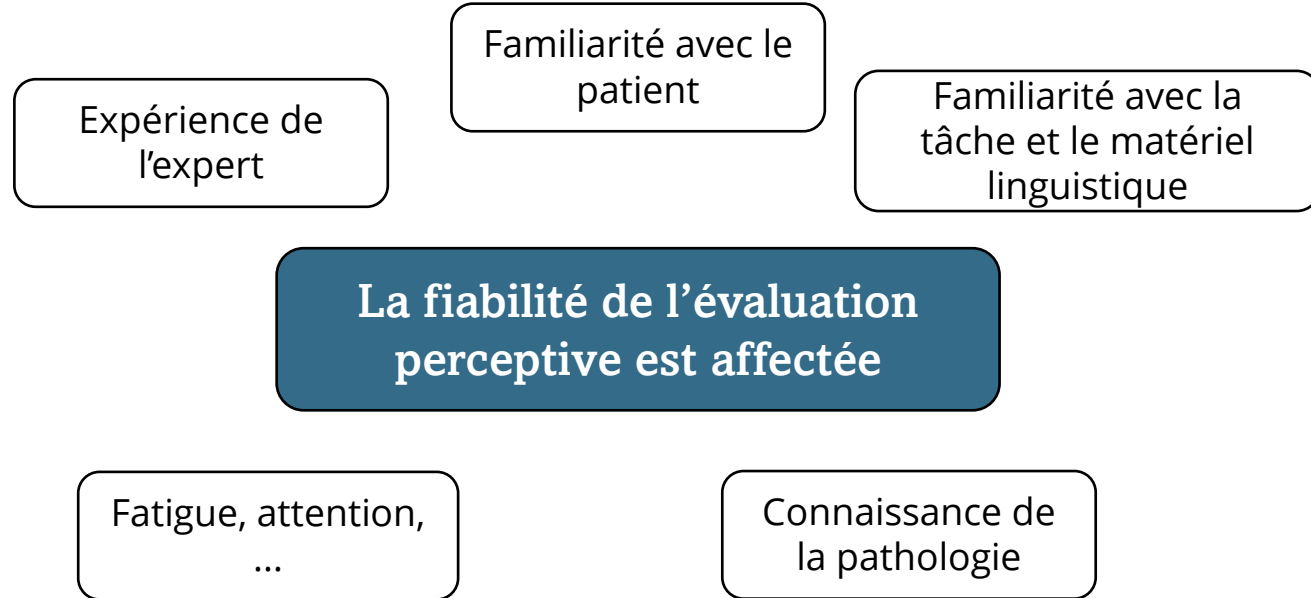


L'intelligibilité est la mesure du degré de précision avec lequel un auditeur décode le message d'un locuteur.

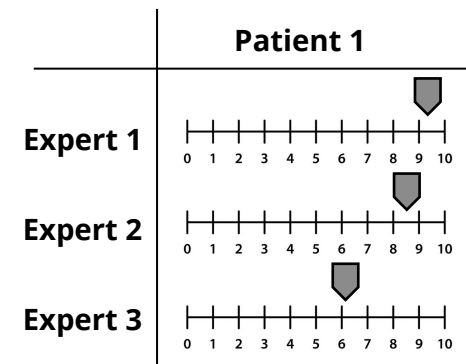
La sévérité est le degré d'altération globale du signal de la parole.

Contexte

Évaluation perceptive : un expert clinicien attribue différents scores pour évaluer la qualité de la parole d'un patient.



➤ Subjectivité



RUGBI

RUGBI: “Looking for Relevant linguistic Units to improve the intelligibility measurement of speech production disorder”

➤ **Projet ANR : 2018 — 2023**



Projet multidisciplinaire impliquant des experts des domaines de la santé, de l’informatique, de la linguistique et de la phonétique clinique.

➤ **Buts:**

- ❖ Développer un outil d’évaluation objective de l’intelligibilité de la parole dans le contexte des troubles de la parole dus principalement au cancer de la tête et du cou.
- ❖ Analyser la contribution des outils d’apprentissage profond.



Travaux en relation



Méfiance des experts (Boîte noire)



Absence d'interprétation par rapport à des altérations locales

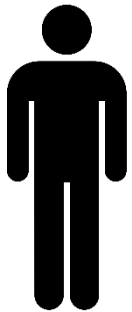
[Bin et al., 2019] Bin, L., Kelley, M. C., Aalto, D., & Tucker, B. V. (2019). Automatic speech intelligibility scoring of head and neck cancer patients with deep neural networks. In International Congress of Phonetic Sciences (ICPHs' 19), Melbourne, Australia.

[Fernández Díaz and Gallardo-Antolín, 2020] Fernández Díaz, M. and Gallardo-Antolín, A. (2020). An attention long short-term memory based system for automatic classification of speech intelligibility. Engineering Applications of Artificial Intelligence.

Méthodologie proposée

STEP 1: Représentation en Phonèmes

Garantir un premier niveau d'interprétation par rapport à des altérations locales

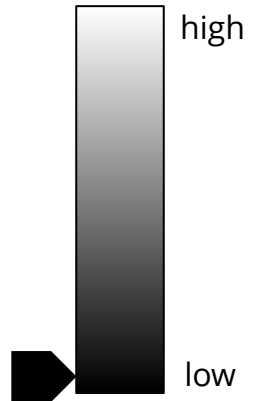


Patient



Score d'intelligibilité

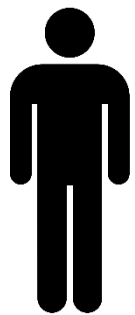
Interprétabilité



Phonème: premier niveau d'interprétabilité

Méthodologie proposée

STEP 2: Exploration des représentations cachées du CNN

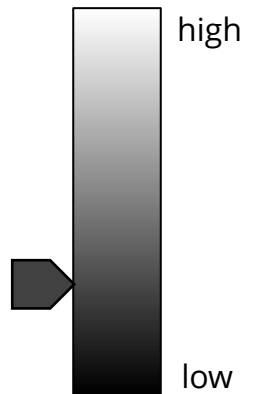


Patient



Score d'intelligibilité

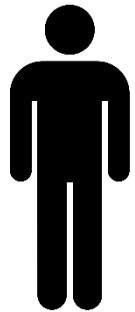
Interprétabilité



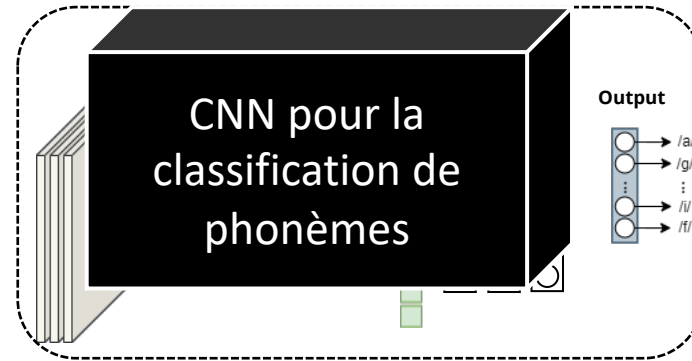
Phonème: premier niveau d'interprétabilité

Méthodologie proposée

STEP 2: Exploration des représentations cachées du CNN



Patient

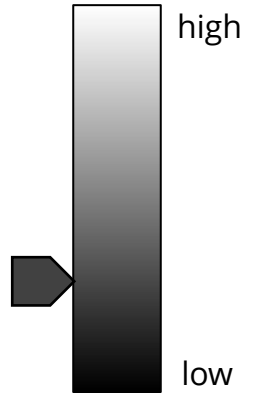


Neuro-based Concept Detector (NCD):
Framework pour l'explicabilité



Score d'intelligibilité

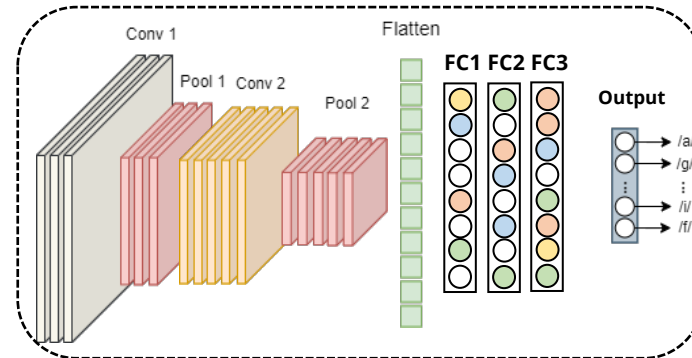
Interprétabilité



Phonème: premier niveau d'interprétabilité

Méthodologie proposée

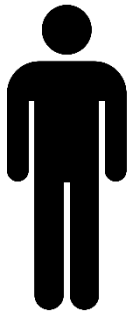
STEP 2: Exploration des représentations cachées du CNN



Neuro-based Concept Detector (NCD):

Framework pour l'explicabilité

Emergence automatique du concept de traits phonétiques

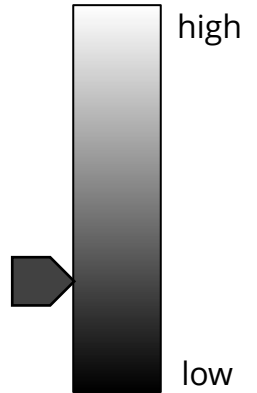


Patient



Score d'intelligibilité

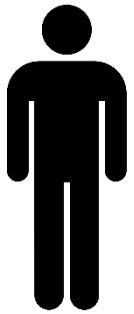
Interprétabilité



Phonème: premier niveau d'interprétabilité

Méthodologie proposée

STEP 2: Exploration des représentations cachées du CNN



Patient

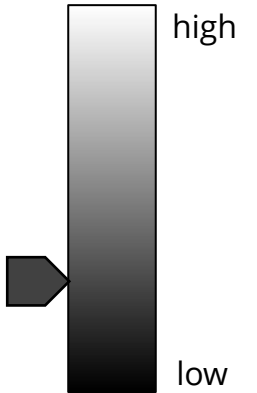


Emergence automatique du concept de traits phonétiques



Score d'intelligibilité

Interprétabilité

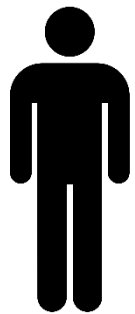


Phonème: premier niveau d'interprétabilité

Trait phonétique: deuxième niveau d'interprétabilité

Méthodologie proposée

STEP 3: Prédiction de l'intelligibilité

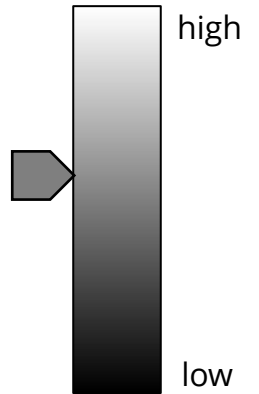


Patient



Score d'intelligibilité

Interprétabilité



Phonème: premier niveau d'interprétabilité

Trait phonétique: deuxième niveau d'interprétabilité

Données

Les corpus de troubles de la parole dus aux cancers de la tête et du cou



Corpus C2SI : 40 sujets contrôles et 87 patients [\[Woisard et al., 2021\]](#)

Recorded Tasks

- Lecture (**LEC**)
- Description d'image (DES)
- Tâches prosodiques
- ...

Perceptual Measures

- Intelligibilité
- Sévérité
- Altération phonémique
- ...

Données

Les corpus de troubles de la parole dus aux cancers de la tête et du cou



Corpus C2SI : 40 sujets contrôles et 87 patients [Woisard et al., 2021]

Recorded Tasks

- Lecture (C2SI-LEC)
- Description d'image (DES)
- Tâches prosodiques
- ...

≈ 1 hour



Perceptual Measures

- Intelligibilité
- Sévérité
- Altération phonémique
- ...

x 6 experts



Expert

Lecture du passage
"Monsieur Seugin
n'avait jamais eu de ..."



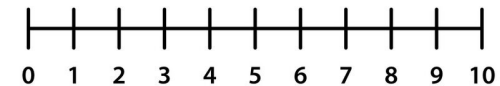
Patient

≈ 15 minutes



Corpus SpeeCOMco : 25 patients [Balaguer, 2021]

Échelle C2SI/ SpeeCOMco pour
l'intelligibilité/sévérité

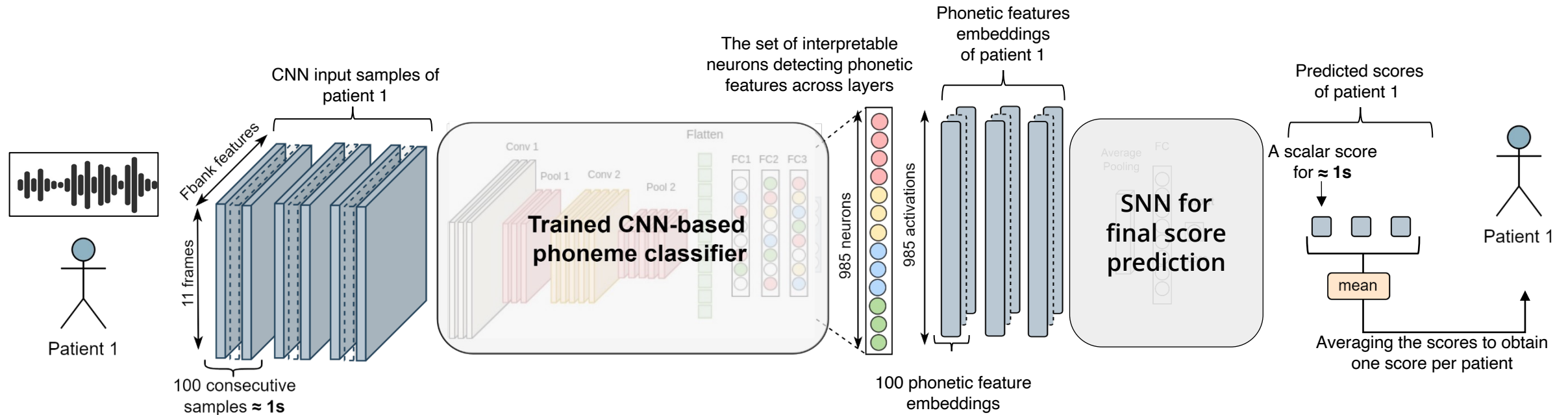


Parole inintelligible /
Altération importante

Discours parfaitement
intelligible /
Absence d'altération

STEP 3: Évaluation

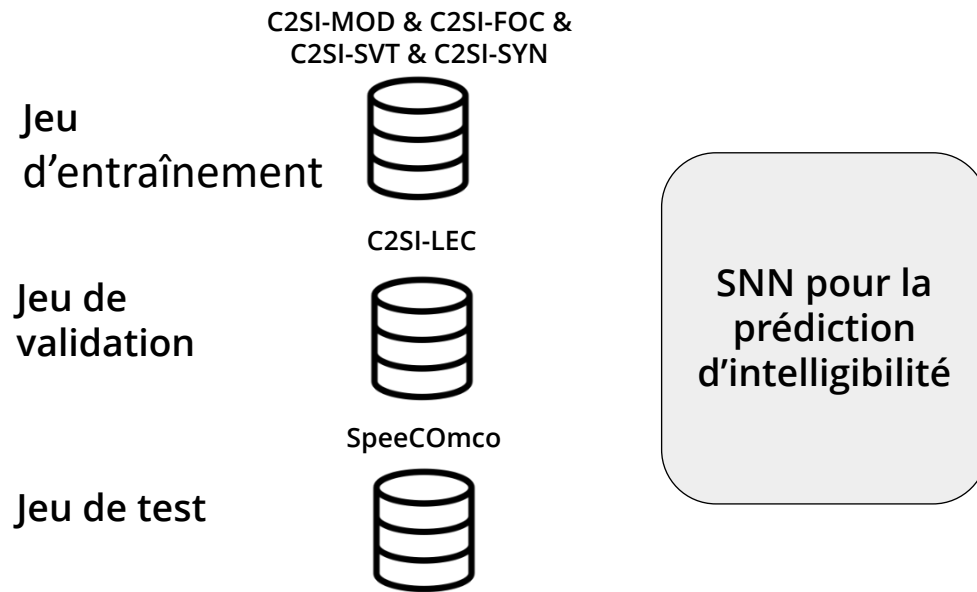
Tâche de regression pour un score final



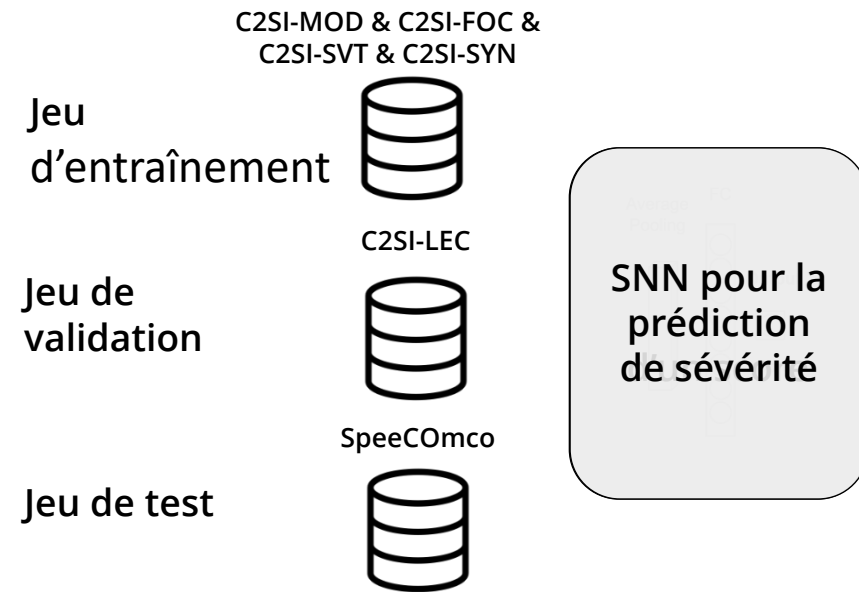
STEP 3: Évaluation

Tâche de regression pour un score final

Tâche 1: Prédiction d'intelligibilité



Tâche 2: Prédiction de sévérité



Résultats

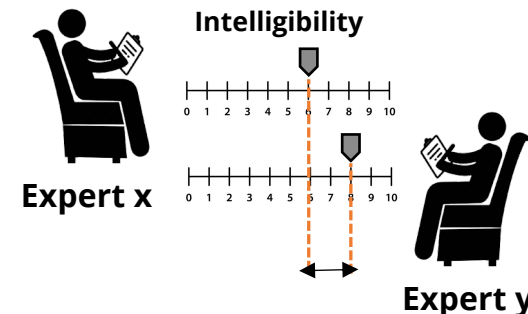
Résultats de la régression

- Métriques d'évaluation :
- $$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$
- $$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

		Prédiction de l'intelligibilité	Prédiction de la sévérité
C2SI-LEC	MAE	1.21	1.25
	MSE	2.97	2.55
SpeeCOmco	MAE	1.32	1.4
	MSE	2.97	2.97

Régression basée sur les embeddings des traits phonétiques

- Erreur inférieure à la différence que l'on peut observer entre l'évaluation perceptive des experts.



Résultats

Analyse de prédiction du modèle sur le corpus C2SI-LEC (jeu de validation)

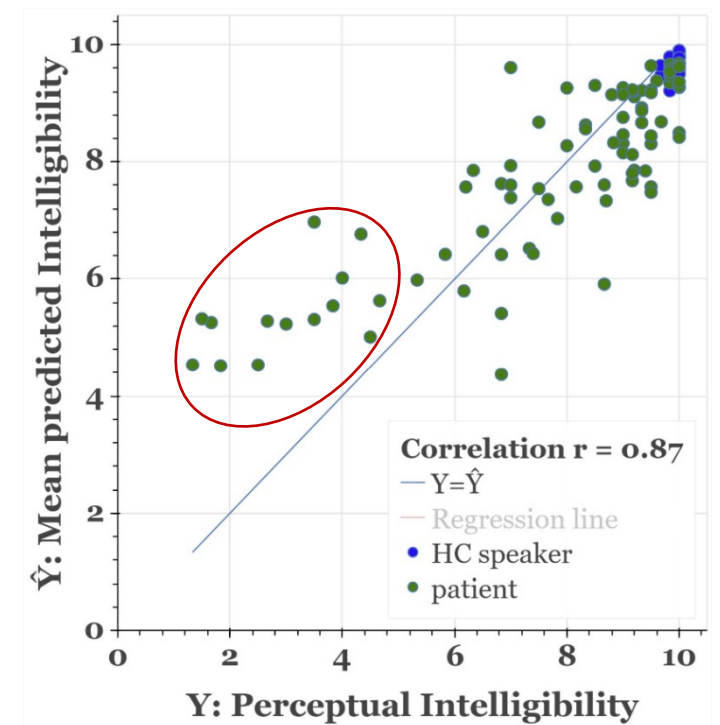
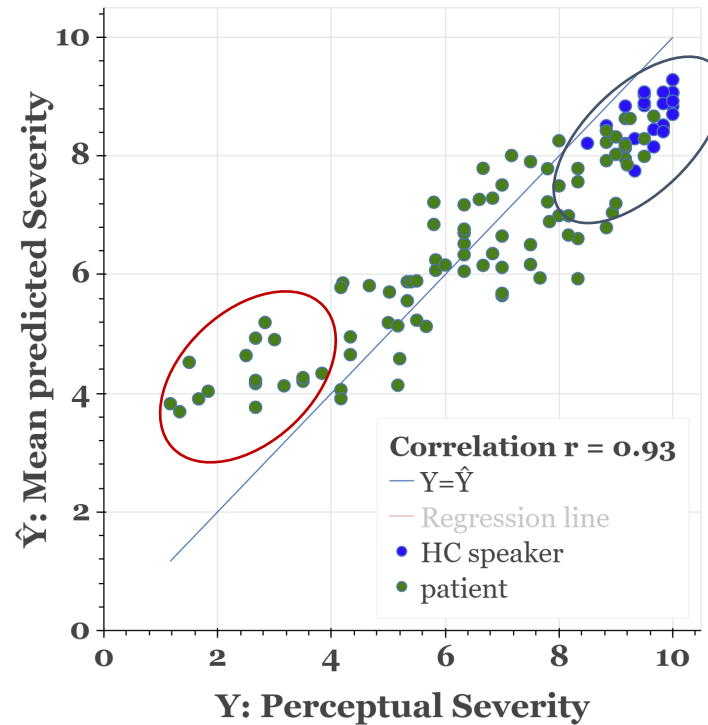
Modèles capables de capturer certaines caractéristiques sous-jacentes liées à l'intelligibilité/ sévérité

Tâche de prédiction de la sévérité :

- Sous-estime les scores de sévérité élevés
- Surestime les scores de sévérité faibles

Tâche de prédiction d'intelligibilité :

- Surestime les faibles scores d'intelligibilité



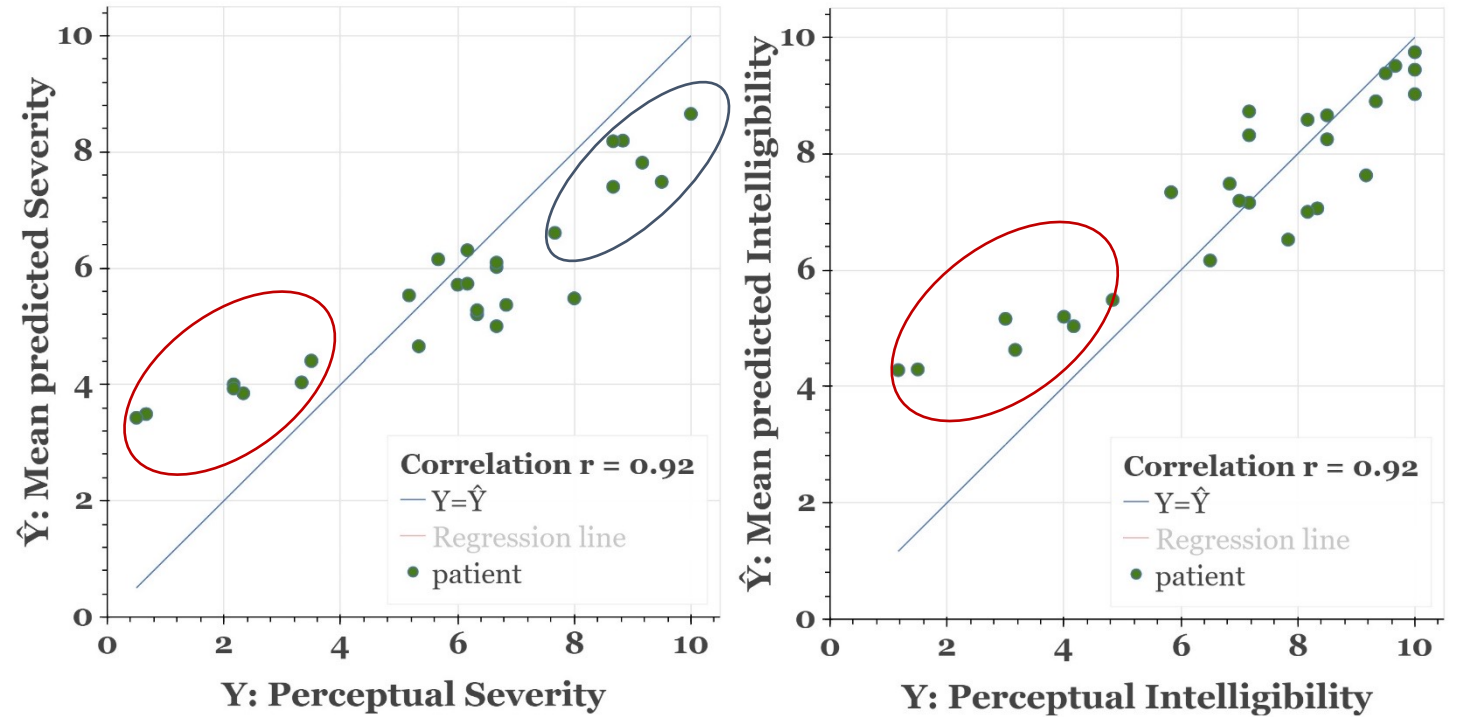
Nuage de points du score moyen prédit par rapport au score perceptif réel des locuteurs C2SI-LEC

Résultats

Analyse de prédiction de modèle sur le corpus SpeeCOMco (jeu de test)

Écarter toute possibilité de conclusions trompeuses de l'analyse précédente.

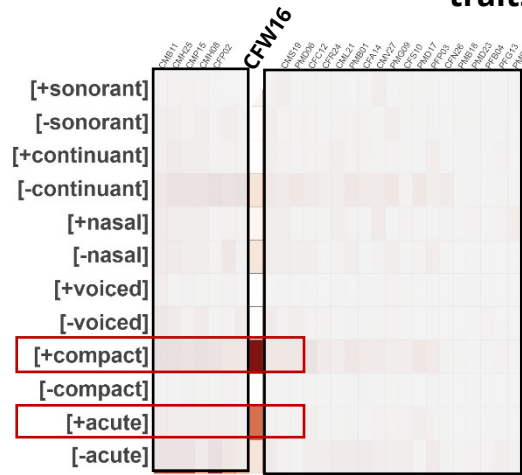
- Exactement les mêmes tendances que celles sur C2SI-LEC.



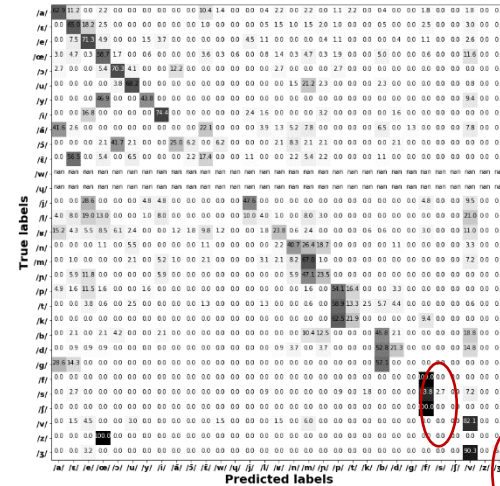
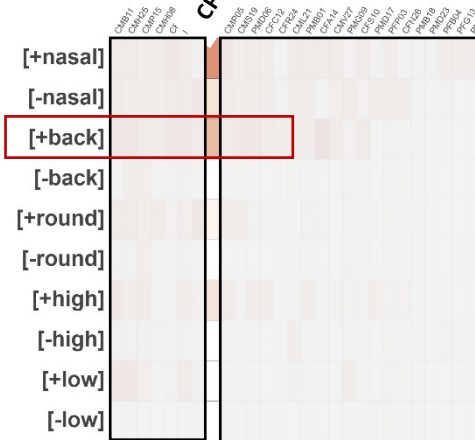
Nuage de points du score moyen prédit par rapport au score perceptif réel des locuteurs SpeeCOMco

Application de bout en bout de la méthodologie proposée

Interprétations au niveau des traits phonétique



Interprétations au niveau des phonèmes



CFW16

Détails de la tumeur :

- ❖ Localisation : Cavité buccale
- ❖ Région : Langue
- ❖ Taille : T4

Traitement: pelvi-glossectomie

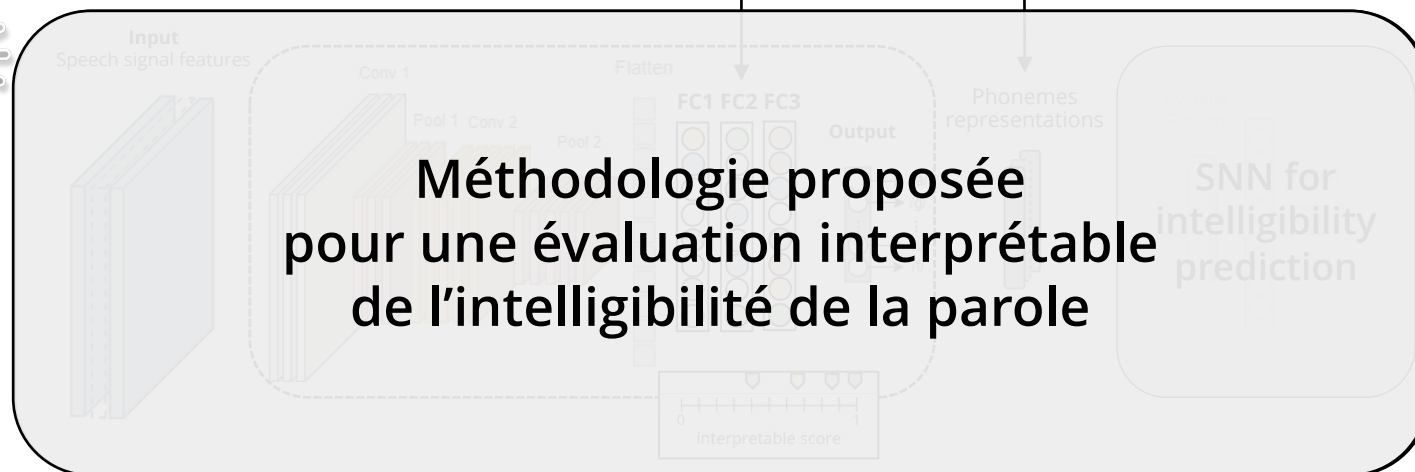
Scores perceptives:

- ❖ Intelligibilité: 5,83
- ❖ Sévérité: 5,16

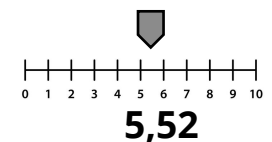
/b/ absorbant les /d/ et /g/
/f/ absorbant les /s/ et /ʃ/



CFW16



Intelligibilité prédite de la parole



Conclusion

Méthodologie en trois étapes basée sur l'apprentissage profond et dédiée à une évaluation interprétable de l'intelligibilité de la parole dans le contexte des troubles de la parole.



Absence d'interprétation de l'évaluation finale en ce qui concerne les altérations locales

- Interprétation du score final au niveau des phonèmes et traits phonétiques.

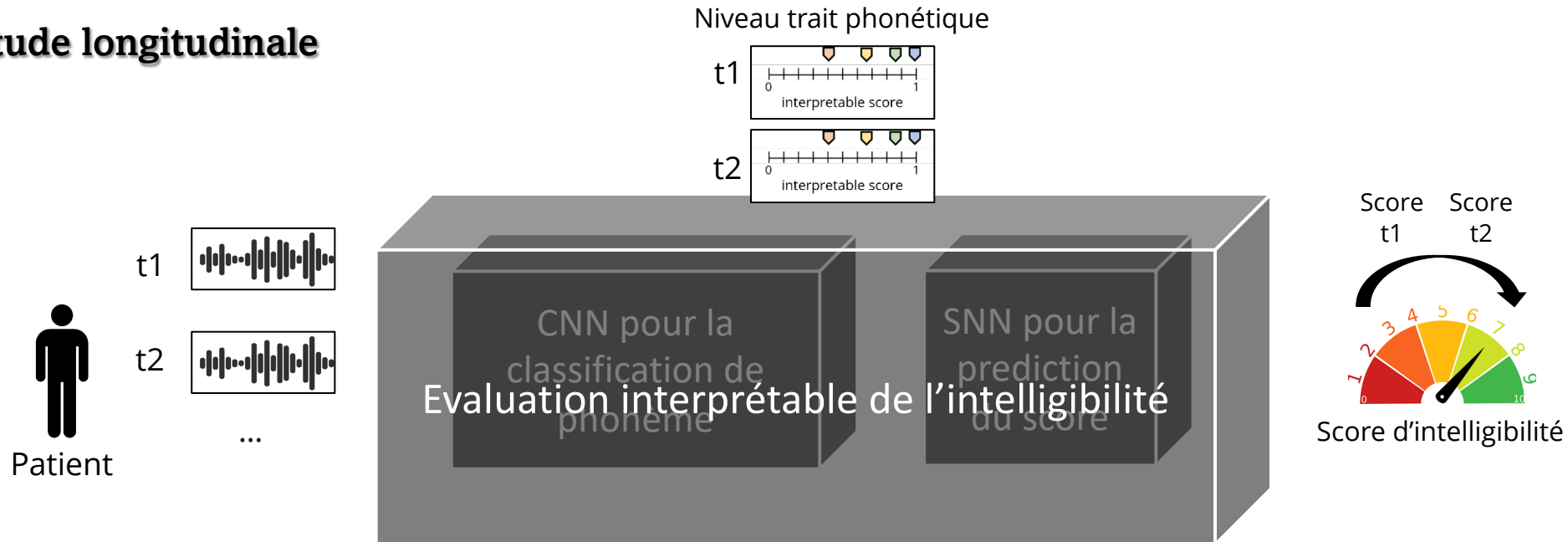


Une solution interprétable basée sur le Deep Learning

- Atténuation de l'effet boîte noire des modèles Deep Learning et donc la méfiance des experts dans un contexte clinique.

Perspectives

Etude longitudinale



Deuxième langue (L2)

Dans les plateformes e-learning dédiées à l'amélioration de la réalisation de locuteurs L2.

Application du framework NCD pour l'explicabilité dans d'autres contextes

Différentes architectures/ tâches/ concepts

Merci 😊