

Stage Master Recherche – 6 mois

Contact : Corinne Fredouille (corinne.fredouille@univ-avignon.fr)

Sujet : Approches à base de Deep Learning appliquées à l'évaluation de l'intelligibilité chez des patients atteints de troubles de la parole.

Le terme « troubles de la parole » fait référence à l'ensemble des déficiences affectant la production de la parole chez un être humain. Le bégaiement est un exemple de troubles de la parole.

Le LIA travaille depuis une dizaine d'années sur les troubles de la parole, et plus particulièrement sur la manière dont les outils de traitements automatiques peuvent aider les cliniciens et les phonéticiens dans leur analyse acoustico-phonétique du signal de parole et/ou perceptive des productions de parole en vue de mieux comprendre les dégradations inhérentes aux troubles de la parole. Les travaux les plus récents portent notamment sur l'étude d'un système de détection automatique de déviations dans des productions de parole dégradée [Laaridh et al., 2015] ou d'un système automatique de prédiction du degré d'intelligibilité basé sur des i-vecteurs [Laaridh et al., 2017 ; Laaridh et al., 2018]. Dans un premier temps, ces approches ont été appliquées sur des productions de parole dégradée produites par des patients atteints de lésions neurologiques localisées dans le système nerveux central ou périphérique. On parle alors de trouble moteur de la parole d'origine neurologique, désigné sous le terme de dysarthrie. La dysarthrie peut être l'un des symptômes de différentes maladies telles que la maladie de Parkinson, la Sclérose Latérale Amyotrophique (SLA), les Accidents Vasculaires Cérébraux, etc. Ces différentes maladies se distinguent notamment par la localisation des lésions neurologiques et, par conséquent, par le type de troubles moteur (faiblesse musculaire, mouvements involontaires, imprécision des mouvements ...) et le type de dégradations de la parole qu'elles peuvent engendrer (distorsion des voyelles, imprécision des consonnes, altération du débit, hypernasalité, ...). Ces approches ont été, dans un deuxième temps, évaluées sur des productions de parole dégradées issues de patients atteints de cancers des voies aérodigestives supérieures (présence de tumeurs) et/ou suite à des traitements thérapeutiques inhérents (exérèse, radiothérapie, etc).

En fonction de la maladie et de son évolution, les troubles moteur du patient pourront être évalués de manière perceptive (« à l'oreille ») par le clinicien sur une échelle de sévérité allant d'une dysarthrie légère à sévère. Sur une échelle similaire, le clinicien pourra également juger du degré d'intelligibilité de la parole d'un patient et de sa capacité à transmettre un message oral à un auditeur.

Malgré son caractère très subjectif, l'évaluation perceptive des troubles de la parole chez les patients reste la seule approche à l'heure d'aujourd'hui utilisée en pratique clinique. Au vu des progrès observés dans le traitement automatique de la parole ces dix dernières années, des solutions technologiques sont âprement attendues dans ce domaine pour aider les cliniciens dans leur bilan clinique. Néanmoins, même si de nombreux travaux scientifiques portent sur l'utilisation d'approches automatiques pour une évaluation objective des troubles de la parole, un besoin de mieux comprendre les dégradations dans le signal de parole est nécessaire.

Le LIA est engagé dans un projet financé par l'Agence Nationale de la Recherche (ANR) avec 3 autres partenaires sur la période 2019-2022 sur ce thème. L'un des objectifs de ce projet est de mieux comprendre quelles unités linguistiques sont majoritairement impliquées dans les processus d'intelligibilité du locuteur. Ainsi, une altération observée ou attendue de ces unités, conséquence d'une pathologie particulière, pourrait permettre de quantifier de manière objective la perte d'intelligibilité chez le patient. D'un point de vue pratique, il s'agira au travers des approches de Deep Learning et d'une comparaison parole normale/parole dégradée associée à des évaluations perceptives de l'intelligibilité d'aborder cette question.

Le sujet du stage proposé ici s'inscrit dans ce cadre. Il aura pour objectif la mise en place du cadre expérimental nécessaire à la recherche des unités linguistiques impliquées dans ces processus d'intelligibilité. Il s'appuiera sur un état de l'art sur les approches de Deep Learning, qui devront être

vues non pas comme des boites noires mais comme un moyen d'extraire de l'information utile et de comprendre les processus étudiés : ici l'intelligibilité du locuteur.

A l'issue de ce stage, un financement de thèse (projet ANR) pourra être proposé au candidat.

Références bibliographiques

[Laaridh et al., 2015] I. Laaridh, C. Fredouille, C. Meunier, « Automatic Detection of Phone-Based Anomalies in Dysarthric Speech », ACM Transactions on Accessible Computing (TACCESS), (Volume : 6 Issue 3), June 2015.

[Laaridh et al., 2017] I. Laaridh, W. Ben-Kheder, C. Fredouille, C. Meunier, « Automatic Prediction of Speech Evaluation Metrics for Dysarthric Speech », Interspeech'2017, Stockholm, Sweden. August 2017.

[Laaridh et al., 2018] I. Laaridh, C. Fredouille, A. Ghio, M. Lalain, V. Woisard, « Automatic evaluation of speech intelligibility based on i-vectors in the context of Head and Neck Cancers », Interspeech'2018, Hyderabad, India. September 2018.