

Sujet de thèse

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

Sujet : Prise en compte de la variabilité inter-juges de l'évaluation perceptive des troubles de la parole et de la voix et son intégration dans un système automatique d'aide à la décision.

Bien que les modes de communication aient largement évolué depuis une vingtaine d'années avec l'ère du numérique, pouvant se substituer dans de nombreuses situations à la parole - messagerie électronique, réseaux sociaux, blogs, messageries instantanées... la parole reste indispensable pour une bonne intégration dans notre société. La communication orale permet de partager avec les autres, que ce soit dans le monde professionnel ou la vie de tous les jours : partager pour exprimer une demande, pour informer, pour conseiller, pour débattre, pour convaincre, pour enseigner, pour divertir, pour exprimer un sentiment, ... pour vivre en société.

La démocratisation de la téléphonie mobile, des réseaux de communication, des logiciels de visioconférences et de l'accès à Internet fait que nous pouvons communiquer oralement partout avec tous nos contacts professionnels ou personnels, quelle que soit leur localisation géographique. Aussi, au vu de l'importance de la communication orale, perdre la parole ou le langage peut être ressenti comme une perte d'humanité.

Les troubles de la communication sont définis par *The American Speech and Hearing Association* de la manière suivante : *"An impairment in the ability to receive, send, process, and comprehend concepts or verbal, nonverbal and graphic symbol systems. A communication disorder may be evident in the processes of hearing, language, and/or speech. A communication disorder may range in severity from mild to profound. It may be developmental or acquired. Individuals may demonstrate one or any combination of the three aspects of communication disorders. A communication disorder may result in a primary disability or it may be secondary to other disabilities"* (ASHA, 1993). Sur la base de cette définition, les troubles de la communication englobent toute altération au niveau de la voix, de la parole, du langage et de l'audition nuisant à la communication. Une enquête réalisée aux Etats-Unis en 2012 estime à 10% le nombre de personnes disant souffrir d'un trouble de la communication (seulement 2% se déclarant sur la base d'un diagnostic clinique) (Morris et al., 2016).

Nous nous intéresserons ici aux troubles de la voix et de la parole. Ces troubles peuvent être liés soit à des dommages de l'appareil phonatoire – *troubles de la voix appelés dysphonie* -, soit à des atteintes neurologiques induisant des troubles moteurs – *appelés dysarthrie* -, ou encore à des dommages liés à une malformation ou à un cancer localisé au niveau de l'appareil de production de la parole - *appelés cancers des Voies Aéro-Digestives Supérieures (VADS), cancers du cou et de la tête ou cancers ORL*.

Pour exemple, les cancers de la tête et du cou sont considérés comme le quatrième type de cancer le plus fréquent chez l'homme. D'après le bulletin épidémiologique hebdomadaire de la santé publique d'avril 2018¹, « *les maladies neurologiques représentent aujourd'hui la principale cause d'invalidité et la deuxième cause de décès dans le monde . Parmi elle, la maladie de Parkinson est celle dont le nombre de cas a le plus augmenté entre 1990 et 2015 : il a plus que doublé* ». La prévalence des Accidents Vasculaires Cérébraux est quant à elle estimée autour de 130000 cas par an en France². L'une des conséquences de ces trois exemples de maladies peut être une dégradation

1 http://beh.santepubliquefrance.fr/beh/2018/8-9/pdf/2018_8-9.pdf

2 https://has-sante.fr/upload/docs/application/pdf/2019-06/app_260_note_de_problematique_ssr_avc_cd_2019_05_22_vfinale.pdf

du fonctionnement de l'appareil de production de la parole, impactant de manière très variable la communication chez les patients atteints. Ces troubles de la communication peuvent avoir des conséquences importantes sur la qualité de vie des patients, pouvant entraver leur vie quotidienne, leurs activités professionnelles, leurs relations sociales et familiales et engendrer d'autres maladies comme la dépression. Améliorer la qualité de vie des patients est un objectif central du parcours de soin des personnes dans lequel le maintien de la communication orale est l'un des points à considérer par les cliniciens. Dans ce contexte, l'évaluation des troubles de la voix et de la parole et leur évolution longitudinale prend une place importante tant au moment du diagnostic que du choix de la prise en charge thérapeutique du patient (incluant la phase de rééducation) et dans son suivi. Dans le même ordre d'idée, les troubles de la production de parole observés au cours de certaines maladies neuro-dégénératives, comme la maladie de Parkinson, sont considérés comme des marqueurs cliniques représentatifs de l'évolution du handicap moteur et cognitif chez les patients. Il est par conséquent important de bien comprendre ces troubles et les observations acoustiques qui en résultent dans le signal de parole afin de pouvoir les caractériser par des processus objectifs, qui font toujours défaut dans la pratique clinique. En effet, l'évaluation perceptive dite à l' « oreille », reste à l'heure actuelle la méthode la plus utilisée en pratique clinique (utilisation de l'échelle GRBAS pour l'évaluation perceptive des dysphonies, de l'échelle UPDRS pour la maladie de Parkinson, ...) malgré son caractère très subjectif, variable et difficilement reproductible largement souligné dans la littérature.

Le LIA travaille depuis une quinzaine d'années sur les troubles de la voix et de la parole, et plus particulièrement sur la manière dont les outils de traitements automatiques peuvent aider les experts cliniciens et phonéticiens dans leur analyse acoustico-phonétique du signal de parole et perceptive des productions des sujets (contrôles et patients). Ces travaux ont pour objectifs de mieux comprendre les dégradations inhérentes aux troubles et de fournir, in fine, des approches d'évaluation objectives de ces dégradations, utilisables en pratique clinique.

Cette thèse s'inscrit dans la continuité de ces travaux et aborde la problématique de l'évaluation perceptive propre à ce contexte et de son implication dans le développement d'approches automatiques d'évaluation objective des troubles de la parole et de la voix.

Prenons l'exemple du corpus d'enregistrements de patients auquel nous avons accès dans le cadre du projet de recherche RUGBI financé par l'ANR. Ce corpus nommé « C2SI » inclut un ensemble de signaux de parole produits par des patients et des sujets contrôles dans le cadre d'un protocole d'enregistrements impliquant différentes tâches de production de parole (production de voyelles tenues et de mots isolés, lecture de phrases ou de textes, parole semi-spontanée à spontanée). Les patients sont atteints de cancers de la tête et du cou et ont subi différents traitements thérapeutiques (chirurgies avec ou sans reconstruction, radio- et/ou chimio-thérapie). Des évaluations perceptives des troubles de la parole chez ces patients ont été réalisées, suivant différents critères, par un jury de 6 experts cliniques sur la base de l'écoute d'une minute de production de parole. Parmi ces critères, nous pouvons citer un score perceptif de sévérité défini comme le degré d'altération du signal de parole (du score 0 – *aucun son identifiable/tous les sons de la parole sont très distordus* - au score 10 - *pas d'altération audible*) ou un score de compréhensibilité du message présent dans un enregistrement de parole (du score 0 – *le message n'est pas compréhensible* – au score 10 – *compréhension de tous les mots et du message*). L'analyse de ces évaluations perceptives et des scores associés sur l'ensemble des patients et contrôles montre des comportements pouvant être très différents d'un juge à l'autre pour un même sujet et, par ailleurs, évolutifs d'un sujet à l'autre.

Comme mentionné plus haut, ces observations sont tout à fait conformes au caractère subjectif de l'évaluation perceptive. Dans ce cadre là, peut-on utiliser ces mesures perceptives comme « Gold standard » que ce soit pour apprendre des modèles pour des tâches d'évaluation des troubles de la parole et pour l'évaluation de leur performance ?

Etant donné qu'aucune autre solution n'est à notre disposition à l'heure d'aujourd'hui, l'objectif de ces travaux est de mieux comprendre ces mesures perceptives au regard du signal de parole, de ses potentielles dégradations et du contexte de production (il n'est pas question d'analyser ici les processus cognitifs en jeu).

Si nous considérons l'ensemble des experts comme fiables et la variabilité de leurs décisions comme le reflet d'une expertise différente, qui se baserait sur des « connaissances » différentes (1) intrinsèques aux productions de parole et à leurs altérations ou (2) extrinsèques (niveaux d'expertise différents, connaissance sur les patients évalués...), il serait intéressant d'identifier sur quelles informations perceptives extraites des productions de parole, ces experts basent leurs évaluations - *même si la question initiale posée pour une évaluation donnée est la même pour tous les experts* - ? Est ce que ces informations perceptives sont similaires en terme qualitatif ? quantitatif ? Est-il possible de retrouver dans les productions de parole des « patrons d'altération » pouvant expliquer des comportements différents/convergenents de chacun de ces experts ?

Un modèle prédictif « explicable » par juge pourrait permettre, par comparaison, d'apporter des réponses à ces questionnements.

Sur la base de ces réponses, la question suivante à laquelle ces travaux devront s'attacher sera d'identifier comment prendre en compte cette variabilité, observée dans les évaluations perceptives produites par un jury d'experts, dans un système automatique amenant à une évaluation rapide, objective et explicable, utilisable en pratique clinique.

Ces travaux de recherche pourront avoir des applications dans d'autres domaines où l'annotation de données nécessaires pour le développement d'un système automatique repose sur une tâche d'évaluation perceptive non triviale et subjective.