



HAL
open science

L'analyse automatisée de la parole et du langage peut-elle contribuer au diagnostic différentiel de dépression et d'apathie chez une population âgée présentant une déficience cognitive légère ?

Aurore Rainouard

► To cite this version:

Aurore Rainouard. L'analyse automatisée de la parole et du langage peut-elle contribuer au diagnostic différentiel de dépression et d'apathie chez une population âgée présentant une déficience cognitive légère ?. Médecine humaine et pathologie. 2019. dumas-02935946

HAL Id: dumas-02935946

<https://dumas.ccsd.cnrs.fr/dumas-02935946>

Submitted on 10 Sep 2020

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.

MÉMOIRE SUR ARTICLE présenté pour l'obtention du
CERTIFICAT DE CAPACITE D'ORTHOPHONISTE
ET DU GRADE DE MASTER 2 D'ORTHOPHONIE

Présenté et soutenu le 14 juin 2019 par

Aurore RAINOUARD

Née le 13 juin 1982

**L'ANALYSE AUTOMATISEE DE LA DE LA PAROLE ET DU LANGAGE
PEUT-ELLE CONTRIBUER AU DIAGNOSTIC DIFFERENTIEL DE
DEPRESSION ET D'APATHIE CHEZ UNE POPULATION AGEE
PRESENTANT UNE DEFICIENCE COGNITIVE LEGERE ?**

Directeur de Mémoire : Auriane Gros

Orthophoniste

Co-directeur(s) de Mémoire : Alexandra König

Psychologue

Nice

2019

Remerciements

Je tiens à remercier ma directrice de mémoire, Auriane Gros, et ma co-directrice de mémoire, Alexandra König, pour leur écoute, leurs conseils et leur confiance à l'image de la liberté qu'elles m'ont octroyée pour réaliser de ce travail.

Ce travail n'aurait pu se réaliser sans l'implication de tous les autres membres du projet, psychologues, orthophonistes, secrétaires, médecin et ingénieurs du CMRR de Nice et en particulier Alexandre Derreumaux, Radia Zeghari et Roxane Fabre ainsi que Nicklas Linz de DFKI et Nelly Darmon du SUPPM du CHU de Nice.

Je remercie également les patients pour leur enthousiasme à contribuer à l'effort de recherche et leur confiance en nous, les soignants équipés de tablettes tactiles. Nous sommes dépositaires de leurs histoires et de leurs voix. Nous ne l'oublions pas.

Je remercie aussi sincèrement mes maîtres de stage, Mmes Florence Lhuilier, Céline Brisack et Nathalie Contencin pour le temps qu'elles m'ont offert alors que les échéances approchaient, leur patience face à un calendrier chaotique et l'humanité de leur pratique. Vous m'avez tant transmis.

Enfin, je tenais à remercier les membres de mon entourage présents durant ces cinq dernières années : Sébastien pour avoir été à mes côtés dans les bons moments comme dans les plus difficiles, la suite c'est maintenant ; Aude et Florie mes sœurs fantastiques, pour leur soutien sans faille malgré mes longs silences ; Timothée et Maxime mes deux soleils du ponant, vos rires et les remous laissés dans votre sillage ont atténué le silence de mon exil oriental ; Marina et François mes chers parents, pour absolument tout, à commencer par la vie ; Eliane et Daniel pour leur accent qui chante et leur porte toujours ouverte ; Michèle « la perle de Nice », ton toit fut si longtemps le mien; et celles qui m'ont permis de constater que l'amitié ne connaît pas d'âge : Laurianne, pour la sincérité de ta grande et belle voix ; Malix, pour ta présence indéfectible, ton appétit et ton courage... immenses ; et Barbara, pour ta grâce, tes rêves qui je le sais deviendront réalité et nos heureux souvenirs d'été.

A mon grand-père, Jean Rainouard, que j'admire.

Tu fus mon premier soutien dans cette entreprise.

Ce travail t'est dédié.

Liste des acronymes

HNR : Harmonic-to-noise ratio

IA: Inventaire Apathie

IAWCD: International Apathy Workgroup Consensus Diagnostic Criteria

MA : Maladie d'Alzheimer

NPI : Inventaire neuropsychiatrique

RAP : Reconnaissance automatique de la parole

SA : Analyse des sentiments (« sentiment analysis »)

TCL: Trouble cognitif léger

TIC : Nouvelles technologies de l'information et de la communication

TNC: Trouble neurocognitif

Sommaire

Introduction.....	4
Article	6
Méthode.....	13
Résultats	19
Discussion	23
Conclusion.....	27
Références	28
Conclusions et perspectives	34
Résumé / Abstract	36

INTRODUCTION

Notre sujet se situe au carrefour des domaines de l'intelligence artificielle, des technologies de la communication, de l'acoustique vocale, du vieillissement pathologique et de l'évaluation médicale. Son choix est motivé par la pratique clinique d'évaluations, réalisées à l'aide d'outils informatiques et numériques au Centre Mémoire de Nice, pour des patients atteints de troubles neurocognitifs et chez lesquels on recherche une apathie, dont la présence augmenterait le risque de développer une maladie d'Alzheimer.

Le diagnostic d'apathie se base traditionnellement sur des hétéro-questionnaires qui sont chronophages et subjectifs. Il est difficile du fait de l'enchevêtrement des troubles comportementaux et cognitifs dus à la pathologie neurodégénérative, mais aussi du fait d'un chevauchement des symptômes avec ceux d'autres pathologies non dégénératives fréquentes dans le vieillissement comme la dépression. Néanmoins, distinguer l'apathie de la dépression s'avère fondamental pour qu'une prise en charge appropriée, médicamenteuse et/ou non médicamenteuse, soit mise en œuvre. C'est pourquoi la Haute Autorité de Santé recommande depuis 2014 la réalisation d'un diagnostic différentiel chez les patients atteints de la maladie d'Alzheimer.

Du point de vue orthophonique ce sujet est justifié par le positionnement particulier du thérapeute lors de la prise en charge de patients âgés avec pathologie neurodégénérative pour lesquels on vise le maintien des capacités le plus longtemps possible. Or, les modifications comportementales et communicationnelles induites par ces pathologies sont progressives et souvent imperceptibles pour le patient comme pour le soignant. Il est pourtant nécessaire de parvenir à les identifier pour les analyser finement, les surveiller au cours du temps et adapter notre prise en charge le cas échéant. La voix est une source d'informations précieuse sur l'état physique et émotionnel du locuteur dont l'étude relève déjà de la compétence orthophonique, notamment lors du bilan des pathologies vocales (dysphonies) et neurologiques (dysarthries). C'est pourquoi l'utilisation du signal de parole nous apparaît comme une mesure pertinente dans un contexte de monitoring, du fait de son objectivité, de sa disponibilité et de la facilité à la prélever en séance.

Dans notre étude, nous cherchons à déterminer si l'analyse de la parole par les nouvelles technologies permet d'identifier des marqueurs vocaux (verbaux et paraverbaux) qui permettraient de différencier une population « avec apathie » d'une population « avec dépression », de différencier une population « avec apathie » d'une population « témoin » et de différencier une population « avec dépression » d'une population « témoin ». Nos hypothèses sont les suivantes : du fait d'un ralentissement

psychomoteur symptomatique observé dans le syndrome apathique comme dans le syndrome dépressif, les voix « pathologiques » (sujets « avec apathie » et sujets « avec dépression ») présentent une prosodie différente de celle des témoins notamment dans le domaine temporel (rythme) ; du fait de l'absence de réaction émotionnelle associée à l'apathie, la voix « apathique » présente une prosodie caractéristique sur le plan rythmique ; du fait d'une humeur excessivement négative associée à la dépression, la parole « dépressive » présente une sémantique caractéristique.

La méthodologie que nous avons mise en place utilise la reconnaissance automatisée de la parole pour extraire à la fois la sémantique sentimentale (analyse des sentiments ou « sentiment analysis ») et la prosodie émotionnelle (fréquence, intensité, stabilité et rythme) des signaux de parole de locuteurs masculins âgés avec des troubles cognitifs légers, alors qu'ils réalisent deux courtes tâches vocales déclenchées par stimulation émotionnelle (un stimulus « négatif » et un stimulus « positif »). Les données sémantiques et acoustiques récoltées, aussi appelées « descripteurs » du signal de parole, sont comparées entre les trois groupes pour mettre en évidence celles qui apparaissent significativement différentes selon le groupe. En découle un panel réduit de descripteurs qui sont ensuite systématiquement comparés entre les groupes appairés deux-à-deux. A l'issue de cette ultime étape, les descripteurs qui continuent de montrer une différence significative entre les paires de groupes sont ensuite discutés comme de potentiels marqueurs vocaux d'une pathologie (apathie ou dépression).

ARTICLE

L'analyse automatisée de la parole et du langage peut-elle contribuer au diagnostic différentiel de dépression et d'apathie chez une population âgée présentant une déficience cognitive légère ?

Soumis le 12 avril 2019, Journal of Voice (Elsevier)

Aurore Rainouard

Département d'Orthophonie de Nice, Faculté de Médecine, Université Nice Sophia Antipolis

CoBTeK (Cognition-Behaviour-Technology) Lab, Université Côte d'Azur

Nicklas Linz

German Research Center for Artificial Intelligence (DFKI)

Radia Zeghari

CoBTeK (Cognition-Behaviour-Technology) Lab, Université Côte d'Azur

Alexandre Derreumaux

CoBTeK (Cognition-Behaviour-Technology) Lab, Université Côte d'Azur

Roxane Fabre

CoBTeK (Cognition-Behaviour-Technology) Lab, Université Côte d'Azur

Centre Hospitalier Universitaire de Nice, Département de Santé Publique

Nelly Darmon

Centre Hospitalier Universitaire de Nice, Service Universitaire de Psychiatrie et de Psychologie Médicale, Université Côte d'Azur

Kevin Rivière

Centre Hospitalier Universitaire de Nice, Service Universitaire de Psychiatrie et de Psychologie Médicale, Université Côte d'Azur

Philippe Robert

Département d'Orthophonie de Nice, Faculté de Médecine, Université Nice Sophia Antipolis

Centre Mémoire de Ressources et Recherches, Centre Hospitalier Universitaire de Nice, Université Côte d'Azur

Association IA

CoBTeK (Cognition-Behaviour-Technology) Lab, Université Côte d'Azur

Alexandra König

INRIA Stars Team, Sophia Antipolis

CoBTeK (Cognition-Behaviour-Technology) Lab, Université Côte d'Azur

Auriane Gros

Département d'Orthophonie de Nice, Faculté de Médecine, Université Nice Sophia Antipolis

CoBTeK (Cognition-Behaviour-Technology) Lab, Université Côte d'Azur

Résumé

Objectif : la pratique du diagnostic différentiel entre l'apathie et la dépression est recommandée pour les patients âgés atteints de troubles neurocognitifs mais demeure compliquée en raison du chevauchement des symptômes. Le risque est alors élevé de mettre en place des traitements inappropriés qui masqueraient une dépression grave sous-jacente ou aggraveraient un processus de vieillissement pathologique. De plus, les échelles d'évaluation traditionnelles ne tiennent pas toujours compte des spécificités du sujet âgé. Parmi les outils issus des nouvelles technologies, la reconnaissance automatique de la parole pourrait constituer une aide précieuse à la prise de décision clinique. Notre objectif est de déterminer si l'analyse croisée de la prosodie et de la sémantique vocales permettrait de différencier une population âgée apathique d'une population âgée dépressive.

Type d'étude : étude analytique et expérimentale

Méthode : trois groupes de patients « apathiques », « dépressifs » et « témoins » (n = 24) présentant un trouble neurocognitif léger ont été enregistrés lors de l'exécution de deux courtes tâches narratives à valence émotionnelle (positive, négative). Les descripteurs verbaux (analyse des sentiments) et paraverbaux (fréquence, énergie, qualité, rythme) relatifs aux caractéristiques sémantiques et acoustiques de la parole ont été extraits automatiquement et examinés entre les groupes.

Résultats : la périodicité moyenne (autocorrélation) combinée au second formant moyen pendant le stimulus négatif permet de différencier les sujets dépressifs des sujets apathiques. Selon la valence du stimulus, les descripteurs liés à l'énergie du signal combinés aux descripteurs de résonance permettent de différencier les sujets apathiques des sujets témoins. Indépendamment du stimulus émotionnel, l'autocorrélation combinée à la valeur minimale de l'énergie permet de différencier les sujets dépressifs des sujets témoins.

Conclusion : l'utilisation de tâches narratives à valence émotionnelle combinées aux technologies de l'information et de la communication ciblant les caractéristiques acoustiques de la parole pourraient contribuer au diagnostic différentiel de la dépression et de l'apathie.

Mots clés : TIC, Voix, Apathie, Dépression, Diagnostic, Personnes âgées

Introduction

Apathie

L'apathie se caractérise par une diminution durable et marquée de la motivation du sujet. Les premiers travaux sur l'apathie l'ont différenciée d'autres syndromes [1] et l'ont modélisée selon trois dimensions : émotionnelle, cognitive et comportementale [2]. Par la suite, des modèles neurobiologiques de comportement orienté vers un but ont considéré les dimensions de l'apathie comme des dysfonctionnements liés à des lésions dans des zones spécifiques du cerveau (cortex préfrontal dorsolatéral, circuit cingulaire antérieur, striatum ventral, circuit orbitofrontal, région frontale antérieure) [3], [4], conduisant à une définition opérationnelle de l'apathie comme « réduction du comportement volontaire et orienté vers un objectif » [5], [6]. Récemment, l'International Apathy Workgroup Consensus Diagnostic Criteria (IAWCD) a remanié la définition et les critères diagnostiques de l'apathie fixés en 2009 [7]: les trois dimensions de l'apathie y sont révisées (comportement et cognition, émotion, interaction sociale), les critères de temps (symptômes présents pendant au moins quatre semaines) et d'impact (handicap fonctionnel évident) restant inchangés [8]. Ces modifications tiennent compte des avancées récentes dans la recherche sur l'apathie et la motivation et permettent d'envisager le développement de nouveaux outils d'évaluation, notamment issus des nouvelles technologies.

Diverses affections peuvent conduire à une apathie comme des affections neurologiques [9]–[13], psychiatriques [14], [15] ou divers affections générales mais c'est dans les maladies neurodégénératives qu'elle est la plus courante [8], [16]. Associée aux troubles neurocognitifs (TNC) majeurs, l'apathie représente un risque de détérioration cognitive plus rapide [17], d'aggravation du pronostic fonctionnel [18], [19] et de perte d'autonomie dans les activités quotidiennes [20], [21]. En présence d'un trouble cognitif léger (TCL), elle augmente le risque de développer une maladie d'Alzheimer (MA) [22], [23]. Même dans une population âgée sans TNC, l'apathie est associée de façon significative à la performance cognitive [24]. Mais ce sont surtout les proches aidants, plus que le patient lui-même, qui perçoivent et souffrent des conséquences de l'apathie de leur proche. En fin de compte, leur qualité de vie est considérablement affectée [25]–[27].

Plusieurs outils d'évaluation sont disponibles dont trois échelles recommandées par l'IAWCD: l'« Apathy Evaluation Scale » (AES) [28], l'Inventaire Apathie (IA) [29] et l'Echelle d'Apathie de Lille (LARS) [30]. Mais au final, l'Inventaire neuropsychiatrique (NPI) [31], créé à l'origine pour fournir une évaluation psychocomportementale chez les patients atteints de démence, reste l'outil de mesure le plus utilisé en recherche clinique en raison de la bonne cohérence interne de son élément « apathie »

[32]. L'apathie se caractérise par une réduction progressive de nombreux comportements comme l'initiation, la persistance, l'intérêt, l'engagement, la perspicacité et la réaction émotionnelle [33]. Cette réduction est progressive et rend difficile la détection de l'apathie, justifiant la pratique d'une évaluation hétérogène de l'apathie. Récemment, des chercheurs ont émis l'hypothèse que la reconnaissance automatique de la parole (RAP) pourrait contribuer à un diagnostic objectif de l'apathie et ont validé l'intérêt d'utiliser des tâches de parole à valence émotionnelle pour détecter l'apathie [34].

L'apathie demeure une construction clinique confrontée à plusieurs écueils : une prévalence variable selon l'étiologie et le stade de la démence ou l'évaluation utilisée, un manque de consensus sur les critères cliniques opérationnels, un chevauchement significatif entre ses manifestations cliniques et d'autres aspects comportementaux de la maladie de Parkinson, de la MA ou de la dépression [35]–[37].

Dépression

La dépression est un trouble psychiatrique de l'humeur dans lequel une personne a de la difficulté à faire face à des événements stressants de la vie, entraînant un sentiment persistant de tristesse, de négativité et de difficulté à faire face aux responsabilités quotidiennes. En 2018, la dépression est classée par l'Organisation Mondiale de la Santé comme la principale cause d'invalidité dans le monde. Le DSM-5 (Manuel diagnostique et statistique des troubles mentaux) est la référence mondiale pour le diagnostic de la dépression [38]. Il exige la présence d'au moins cinq critères cliniques parmi une liste de neuf critères (humeur dépressive, diminution de l'intérêt et du plaisir, perte ou gain de poids, insomnie ou hypersomnie, agitation ou ralentissement psychomoteur, fatigue ou perte d'énergie, sentiment de valeur ou de culpabilité, capacité réduite de penser ou de se concentrer, pensées récurrentes de décès ou de suicide), un critère de durée (pour au moins deux semaines), un critère d'impact (« souffrance significative » ou « déficience fonctionnelle sociale, professionnelle... »).

Chez les personnes âgées vivant en institution, la prévalence de la dépression est préoccupante [39] mais reste sous-diagnostiquée, en raison des caractéristiques sémiologiques propres au vieillissement, telles que les comorbidités multiples avec des troubles cognitifs (altération de la vitesse de traitement et des fonctions exécutives), somatiques (facteurs de risque vasculaire) et psychiatriques (atrophie cérébrale, lésions de la substance blanche), mais aussi des dépendances et troubles du comportement [40], [41], [42]. Ces facteurs, fréquents avec l'âge, modifient l'expression de la dépression, aidant à masquer ses symptômes cardinaux ou, au contraire, à augmenter son intensité [43]. Par conséquent, détecter la dépression paraît d'autant plus important que le sujet est âgé.

Les instruments les plus couramment utilisés pour évaluer la gravité de la dépression sont l'Inventaire de dépression de Beck (BDI) [44] (auto-questionnaire), l'échelle de dépression de Montgomery-Asberg (MADRS) [45] et l'échelle Hamilton Depression Rating (HDRS) [46] (hétéro-questionnaires). Le temps qu'il a fallu pour remplir ces questionnaires a encouragé l'élaboration d'outils de dépistage plus adaptés, comme la Mini International Neuropsychiatric Interview (MINI) [47] et l'Hospital Anxiety and Depression scale (HAD) [48].

La dépression provoque une détresse émotionnelle qui se manifeste par des larmes, de la tristesse, de l'anxiété, des sentiments d'inutilité et de désespoir qui sont autant d'affects négatifs pouvant être exprimés par le système vocal humain. Parce que la dépression a des effets sur la voix, les chercheurs ont étudié la parole dépressive et ont été capables d'identifier des marqueurs prosodiques de la dépression dans la parole tels que les durées de latence, les temps de réponse, la diminution du débit de parole, l'intensité réduite, la diminution des inflexions, l'intonation et l'accentuation réduites, l'articulation lente et la voix monotone [49]–[53]. Depuis quelques décennies, les nouvelles technologies de l'information et de la communication (TIC), basées sur les RAP et les techniques d'apprentissage machine (« machine learning »), permettent d'étudier, de comparer et de classer ces marqueurs acoustiques à des échelles encore plus fines permettant de concevoir des dispositifs d'aide à la décision clinique pour la détection, le diagnostic et le suivi des états dépressifs [54]–[58].

Concernant le diagnostic différentiel, un problème soulevé par les échelles de dépression est qu'elles favorisent l'amalgame entre l'apathie et la dépression. Ce manque de différenciation s'explique par la présence d'éléments évoquant l'apathie qui sont associés à des scores élevés, entraînant des convergences statistiques entre les échelles d'apathie et de dépression qui biaisent le diagnostic [59], [60].

Intérêt du diagnostic différentiel entre apathie et dépression chez la population âgée

En général, la population âgée est souvent sujette à la dépression ou à l'apathie, ou aux deux simultanément [61][59], [60]. Des erreurs de diagnostic peuvent survenir en raison du chevauchement de plusieurs de leurs symptômes cardinaux comme la perte d'intérêt et de motivation ainsi que de certains symptômes végétatifs comme le retard psychomoteur, la fatigue et l'hypersomnie, le manque de lucidité [62]–[64]. De plus, on sait que chez le sujet très âgé, la présentation clinique de la dépression change avec une diminution des éléments dysphoriques en faveur de symptômes apathiques [65], [66]. Pour les personnes âgées en institution, le risque majeur est alors de se trouver négligé, mais aussi d'être exclu des soins [67]. En ce qui concerne la prise en charge pharmacologique, de nombreuses études ont rapporté que les inhibiteurs sélectifs de la recapture de la sérotonine (ISRS) utilisés dans le

traitement de la dépression peuvent augmenter l'apathie et le retrait de l'engagement des patients âgés, ce qui constitue un argument supplémentaire en faveur d'un diagnostic différentiel fiable entre apathie et dépression [68]–[72].

De même, ces deux syndromes peuvent être associés de façon indépendante et différenciée dans la démence [73], [74]. Mais si la dépression chez les personnes âgées est associée à un risque accru de décès par suicide et nécessite donc un traitement approprié [75], seule une faible proportion d'entre elles bénéficie d'un antidépresseur (alors que la majorité bénéficie des psychotropes) [76]. Dans le contexte d'une MA, distinguer l'apathie de la dépression contribuerait donc à une meilleure caractérisation des patients, ce qui conduirait à une meilleure gestion pharmacologique de leurs symptômes comportementaux [77].

Nouveaux outils d'évaluation

Les cliniciens ont besoin d'outils d'évaluation simples, rapides et fiables qu'ils pourraient utiliser systématiquement dans leurs consultations de routine. Ainsi, plusieurs d'entre eux plaident pour le développement de nouveaux outils, potentiellement basés sur les TIC, qui permettraient de dépasser la subjectivité de l'évaluateur [78], [79].

Dans le domaine de l'évaluation de la dépression, les progrès technologiques combinés aux modèles mathématiques appliqués ensuite aux évaluations subjectives traditionnelles ont donné naissance à une nouvelle génération d'instruments appelés tests adaptatifs informatisés (TAI), une sorte d' « échelle 2.0 » dans laquelle les questionnaires sont adaptés aux réponses des patients [80], [81].

Plusieurs techniques pour mesurer l'activité biologique ont été utilisées avec succès par les chercheurs dans le domaine médical : les jeux vidéo pour détecter les troubles potentiels d'initiation, de planification et de motivation dans la maladie de Parkinson [82]; l'actigraphie pour mesurer l'activité locomotrice dans la démence et la schizophrénie [83], [84]; l'activité électrodermale pour détecter la présence/absence d'émotion dans les dépressions bipolaires et unipolaires, la schizophrénie et la démence [85]–[89]; l'oculométrie pour mesurer le mouvement et la fixation oculaire dans les apathies [90].

Dans le cadre de la mesure du comportement moteur et émotionnel, certains capteurs de haute technologie sont particulièrement intéressants pour détecter et surveiller certains troubles car ils peuvent intercepter les actions des sujets en temps réel. Ils peuvent également détecter des changements à un niveau indétectable pour l'œil humain [91]. C'est le cas de l'analyse vidéo automatisée qui utilise des caméras placées dans l'environnement du patient : un algorithme

informatique utilise des signaux vidéo et identifie les variations du positionnement de la personne (posture, vitesse du mouvement, localisation précise) en évaluant son niveau d'autonomie dans l'exécution des activités quotidiennes simples ou la qualité du sommeil [92], [93]. Enfin, la reconnaissance automatique de la parole (RAP) identifie les caractéristiques de la parole qui reflètent la présence de l'émotion dans la parole à travers des changements dans la stabilité vocale, la temporalité de la parole et la hauteur de la fréquence fondamentale [94]. Cette technologie a récemment amélioré la précision du diagnostic des pathologies neurodégénératives [95]–[97].

A terme, la RAP vise à détecter la valence émotionnelle du patient et son niveau de contrôle [98] mais les chercheurs peinent à modéliser et identifier un panel d'éléments acoustiques suffisamment sensibles pour permettre la reconnaissance des émotions dans des conditions réelles, qui seraient valables pour des sous-ensembles réduits ou évaluables quel que soit le corpus analysé [99]. Bien que les défis techniques soient énormes, de nombreuses fonctionnalités acoustiques développées pour l'analyse des signaux vocaux ou l'extraction d'informations musicales permettent aujourd'hui l'étude de la parole émotionnelle [100]–[104] et contribuent au relatif consensus sur les meilleures caractéristiques acoustiques modélisant le discours émotionnel [99], [105], [106] : principalement prosodiques (fréquence, énergie, rythme et qualité du signal sonore) mais aussi spectrales (spectre et cepstre du signal sonore). Certaines sont segmentaires (calculées au niveau de la trame), d'autres sont supra-segmentaires (calculées sur une phrase entière, ou instance émotionnelle), certaines sont calculés sur la voix et d'autres sur des signaux non vocalisés. Enfin, une autre technologie émergente issue de l'informatique affective (« affective computing ») : l'analyse des sentiments (« sentiment analysis »), étudie les émotions humaines en se basant sur des sources textuelles dématérialisées (opinions exprimées sur les réseaux sociaux) tout en nécessitant de grandes quantités de données [107].

A notre connaissance, aucune étude n'a tenté d'utiliser les TIC pour distinguer l'apathie de la dépression. L'étude de la valence du langage par l'analyse automatisée des sentiments (SA) d'une part, et des variations de la prosodie émotionnelle d'autre part, pourrait permettre d'objectiver la présence de ces deux syndromes voire de les distinguer l'un de l'autre. C'est pourquoi nous proposons dans cette étude préliminaire de comparer un panel d'indices paraverbaux (prosodiques) et verbaux (sémantiques) extraits par RAP lors de tâches vocales déclenchées par l'émotion dans le but d'identifier des marqueurs vocaux qui seraient distinctifs de l'apathie et de la dépression. Postulant que la voix reflète l'état émotionnel et que ce dernier est perturbé différemment dans l'apathie et la dépression, nos objectifs consistent principalement à vérifier si certains indices vocaux permettent de différencier une voix « apathique » d'une voix « dépressive », puis à vérifier si des indices vocaux permettent de

distinguer des sujets souffrant d'apathie parmi des témoins et des sujets souffrant de dépression parmi des témoins.

Méthode

Participants

Vingt-quatre hommes atteints de TNC légers ont été sélectionnés pour cette étude (n = 24).

Les participants ont été recrutés au CHU de Nice, au Centre Mémoire de Ressources et Recherches (CMRR) de l'Institut Claude Pompidou et au Service Universitaire de Psychiatrie et de Psychologie Médicale (USPPM) de l'Hôpital Pasteur, dans le cadre de leur suivi (primo consultant ou consultation de routine) ou comme aidant familial principal.

En raison d'un effet de genre très significatif sur la fréquence fondamentale (F0) de la voix [108] et sur la gestion de l'humeur dans la dépression [109], seuls les participants masculins ont été retenus.

La présence de troubles cognitifs a été démontrée chez tous les consultants par une évaluation clinique comprenant le Mini Mental State Examination (MMSE) [110] et une série de tests neuropsychologiques pour évaluer le fonctionnement exécutif et la mémoire.

Pour les consultants en mémoire, la présence de troubles apathiques et dépressifs du comportement a été identifiée via l'IA [29] et le NPI. L'apathie a été diagnostiquée sur la base du score IA total (≥ 4) et l'absence de dépression a été objectivée par un rapport nul ou insignifiant du score « dépression/dysphorie » par rapport au score NPI total ($\leq 20\%$).

Parmi les consultants en psychiatrie, les sujets présentant des épisodes dépressifs sévères (avec/sans symptômes psychotiques) et des épisodes dépressifs légers ont été sélectionnés. Les patients dépressifs ont été diagnostiqués selon le DSM-5.

Pour être inclus dans l'étude, les participants devaient avoir plus de 65 ans, obtenir un score au MMSE compris entre 20 et 26 et parler le français qui devait être leur langue maternelle.

Les sujets présentant des problèmes auditifs majeurs, des troubles du langage, des antécédents de traumatisme crânien, de perte de conscience, un comportement moteur aberrant ou psychotique, ou une dépendance n'ont pas été inclus dans l'étude.

Les participants inclus ont été classés en trois groupes : un groupe « avec dépression » (n = 8) appelé « groupe 1 » (G1), un groupe « témoin » sans dépression et sans apathie (n = 8) appelé « groupe 2 » (G2) et un groupe « avec apathie » (n = 8) appelé « groupe 3 » (G3).

Un consentement écrit de tous les sujets a été obtenu avant l'enregistrement.

L'étude a été approuvée par le Comité d'éthique de Nice (ELEMENT ID RCB 2017-A01896-45, MoTap ID RCB 2017-A01366-47) et menée conformément à la Déclaration d'Helsinki.

Protocole

Notre étude a porté sur deux tâches vocales considérées comme des tâches de parole semi-spontanée. Les tâches d'élocution spontanée présentent plusieurs avantages : elles libèrent le sujet de toute contrainte linguistique en nécessitant moins d'effort cognitif, elles annulent les effets de la pratique et elles induisent un large éventail d'effets émotionnels, notamment lors de la restitution d'événements personnels ayant provoqué des émotions significatives [111]. De plus, la parole spontanée est sensible à la dépression : elle permet d'observer des différences significatives dans la durée des pauses vocales [112].

Les participants étaient invités à effectuer deux tâches vocales : parler d'un événement positif de leur vie (« *pouvez-vous me raconter en une minute, un évènement agréable de votre vie ?* ») et parler d'un événement négatif dans leur vie (« *pouvez-vous me raconter en une minute, un évènement désagréable de votre vie ?* »). Les tâches à valence émotionnelle positive et négative sont pertinentes pour notre étude parce que si les changements dans le comportement vocal de l'orateur peuvent être dus à des changements dans son affect, ces changements dépendent aussi du sujet lui-même [113], [114]. Ainsi, un sujet souffrant de dépression ne réussira pas à réduire les inhibitions liées à l'information négative et il y aura une augmentation du contenu négatif de son discours, quel que soit son niveau de dépression [52], [115].

Les instructions pour les tâches vocales ont été pré-enregistrées par un psychologue du CMRR (voix humaine) et lues par la tablette numérique avant de procéder aux enregistrements, garantissant des instructions standardisées pour les deux expériences.

Les voix des sujets ont été enregistrées avec le microphone interne d'une tablette numérique. L'application a géré et enregistré automatiquement les tâches. Pour améliorer la comparabilité des enregistrements, ils ont été échantillonnés à 22 050 Hz et codés à 16 bits au format wav.

Données

Descripteurs verbaux

Le canal verbal s'intéresse aux unités du langage qui transmettent le sens (les mots). C'est l'un des canaux de communication des émotions qui peut parfois être désynchronisé des comportements

paraverbaux (un visage souriant accompagné d'une voix agressive). Il appartient au domaine de la linguistique et il est de nos jours, fortement investi par la linguistique informatique.

Les données verbales extraites pour cette étude portent sur le niveau sémantique (par le biais d'une analyse des sentiments - SA). La SA est le processus de classification des mots en fonction des émotions ou des opinions qu'ils expriment. La forme la plus simple de SA distingue les sentiments positifs des négatifs, mais d'autres formes plus complexes identifient différents types d'émotions.

Pour notre étude, nous proposons de classer le sentiment général des réponses aux deux tâches. Le lexique « French Expanded Emotion Lexicon » (FEEL) a été utilisé à cette fin [116]. Le FEEL est traduit du lexique « English word-emotion association lexicon » [117], et caractérise positivement ou négativement 14 000 mots français ; il a récemment été utilisé dans des essais de classification par apprentissage machine pour distinguer les réponses des experts médicaux de celles des non-experts dans des forums de santé sur Internet [118].

Pour notre étude, nous avons tenté de déceler le sentiment général se dégageant des réponses des participants. Pour ce faire, nous avons transcrit manuellement les enregistrements à l'aide de l'outil Praat [119]. Tous les mots des transcriptions sont ensuite comparés aux mots présents dans FEEL et si oui, leur valence est enregistrée. Finalement, chaque enregistrement présentait un certain nombre de mots positifs [pos FEEL] et négatifs [neg FEEL]. Ce nombre de mots a été extrait pour chaque tâche. **(Tableau 1)**

Descripteurs paraverbaux

Les descripteurs paraverbaux permettent de décoder la manière dont un message est transmis. Ils sont nombreux et peuvent concerner d'autres signaux que le signal acoustique : expression faciale, mouvements et postures du corps, contact visuel, utilisation de l'espace et de l'environnement pendant la communication, toucher, apparence physique. L'étude du signal acoustique permet d'étudier le comportement vocal des locuteurs [120]. Pour cette étude, les descripteurs paraverbaux ne concernent que le signal acoustique.

Dans la détection des émotions, les descripteurs paraverbaux sont empruntés aux domaines de la prosodie, de la phonétique, de la reconnaissance vocale et du discours. Tous ces descripteurs incluent des paramètres des domaines de la fréquence (par exemple, la F0), de l'amplitude (par exemple, l'énergie), du tempo (par exemple, le rythme) et du spectre (par exemple, l'enveloppe spectrale) [106].

Dans notre étude, seuls les descripteurs acoustiques du niveau prosodique ont été considérés : fréquence, énergie, qualité/stabilité et rythme du signal. Les données ont été automatiquement extraites

du signal audio par l'application. Aucun traitement préalable n'a été nécessaire, ce qui augmente l'applicabilité des résultats dans un contexte clinique et écologique. Les données ont été extraites séparément pour chaque tâche vocale. (**Tableau 1**)

Le logiciel Praat fournit des statistiques sur les longueurs des segments silencieux et des segments sonores, déterminés en fonction de l'intensité (elle-même calculée à partir du signal sonore filtré par un filtre passe-bande), des statistiques sur la hauteur audible (comme F0) et des statistiques sur le tempo vocal (approximation par noyaux de syllabes [121]). Un script Matlab a été utilisé pour calculer le rapport harmonique sur bruit (« Harmonic to Noise Ratio » - HNR) et les statistiques sur les trois premiers formants. Le logiciel openSmile a fourni toutes les autres statistiques [104].

[Tableau 1] Descripteurs - définitions, statistiques et sources

Type	Niveau	Domaine	Intitulés	Statistiques	Définition	Source
VERBAL	Sémantique	Analyse des sentiments	<i>[FEEL]</i>	Neg, Pos		FEEL
PARAVERBAL	Prosodique	Fréquence	<i>[F0]</i>	Max, Min, Moyenne (Moy), Plage, Ecart-type (ET), Variance (Var)	La fréquence fondamentale du signal acoustique.	Praat
			<i>[F0 final]</i>	Max, Min, Moy, ET, Quartiles (Q1, Q2, Q3), Interquartiles (IQR 1-1, IQR 1-3, IQR 2-3)	La fréquence fondamentale du signal acoustique corrigée (suppression des perturbations comme les bruits non vocaux).	openSmile
			<i>[x corr]</i>	Max, Min, Moy	Corrélation croisée du signal de parole. Mesure de la régularité du signal vocal.	Matlab
			Energie	<i>[Loudness]</i>	Max, Min, Moy, ET, Q1, Q2, Q3, IQR 1-1, IQR 1-3, IQR 2-3	Intensité du signal acoustique par Modulation d'Impulsion Codée.
		Qualité	<i>[Formant 1]</i>	Moy, Var		Matlab

	<i>[Formant 2]</i>	Moy, Var		
	<i>[Formant 3]</i>	Moy, Var	Les trois premières harmoniques du signal acoustique.	
	<i>[Harmonic-to-Noise Ratio]</i>		La proportion de sons harmoniques par rapport au bruit dans le signal acoustique.	Matlab
	<i>[Jitter]</i>	Max, Min, Moy, ET, Q1, Q2, Q3, IQR 1-1, IQR 1-3, IQR 2-3	Le jitter est la fluctuation temporelle du signal acoustique.	openSmile
	<i>[Shimmer]</i>	Max, Min, Moy, ET, Q1, Q2, Q3, IQR 1-1, IQR 1-3, IQR 2-3	Le shimmer est la variabilité de l'intensité du son par rapport à l'intensité moyenne.	openSmile
Rythme	<i>[Sound]</i>	Moy, Durée, Ratio de durée du son	Segment sonore.	Praat
	<i>[Pause]</i>	Moy, Durée moyenne des pauses, Taux de pause	Segment de pause.	Praat
	<i>[Words]</i>		Nombre de mots.	Praat
	<i>[Pause to Word Ratio]</i>		Rapport entre les segments de pause et le nombre de mots.	Praat
	<i>[Articulation Rate]</i>		Rapport entre les syllabes et la durée de parole.	Praat
	<i>[Phon Rate]</i>		Rapport entre les pseudo-syllabes et la durée.	Praat
	<i>[Speech Rate]</i>		Rapport entre le nombre de syllabes et la durée.	Praat
	<i>[Syllables]</i>	Syllabes	Nombre de syllabes.	Praat
	<i>[Total Duration]</i>		Durée totale d'enregistrement.	Praat

Analyse statistique

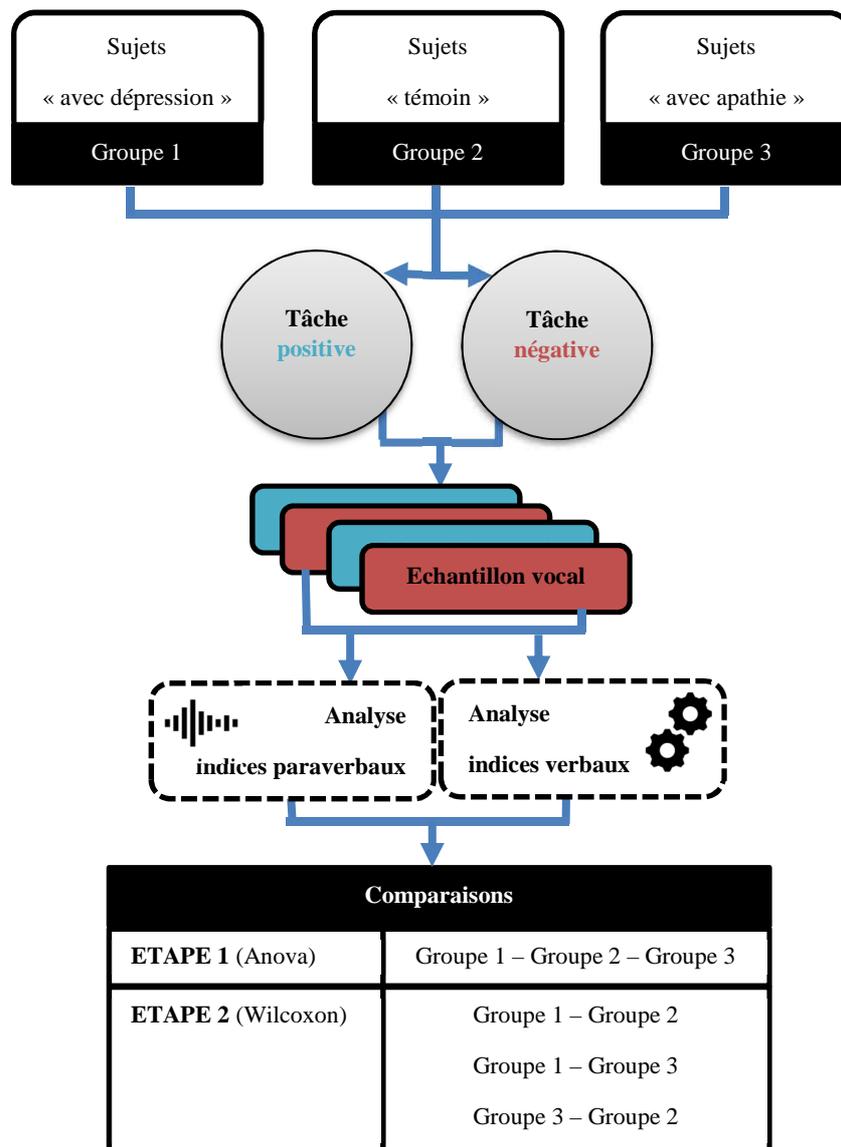
Les variables qualitatives ont été décrites à l'aide de la fréquence et du pourcentage (niveaux d'études) et les variables quantitatives ont été décrites à l'aide de la moyenne et de l'écart-type (âge, score MMSE, descripteurs paraverbaux et verbaux).

Afin de comparer les variables selon les trois groupes de patients, un test du Khi-deux a été effectué pour les variables qualitatives et une analyse de variance (ANOVA) pour les variables quantitatives. Pour les scores d'intérêt, un test statistique de Mann-Whitney-Wilcoxon a été effectué sur chacune des données avec une valeur $p < 0,05$ pour ANOVA, qui a révélé une possible proximité des médianes des données entre les trois paires de groupes (dans un ensemble de 3 groupes).

Une valeur $p < 0,05$ est considérée comme significative.

Toutes les analyses statistiques ont été effectuées à l'aide du logiciel en ligne BiostaTGV, dont les calculs sont effectués avec le logiciel statistique R version 3.5.3.

[Schéma 1] Plan de l'étude



Résultats

Données démographiques

[Tableau 2] Population

Variables	Population totale (n=24)		Groupe 1 (n=8)		Groupe 2 (n=8)		Groupe 3 (n=8)		p-valeur*
	Moyenne	[ET]	Moyenne	[ET]	Moyenne	[ET]	Moyenne	[ET]	
Age	71,5	8,6	63,1	6,1	77,1	5,6	74,3	7,0	<,001
MMSE – score total	23,3	2,1	23,6	2,4	23,1	2,2	23,3	1,8	0,891
	n	(%)	n	(%)	n	(%)	n	(%)	p-valeur**
Etudes (0-5 ans)	6	25	3	38	0	0	3	38	0,099
Etudes (6-12 ans)	10	42	3	38	6	75	1	13	
Etudes (>12 ans)	8	33	2	25	2	25	4	50	
	24		8		8		8		

* ANOVA

** Test du Khi-2

ET : écart-type

Les sujets avaient entre 55 et 84 ans ($71,5 \pm 8,6$ ans). Les hommes souffrant de dépression étaient significativement plus jeunes ($63,13 \pm 6,1$) que les hommes souffrant d'apathie ($74,25 \pm 7$) et les hommes témoins ($77,1 \pm 5,6$). Les trois groupes n'étaient pas différents en termes de score MMSE ($23,3 \pm 2,1$, ce qui correspond à une déficience cognitive légère). Ils n'étaient pas non plus différents sur le plan du niveau d'études.

Notre échantillon de vingt-quatre patients peut donc être considéré comme homogène.

Données acoustiques

Choix des descripteurs verbaux et paraverbaux (parmi trois groupes)

Sur les 3 243 descripteurs acoustiques reçus après leur extraction, 202 descripteurs ont été sélectionnés pour cette étude, dont 198 descripteurs paraverbaux et 4 descripteurs verbaux. L'analyse de variance a été effectuée sur les trois groupes d'échantillons indépendants (G1, G2, G3) pour les deux tâches vocales. Seules les valeurs significatives sont rapportées.

[Tableau 3] Choix des descripteurs verbaux et paraverbaux

		G1 (n=8)		G2 (n=8)		G3 (n=8)		
Tâche	Descripteur	Moyenne [ET]		Moyenne [ET]		Moyenne [ET]		Significativité
Négative	xcorr Moy (*10 ⁶)	0,40	0,95	-19,16	10,95	-8,34	11,63	**
	F0 sma ET	46,52	13,05	38,02	7,83	32,06	9,15	*
	loudness sma Q1	0,06	0,02	0,12	0,07	0,06	0,02	*
	loudness sma Min	0,00	0,00	1539,50	1828,47	517,25	871,19	*
	F2 Moy	4282,41	3095,30	4321,62	2456,01	9410,50	4143,96	**
	Shimmer sma Max	201,00	274,96	936,63	466,40	748,00	585,11	*
Positive	xcorr Moy (*10 ⁶)	0,37	0,00	-19,35	11,47	-7,04	12,84	**
	loudness sma Q1	0,06	0,02	0,12	0,05	0,07	0,03	**
	loudness sma Min	0,00	0,00	1709,88	1370,11	405,13	772,42	**
	F3 Var (/10 ⁶)	14,62	20,27	10,90	10,72	34,06	20,97	*
	Jitter sma ET	0,17	0,02	0,13	0,03	0,15	0,03	*

* p < 0,05

** p < 0,01

ET: écart-type

Q1: quartile 1

Min: position absolue de la valeur minimale (en frames)

Max: position absolue de la valeur maximale (en frames)

sma: moyenne glissante (« simple moving average »)

G1 : sujets « avec dépression » ; G2 : sujets « témoins » ; G3 : sujets « avec apathie »

Au total, entre les trois groupes, onze descripteurs paraverbaux ont présenté des différences significatives de valeurs et aucun descripteur verbal n'a présenté de différence significative de valeurs. Pour les descripteurs paraverbaux, les valeurs les plus significativement différentes appartenaient aux domaines de la *fréquence*, de l'*intensité* et de la *stabilité/qualité* du signal. En revanche, les descripteurs du domaine du *rythme* n'ont montré aucune différence significative dans leurs valeurs entre les trois groupes.

Indépendamment de la valence du stimulus émotionnel les descripteurs des domaines de la *fréquence* (périodicité) et de l'*intensité* (énergie) présentaient des différences de valeurs significatives entre les trois groupes: on retrouve le descripteur « périodicité moyenne » [xcorr Moy] relatif aux variations de fréquence du signal (p < 0,01 dans les deux tâches) ; le descripteur « énergie dans le 1^{er} quartile » [loudness Q1] relatif à l'intensité du signal sonore (p < 0,01 dans la tâche positive et p < 0,05 dans la

tâche négative) ; et le descripteur « valeur minimale de l'énergie » [loudness Min] ($p < 0,01$) pendant le stimulus positif et $p < 0,05$ pendant le stimulus négatif).

Pendant le stimulus négatif, les descripteurs de *stabilité/qualité* ont montré des valeurs d'intérêt entre les trois groupes: les valeurs de « moyenne du second formant » [F2 Moy] étaient significativement différentes ($p < 0,01$) tout comme les valeurs du descripteur « valeur maximale de shimmer » [Shimmer Max] ($p < 0,05$). Les valeurs du descripteur de *fréquence* « dispersion de la fréquence fondamentale » [F0 ET] différaient aussi significativement entre les trois groupes pendant la tâche négative ($p < 0,05$).

Pendant le stimulus positif, les descripteurs de *stabilité/qualité* montraient aussi des valeurs d'intérêt avec des valeurs de « variance du troisième formant » [F3 var] et de « dispersion du jitter » [Jitter ET] qui étaient significativement différentes entre les trois groupes ($p < 0,05$).

Si les descripteurs verbaux n'ont pas présenté de différence de valeurs significatives entre les trois groupes, empêchant leur utilisation comme marqueur vocal d'intérêt dans cette étude, les descripteurs paraverbaux offrent quant à eux plusieurs interprétations possibles, détaillées dans la discussion.

Comparaison des groupes (deux à deux)

[Tableau 4] Comparaison des groupes (deux à deux)

Tâche vocale	Descripteur	Significativité G1-G2-G3	p-valeur					
			G1 - G2	Direction G1	G1 - G3	Direction G1	G3 - G2	Direction G3
Négative	xcorr Moy	**	**	↑	*	↑	0,10	-
	F0 sma ET	*	0,65	-	0,13	-	0,23	-
	loudness sma Q1	*	*	↓	0,88	-	*	↓
	loudness sma Min	*	**	↓	0,08	-	0,17	-
	F2 Moy	**	0,65	-	*	↓	*	↑
	Shimmer sma Max	*	**	↓	0,05	-	0,51	-
Positive	xcorr Moy	**	**	↑	0,43	-	0,17	-
	loudness sma Q1	**	**	↓	0,16	-	0,06	-
	loudness sma Min	**	**	↓	0,17	-	*	↓
	F3 Var	*	0,78	-	0,08	-	*	↑
	Jitter sma ET	*	**	↑	0,10	-	0,14	-

* $p < 0,05$

** $p < 0,01$

ET: écart-type

Q1: quartile 1

Min: position absolue de la valeur minimale (en frames)

Max: position absolue de la valeur maximale (en frames)

sma: moyenne glissante (« simple moving average »)

G1 : sujets « avec dépression » ; G2 : sujets « témoins » ; G3 : sujets « avec apathie »

Au total, entre les trois paires de groupes, dix descripteurs paraverbaux parmi onze ont présenté des différences significatives de valeurs. Le descripteur « dispersion de la fréquence fondamentale » [F0 ET] n'a pas présenté de valeurs d'intérêt lors de cette nouvelle étape de comparaison des descripteurs ce qui n'a pas permis de le considérer comme un marqueur vocal intéressant dans cette étude.

Différencier l'apathie de la dépression : comparaison des groupes « avec dépression » (G1) et « avec apathie » (G3)

Les valeurs des descripteurs de *fréquence* « périodicité moyenne » [xcorr Moy] et de *qualité* du signal « moyenne du second formant » [F2 Moy] étaient significativement différentes dans les deux groupes ($p < 0,05$) mais seulement en condition négative.

Lors du stimulus négatif, les valeurs des « périodicités moyennes » de la fréquence étaient significativement plus grandes (car positives) dans le groupe des sujets « avec dépression » comparativement à celles du groupe des sujets « avec apathie » (car négatives). Alors que les valeurs des « moyennes des seconds formants » étaient significativement plus petites dans le groupe des sujets « avec dépression » comparativement aux sujets « avec apathie ».

Ces résultats suggèrent que les descripteurs « périodicité moyenne » [xcorr Moy] de la fréquence et « moyenne du second formant » [F2 Moy] seraient utiles pour différencier la dépression de l'apathie via l'analyse vocale automatisée.

Détecter l'apathie : comparaison des groupes « avec apathie » (G3) et « témoin » (G2)

Entre le groupe « avec apathie » et le groupe « témoin », quatre descripteurs différaient significativement ($p < 0,05$).

Pendant la tâche positive, les valeurs du descripteur d'*intensité* du signal « valeur minimale d'énergie » [loudness Min] étaient significativement moins élevées dans le groupe « avec apathie » que dans le groupe « témoin » et les valeurs du descripteur de *qualité* du signal « variance du troisième formant » [F3 var] étaient significativement plus élevées dans le groupe « avec apathie » que dans le groupe « témoin ».

Pendant la tâche négative, les valeurs du descripteur d'*intensité* du signal « énergie dans le 1^{er} quartile » [loudness Q1], comme celles du descripteur de *qualité* « moyenne du deuxième formant » [F2 Moy], étaient significativement moins élevées dans le groupe « avec apathie » que dans le groupe « témoin ».

Ces résultats suggèrent que l'observation des descripteurs d'*intensité* et de *qualité*, cette dernière étant influencé par la valence des stimuli, serait utile pour détecter l'apathie via l'analyse vocale automatisée.

Détecter la dépression : comparaison des groupes « avec dépression » (G1) et « témoin » (G2)

Entre le groupe « avec dépression » et le groupe « témoin », huit descripteurs différaient significativement.

Quelle que soit la valence du stimulus émotionnel, les descripteurs de *fréquence* et d'*intensité* du signal présentaient des différences significatives de valeurs entre les deux groupes : les valeurs des « périodicités moyennes » [xcorr Moy] étaient significativement plus grandes (car positives) dans le groupe des sujets « avec dépression » ($p < 0,01$) que dans le groupe des sujets « témoins » ; les « valeurs minimales de l'énergie » [loudness Min] étaient significativement plus petites (car nulles) dans le groupe des sujets « dépressifs » ($p < 0,01$) que celles des « témoins » ; les valeurs d' « énergie dans le 1^{er} quartile » [loudness Q1] étaient aussi significativement plus petites dans le groupe « avec dépression » que celles du groupe « témoin » ($p < 0,01$ lors du stimulus positif et $p < 0,05$ lors du stimulus négatif).

Pendant la tâche négative, le descripteur lié à la *stabilité* du signal « valeur maximale du shimmer » [Shimmer Max] était significativement plus petit dans le groupe « avec dépression » comparativement au groupe « témoin ».

Pendant la tâche positive, le descripteur de *stabilité* du signal « dispersion du jitter » [Jitter ET] était significativement plus grand dans le groupe « dépression » comparativement au groupe « témoin ».

Ces résultats suggèrent que de la plupart des descripteurs paraverbaux de notre panel permettraient de détecter la dépression via l'analyse vocale automatisée.

Discussion

Dans un trouble cognitif léger (TCL), la présence d'une apathie accélère le déclin fonctionnel et cognitif du sujet. Au début d'une maladie d'Alzheimer (MA), la présence d'une apathie est un prédicteur significatif d'une dépression une à quatre années plus tard [17]. Le chevauchement de certains symptômes de l'apathie avec la dépression rend nécessaire un diagnostic différentiel pour lequel des outils font cruellement défaut. L'objectif principal de cette étude était de vérifier si certains

indices vocaux, collectés à l'aide des technologies de l'information et de la communication (TIC), permettraient de différencier une personne souffrant d'apathie d'une personne souffrant de dépression et de différencier un groupe témoin d'un groupe avec apathie ou avec dépression. Nos résultats montrent que l'observation simultanée de deux indices vocaux lors du stimulus négatif permettrait de différencier l'apathie de la dépression. Ils montrent aussi que l'observation de deux paires d'indices vocaux différents selon la valence de la tâche permettrait de différencier un groupe apathique d'un groupe témoin. Ils montrent enfin que l'observation de la plupart de nos descripteurs de fréquence et d'énergie du signal sonore permettrait de différencier un groupe dépressif d'un groupe témoin indépendamment de la valence de la tâche.

Nous avons observé une différence entre le groupe « avec apathie » et le groupe « dépressif » qui portait sur un descripteur de *fréquence* [xcorr Moy] et un descripteur de *qualité* [F2 Moy] lors du stimulus négatif.

Concernant le descripteur de *fréquence* du signal [xcorr Moy] qu'on appelle aussi « autocorrélation du signal » ou corrélation d'un signal par lui-même, il s'agit d'une fonction mathématique utilisée en traitement du signal qui permet d'estimer la périodicité d'un signal périodique (comme le signal de la parole) en repérant le temps pour lequel elle atteint son maximum. Elle permet donc de détecter des régularités, des profils répétés dans un signal. Une autocorrélation « négative » traduit la présence de variations brusques de la fréquence fondamentale (F0) du signal alors qu'une autocorrélation « positive » traduit la présence de variations faibles de la F0. Sur le plan perceptif, une autocorrélation négative évoque donc la présence d'intonations dans la voix, ce qui est le cas de notre groupe « avec apathie ». Inversement une autocorrélation positive évoque une monotonie vocale, ce qui est le cas de notre groupe « avec dépression » à l'évocation d'un souvenir désagréable. En d'autres termes, les sujets « avec dépression » de notre étude présentent une intonation particulièrement appauvrie se traduisant par une certaine monotonie vocale, ce qui est cohérent avec les manifestations cliniques du ralentissement moteur dans la dépression décrites dans la littérature [122]. Cette monotonie vocale est prégnante et distinctive lors de la restitution d'un évènement désagréable.

Concernant le descripteur de *qualité* du signal [F2 Moy], il exprime une valeur moyenne de l'harmonique F2 appelé « formant 2 » ou encore second harmonique. Lors de la mise en vibration des cordes vocales, des oscillations de pression d'air sont propagées dans le conduit vocal (pharynx, cavité buccale, cavité nasale, cavité labiale) sous la forme d'une onde périodique complexe, elle-même

constituée d'un ensemble d'ondes périodiques simples appelées « harmoniques ». Ces harmoniques sont modulées par la taille, la forme, le volume, la matière des cavités (résonateurs) qu'elles traversent. Les résonateurs agissent donc comme des filtres, réduisant certains harmoniques, en renforçant d'autres. Sur le plan perceptif, on parle du timbre, défini comme l'audibilité des harmoniques. Notre groupe « avec apathie » présente des moyennes de F2 deux fois plus grandes que dans notre groupe « avec dépression », ce qui évoque une différence de timbre vocalique entre les deux groupes, caractérisée par une position de la langue plus antérieure chez les sujets « avec apathie » que chez les sujets « avec dépression » lorsqu'ils restituent un évènement désagréable [123], [124]. Une étude des voyelles antérieures présentes dans le discours des sujets souffrant d'apathie pourrait certainement présenter un intérêt dans le cadre du diagnostic de l'apathie. Néanmoins, l'étude objective des indices de qualité vocale comme le timbre, par définition empreinte sonore unique et individuelle, n'a pas encore permis d'accéder à une définition et des outils de mesure consensuels. L'utilisation de l'outil statistique (Matlab) ayant permis de calculer notre descripteur [F2 Moy] peut donc être discutée. C'est pourquoi nos résultats nécessitent d'être confrontés à d'autres expérimentations et outils de mesure issus de laboratoires spécialisés dans l'acoustique et/ou la phonétique. On pense notamment à l'étude des coefficients cepstraux à fréquence Mel (MFCC), descripteurs de haut niveau issus du spectre du signal qui capturent précisément les fréquences de résonance du conduit vocal et sont facilement extractibles avec l'outil OpenSMILE.

Enfin, dans le cadre du diagnostic différentiel entre apathie et dépression, un dernier paramètre est à prendre en compte simultanément à la source sonore (régularité du signal) et au filtre (résonance du signal) : la valence du stimulus. Dans cette étude, la différenciation des voix « apathiques » des voix « dépressives » sur le plan acoustique est permise par la tâche à valence négative. Cette influence de la tâche sur la qualité vocale (et réciproquement) a d'ailleurs été montrée dans d'autres paradigmes expérimentaux [125], [126].

Nous avons aussi observé une différence entre le groupe « avec apathie » et le groupe « témoin » portant sur une paire d'indices différents selon la valence du stimulus.

Lors de la restitution d'un évènement agréable le groupe « avec apathie » diffère du groupe « témoin » sur la paire de descripteurs d'*intensité* [loudness Min] et de *qualité* du signal [F3 var] : le point d'énergie le plus bas atteint par le signal sonore (exprimé par [loudness Min]) est beaucoup plus bas dans le groupe « avec apathie » suggérant une plus faible intensité vocale des sujets « avec apathie » par rapport aux sujets « témoins » ; quant à la dispersion des valeurs autour des moyennes des troisièmes harmoniques du signal (exprimée par [F3 var]), elle plus grande dans le groupe « avec

apathie » que chez les sujets « témoins » à l'évocation d'un un évènement agréable. Sur le plan perceptif (en termes de timbre), cette tendance est difficilement interprétable.

Lors de la restitution d'un évènement désagréable, c'est une nouvelle paire de descripteurs d'*intensité* [loudness Q1] et de *qualité* du signal [F2 Moy] qui distingue les sujets « apathiques » des « témoins » : le niveau d'énergie contenu dans le premier quartile du signal sonore (exprimé par [loudness Q1]) est beaucoup plus bas dans le groupe « avec apathie » suggérant à nouveau une faible intensité vocale des sujets « avec apathie » par rapport aux sujets « témoins » ; quant aux moyennes de F2 ([F2 Moy]), elles sont deux fois plus grandes dans le groupe « avec apathie » suggérant une différence de timbre vocalique entre les deux groupes, caractérisée par une position de la langue plus antérieure chez les sujets « avec apathie » que chez les sujets « témoins » lors de la remémoration d'un un évènement désagréable.

Dans notre étude, l'*énergie* et la *qualité* du signal semblent donc influencées par l'apathie et on parvient à dégager un hypothétique profil vocal d'homme souffrant d'apathie se caractérisant par une intonation vocale légèrement abaissée et par des nuances dans le timbre vocal découlant d'une articulation vocalique plus antérieure.

Enfin, nous avons observé une différence entre le groupe « avec dépression » et le groupe « témoin » qui portait sur la majorité des indices vocaux considérés dans cette étude ; notamment l'autocorrélation et l'énergie du signal sonore dont les valeurs très abaissées quelle que soit la valence de la tâche, leur conférant une valeur distinctive.

Concernant le descripteur de *fréquence*, exprimé par l'autocorrélation qui permet de quantifier les variations de fréquences sur l'ensemble de l'enregistrement [xcorr Moy], le groupe « avec dépression » présente une autocorrélation positive s'exprimant sur le plan perceptif par une monotonie vocale symptomatique dans ce groupe, qu'il évoque un évènement agréable ou un évènement désagréable de sa vie. Ceci est à mettre en lien avec la littérature scientifique déjà évoquée au sujet du ralentissement psychomoteur dans la dépression.

Concernant les descripteurs d'*intensité* exprimés par [loudness Min] et [loudness Q1], ils sont tous les deux fortement abaissés dans le groupe « avec dépression » lors du rappel d'un évènement désagréable ou d'un évènement agréable. Pour ce groupe, le point d'énergie le plus bas atteint par le signal est systématiquement « zéro » (valeur nulle) témoignant de phases où l'intensité vocale est nulle. Cela paraît cohérent avec d'autres études qui ont rapporté une diminution caractéristique de l'intensité et de l'inflexion vocales chez les dépressifs [49], [114], [127]. Cela pourrait aussi signifier la présence

de micro-pauses au sein des signaux vocaux des dépressifs, mais nous n'avons pas pu identifier de publication à ce sujet.

Il est aussi intéressant de noter que les descripteurs verbaux (analyse des sentiments) et paraverbaux temporels (domaine du *rythme* de la parole) n'ont pas pu être explorés dans cette étude puisqu'ils n'ont pas montré de valeurs significativement différentes entre les trois groupes (première étape statistique). Nous estimons que cela est dû à la très petite taille de notre échantillon et considérons qu'il faudrait continuer la recherche de ces marqueurs dans des études futures dans le champ de la dépression (détection de la valence négative du discours prévalente dans les deux tâches).

Cette étude admet d'autres limites puisqu'elle ne propose pas de tâche standardisée de langage (lecture d'un texte) empêchant sa comparaison avec d'autres protocoles. Elle paraît cependant adaptée pour la recherche de troubles émotionnels même si la question éthique des risques associés à ce type de sollicitations se pose du fait de la fragilité des sujets, notamment dépressifs [128]. Notre échantillon étant seulement composé d'hommes, il serait aussi intéressant de reproduire ce protocole avec des voix féminines pour observer un éventuel effet de genre. Enfin, la différence d'âge entre le groupe des « dépressifs » et les autres groupes constitue certainement un écueil important dans cette étude même si l'impact de vieillissement sur la voix reste toujours discuté [111].

Conclusion

Ces résultats donnent plusieurs arguments en faveur de la contribution de l'analyse automatisée de la parole en confirmant la validité d'un panel basique de descripteurs acoustiques. Ils confirment l'intérêt des tâches vocales semi-spontanées déclenchées par stimulus à valence émotionnelle et identifient des marqueurs vocaux potentiels pour le diagnostic et le diagnostic différentiel avec notamment : (1) la périodicité et la qualité (deuxième formant) du signal sonore lors de la tâche à valence négative pour le diagnostic différentiel de l'apathie et de la dépression ; (2) l'intensité et la qualité (deuxième et troisième formants) du signal sonore pour le diagnostic de l'apathie ; (3) la périodicité et l'intensité du signal sonore pour le diagnostic de la dépression.

Références

- [1] R. S. Marin, « Differential diagnosis and classification of apathy », *Am. J. Psychiatry*, vol. 147, n° 1, p. 22-30, janv. 1990.
- [2] R. S. Marin, « Apathy: A neuropsychiatric syndrome », *J. Neuropsychiatry Clin. Neurosci.*, vol. 3, n° 3, p. 243-254, 1991.
- [3] D. T. Stuss, R. Van Reekum, et K. J. Murphy, « Differentiation of states and causes of apathy », in *The neuropsychology of emotion*, New York, NY, US: Oxford University Press, 2000, p. 340-363.
- [4] C. Le Heron, M. A. J. Apps., et M. Husain, « The anatomy of apathy: A neurocognitive framework for amotivated behaviour », *Neuropsychologia*, vol. 118, n° Pt B, p. 54-67, sept. 2018.
- [5] R. Levy et B. Dubois, « Apathy and the functional anatomy of the prefrontal cortex-basal ganglia circuits », *Cereb. Cortex N. Y. N 1991*, vol. 16, n° 7, p. 916-928, juill. 2006.
- [6] S. E. Starkstein et A. F. G. Leentjens, « The nosological position of apathy in clinical practice », *J. Neurol. Neurosurg. Psychiatry*, vol. 79, n° 10, p. 1088-1092, oct. 2008.
- [7] P. H. Robert *et al.*, « Proposed diagnostic criteria for apathy in Alzheimer's disease and other neuropsychiatric disorders », *Eur. Psychiatry*, vol. 24, n° 2, p. 98-104, mars 2009.
- [8] P. H. Robert *et al.*, « Is it time to revise the diagnostic criteria for apathy in brain disorders? The 2018 international consensus group », *Eur. Psychiatry*, vol. 54, p. 71-76, oct. 2018.
- [9] Q.-F. Zhao *et al.*, « The prevalence of neuropsychiatric symptoms in Alzheimer's disease: Systematic review and meta-analysis », *J. Affect. Disord.*, vol. 190, p. 264-271, janv. 2016.
- [10] D. Craufurd, J. C. Thompson, et J. S. Snowden, « Behavioral changes in Huntington Disease. », *Neuropsychiatry. Neuropsychol. Behav. Neurol.*, vol. 14, n° 4, p. 219-226, 2001.
- [11] M. G. Den Brok, J. W. Van Dalen, W. A. Van Gool, E. P. Moll van Charante, R. M. De Bie, et E. Richard, « Apathy in Parkinson's disease: A systematic review and meta-analysis. », *Mov. Disord. Off. J. Mov. Disord. Soc.*, vol. 30, n° 6, p. 759-769, mai 2015.
- [12] S. S. Staekenborg *et al.*, « Behavioural and psychological symptoms in vascular dementia; differences between small- and large-vessel disease », *J. Neurol. Neurosurg. Psychiatry*, vol. 81, n° 5, p. 547-551, mai 2010.
- [13] S. E. Starkstein et J. Pahissa, « Apathy Following Traumatic Brain Injury », *Psychiatr. Clin. North Am.*, vol. 37, n° 1, p. 103-112, mars 2014.
- [14] G. S. Yuen *et al.*, « Apathy in Late-Life Depression: Common, Persistent, and Disabling », *Am. J. Geriatr. Psychiatry*, vol. 23, n° 5, p. 488-494, mai 2015.
- [15] H. Yazbek *et al.*, « The Lille Apathy Rating Scale (LARS): exploring its psychometric properties in schizophrenia », *Schizophr. Res.*, vol. 157, n° 1-3, p. 278-284, août 2014.
- [16] S. Ishii, N. Weintraub, et J. R. Mervis, « Apathy: a common psychiatric syndrome in the elderly », *J. Am. Med. Dir. Assoc.*, vol. 10, n° 6, p. 381-393, juill. 2009.
- [17] S. E. Starkstein, R. Jorge, R. Mizrahi, et R. G. Robinson, « A prospective longitudinal study of apathy in Alzheimer's disease », *J. Neurol. Neurosurg. Psychiatry*, vol. 77, n° 1, p. 8-11, janv. 2006.
- [18] M. Benoit *et al.*, « Apathy and depression in Alzheimer's disease are associated with functional deficit and psychotropic prescription », *Int. J. Geriatr. Psychiatry*, vol. 23, n° 4, p. 409-414, avr. 2008.
- [19] P. Thomas, J. P. Clement, C. Hazif-Thomas, et J. M. Leger, « Family, Alzheimer's disease and negative symptoms », *Int. J. Geriatr. Psychiatry*, vol. 16, n° 2, p. 192-202, févr. 2001.
- [20] J. Vilalta-Franch, L. Calvó-Perxas, J. Garre-Olmo, O. Turró-Garriga, et S. López-Pousa, « Apathy Syndrome in Alzheimer's Disease Epidemiology: Prevalence, Incidence, Persistence, and Risk and Mortality Factors », *J. Alzheimers Dis.*, vol. 33, n° 2, p. 535-543, déc. 2012.
- [21] C. U. Onyike *et al.*, « Epidemiology of Apathy in Older Adults: The Cache County Study », *Am. J. Geriatr. Psychiatry*, vol. 15, n° 5, p. 365-375, mai 2007.
- [22] P. H. Robert *et al.*, « Neuropsychological Performance in Mild Cognitive Impairment with and without Apathy », *Dement. Geriatr. Cogn. Disord.*, vol. 21, n° 3, p. 192-197, 2006.
- [23] E. Papastavrou, A. Kalokerinou, S. S. Papacostas, H. Tsangari, et P. Sourtzi, « Caring for a relative with dementia: family caregiver burden. », *J. Adv. Nurs.*, vol. 58, n° 5, p. 446-457, juin 2007.
- [24] G. Montoya-Murillo, N. Ibarretxe-Bilbao, J. Peña, et N. Ojeda, « The impact of apathy on cognitive performance in the elderly », *Int. J. Geriatr. Psychiatry*, févr. 2019.
- [25] M. De Vugt *et al.*, « Behavioural disturbances in dementia patients and quality of the marital relationship. », *Int. J. Geriatr. Psychiatry*, vol. 18, n° 2, p. 149-154, févr. 2003.

- [26] C. Derouesne, S. Thibault, S. Lagha-Pierucci, V. Baudouin-Madec, D. Ancrì, et L. Lacomblez, « Decreased awareness of cognitive deficits in patients with mild dementia of the Alzheimer type », *Int. J. Geriatr. Psychiatry*, vol. 14, n° 12, p. 1019-1030, déc. 1999.
- [27] C. A. Yeager et L. Hyer, « Apathy in dementia: Relations with depression, functional competence, and quality of life », *Psychol. Rep.*, vol. 102, n° 3, p. 718-722, 2008.
- [28] R. S. Marin, R. C. Biedrzycki, et S. Firinciogullari, « Reliability and validity of the apathy evaluation scale », *Psychiatry Res.*, vol. 38, n° 2, p. 143-162, août 1991.
- [29] P. H. Robert *et al.*, « The Apathy Inventory: assessment of apathy and awareness in Alzheimer's disease, Parkinson's disease and mild cognitive impairment », *Int. J. Geriatr. Psychiatry*, vol. 17, n° 12, p. 1099-1105, déc. 2002.
- [30] P. Sockeel, K. Dujardin, D. Devos, C. Denève, A. Destée, et L. Defebvre, « The Lille apathy rating scale (LARS), a new instrument for detecting and quantifying apathy: validation in Parkinson's disease. », *J. Neurol. Neurosurg. Psychiatry*, vol. 77, n° 5, p. 579-584, mai 2006.
- [31] J. L. Cummings, M. Mega, K. Gray, S. Rosenberg-Thompson, D. A. Carusi, et J. Gornbein, « The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia », *Neurology*, vol. 44, n° 12, p. 2308-2314, déc. 1994.
- [32] J. L. Cummings, « The Neuropsychiatric Inventory: Assessing psychopathology in dementia patients », *Neurology*, vol. 48, n° Issue 5, Supplement 6, p. 10S-16S, mai 1997.
- [33] S. E. Starkstein, G. Petracca, E. Chmerinski, et J. Kremer, « Syndromic Validity of Apathy in Alzheimer's Disease », *Am. J. Psychiatry*, vol. 158, n° 6, p. 872-877, juin 2001.
- [34] N. Linz *et al.*, « Automatic Detection of Apathy using Acoustic Markers extracted from Free Emotional Speech », présenté à 2nd workshop on AI for Aging, Rehabilitation and Independent Assisted Living (ARIAL) @IJCAI'18, 2018.
- [35] G. C. Pluck, « Apathy in Parkinson's disease », *J. Neurol. Neurosurg. Psychiatry*, vol. 73, n° 6, p. 636-642, déc. 2002.
- [36] A. M. Landes, S. D. Sperry, M. E. Strauss, et D. S. Geldmacher, « Apathy in Alzheimer's Disease », *J. Am. Geriatr. Soc.*, vol. 49, n° 12, p. 1700-1707, déc. 2001.
- [37] M. L. Levy *et al.*, « Apathy Is Not Depression », *J. Neuropsychiatry Clin. Neurosci.*, vol. 10, n° 3, p. 314-319, août 1998.
- [38] American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub, 2013.
- [39] S. Kindynis, S. Burlacu, P. Louville, et F. Limosin, « Thérapie des schémas du sujet âgé : impact sur la dépression, l'anxiété et les schémas cognitifs typiques », *L'Encéphale*, vol. 39, n° 6, p. 393-400, déc. 2013.
- [40] R. C. Baldwin et J. O'Brien, « Vascular basis of late-onset depressive disorder. », *Br. J. Psychiatry J. Ment. Sci.*, vol. 180, p. 157-160, févr. 2002.
- [41] K. E. Heek et R. Ho, « The many faces of geriatric depression », *Curr. Opin. Psychiatry*, vol. 21, n° 6, p. 540-545, nov. 2008.
- [42] Y. I. Sheline *et al.*, « Cognitive Function in Late Life Depression: Relationships to Depression Severity, Cerebrovascular Risk Factors and Processing Speed », *Biol. Psychiatry*, vol. 60, n° 1, p. 58-65, juill. 2006.
- [43] T. Gallarda et H. Lôo, « Dépression et personnes âgées », *L'Encéphale*, vol. 35, n° 3, p. 269-280, juin 2009.
- [44] A. T. Beck, C. H. Ward, M. Mendelson, J. Mock, et J. Erbaugh, « An Inventory for Measuring Depression », *Arch. Gen. Psychiatry*, vol. 4, n° 6, p. 561-571, juin 1961.
- [45] S. A. Montgomery, N. Smeyatsky, M. de Ruitter, et D. B. Montgomery, « Profiles of antidepressant activity with the Montgomery-Asberg Depression Rating Scale », *Acta Psychiatr. Scand.*, vol. 72, n° Suppl 320, p. 38-42, 1985.
- [46] M. Hamilton, « A rating scale for depression », *J. Neurol. Neurosurg. Psychiatry*, vol. 23, n° 1, p. 56-62, févr. 1960.
- [47] Y. Lecrubier *et al.*, « The Mini International Neuropsychiatric Interview (MINI): A short diagnostic structured interview: Reliability and validity according to the CIDI », *Eur. Psychiatry*, vol. 12, n° 5, p. 224-231, 1997.

- [48] A. S. Zigmond et R. P. Snaith, « The Hospital Anxiety and Depression Scale », *Acta Psychiatr. Scand.*, vol. 67, n° 6, p. 361-370, 1983.
- [49] J. K. Darby et H. Hollien, « Vocal and speech patterns of depressive patients », *Folia Phoniatr. (Basel)*, vol. 29, n° 4, p. 279-291, 1977.
- [50] P. Hardy, R. Jouvent, et D. Widlöcher, « Speech pause time and the retardation rating scale for depression (ERD). Towards a reciprocal validation », *J. Affect. Disord.*, vol. 6, n° 1, p. 123-127, févr. 1984.
- [51] A. Nilsson, J. Sundberg, S. Ternström, et A. Askenfelt, « Measuring the rate of change of voice fundamental frequency in fluent speech during mental depression. », *J. Acoust. Soc. Am.*, vol. 83, n° 2, p. 716-728, févr. 1988.
- [52] P. Vanger, A. B. Summerfield, B. K. Rosen, et J. P. Watson, « Effects of communication content on speech behavior of depressives », *Compr. Psychiatry*, vol. 33, n° 1, p. 39-41, 1992.
- [53] M. Villain *et al.*, « Affective Prosody and Depression After Stroke: A Pilot Study », *Stroke*, vol. 47, n° 9, p. 2397-2400, 2016.
- [54] N. Cummins, « An investigation of depressed speech detection: features and normalization », in *Interspeech, 2011*, 2011, p. 6-9.
- [55] S. Alghowinem, R. Goecke, M. Wagner, J. Epps, M. Breakspear, et G. Parker, « Detecting depression: A comparison between spontaneous and read speech », in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, p. 7547-7551.
- [56] M. Faurholt-Jepsen *et al.*, « Behavioral activities collected through smartphones and the association with illness activity in bipolar disorder. », *Int. J. Methods Psychiatr. Res.*, vol. 25, n° 4, p. 309-323, déc. 2016.
- [57] F. Or, J. Torous, et J. P. Onnela, « High potential but limited evidence: Using voice data from smartphones to monitor and diagnose mood disorders. », *Psychiatr. Rehabil. J.*, vol. 40, n° 3, p. 320-324, sept. 2017.
- [58] W. Pan *et al.*, « Depression recognition based on speech analysis », *Chin. Sci. Bull.*, vol. 63, n° 20, p. 2081-2092, juill. 2018.
- [59] R. S. Marin, S. Firinciogullari, et R. C. Biedrzycki, « The sources of convergence between measures of apathy and depression », *J. Affect. Disord.*, vol. 28, n° 2, p. 117-124, 1993.
- [60] J. G. Rabkin, S. J. Ferrando, W. van Gorp, R. Rieppi, M. McElhiney, et M. Sewell, « Relationships among apathy, depression, and cognitive impairment in HIV/AIDS », *J. Neuropsychiatry Clin. Neurosci.*, vol. 12, n° 4, p. 451-457, 2000.
- [61] P. Frémont, « Aspects cliniques de la dépression du sujet âgé », *Psychol. Neuropsychiatr. Vieil.*, vol. 2, n° 1, p. 19-27, sept. 2004.
- [62] C. G. Lyketsos, O. Lopez, B. Jones, A. L. Fitzpatrick, J. Breitner, et S. DeKosky, « Prevalence of Neuropsychiatric Symptoms in Dementia and Mild Cognitive Impairment: Results From the Cardiovascular Health Study », *JAMA*, vol. 288, n° 12, p. 1475, sept. 2002.
- [63] A. M. Landes, S. D. Sperry, et M. E. Strauss, « Prevalence of Apathy, Dysphoria, and Depression in Relation to Dementia Severity in Alzheimer's Disease », *J. Neuropsychiatry Clin. Neurosci.*, vol. 17, n° 3, p. 342-349, 2005.
- [64] M. Benoit *et al.*, « Apathy and Depression in Mild Alzheimer's Disease: A Cross-Sectional Study Using Diagnostic Criteria », *J. Alzheimers Dis.*, vol. 31, n° 2, p. 325-334, juill. 2012.
- [65] K. R. R. Krishnan, J. C. Hays, L. A. Tupler, L. K. George, et D. G. Blazer, « Clinical and phenomenological comparisons of late-onset and early-onset depression », *Am. J. Psychiatry*, vol. 152, n° 5, p. 785-788, 1995.
- [66] M. Mehta *et al.*, « Depressive symptoms in late life: Associations with apathy, resilience and disability vary between young-old and old-old », *Int. J. Geriatr. Psychiatry*, vol. 23, n° 3, p. 238-243, 2008.
- [67] J. W. van Dalen, L. L. van Wanrooij, E. P. Moll van Charante, C. Brayne, W. A. van Gool, et E. Richard, « Association of Apathy With Risk of Incident Dementia: A Systematic Review and Meta-analysis », *JAMA Psychiatry*, vol. 75, n° 10, p. 1012, oct. 2018.
- [68] F. Harrison, L. Aerts, et H. Brodaty, « Apathy in Dementia: Systematic Review of Recent Evidence on Pharmacological Treatments », *Curr. Psychiatry Rep.*, vol. 18, n° 11, nov. 2016.
- [69] N. Wongpakaran, R. van Reekum, T. Wongpakaran, et D. Clarke, « Selective serotonin reuptake inhibitor use associates with apathy among depressed elderly: a case-control study », *Ann. Gen. Psychiatry*, vol. 6, n° 1, p. 7, févr. 2007.

- [70] W. J. Barnhart, E. H. Makela, et M. J. Latocha, « SSRI-Induced Apathy Syndrome: A Clinical Review », *J. Psychiatr. Pract.*, vol. 10, n° 3, p. 196-199, mai 2004.
- [71] R. S. Marin, B. S. Fogel, J. Hawkins, J. Duffy, et B. Krupp, « Apathy: A treatable syndrome », *J. Neuropsychiatry Clin. Neurosci.*, vol. 7, n° 1, p. 23-30, 1995.
- [72] R. Hoehn-Saric, J. R. Lipsey, et D. R. McLeod, « Apathy and Indifference in Patients on Fluvoxamine: and Fluoxetine », *J. Clin. Psychopharmacol.*, vol. 10, n° 5, p. 343-345, oct. 1990.
- [73] K. Palmer *et al.*, « Neuropsychiatric Predictors of Progression from Amnesic-Mild Cognitive Impairment to Alzheimer's Disease: The Role of Depression and Apathy », *J. Alzheimers Dis.*, vol. 20, n° 1, p. 175-183, mars 2010.
- [74] E. Richard *et al.*, « Symptoms of Apathy Are Associated with Progression from Mild Cognitive Impairment to Alzheimer's Disease in Non-Depressed Subjects », *Dement. Geriatr. Cogn. Disord.*, vol. 33, n° 2-3, p. 204-209, 2012.
- [75] M. Mouquet, V. Bellamy, et V. Carasco, « Suicides et tentatives de suicide en France », *DREES Etudes Résultats*, vol. 488, p. 8, 2006.
- [76] P. Frémont, « L'utilisation des psychotropes chez le sujet âgé: trop ou trop peu?: Médicaments et sujets âgés », *Presse Médicale*, vol. 28, n° 32, p. 1794-1799, 1999.
- [77] P. Tagariello, P. Girardi, et M. Amore, « Depression and apathy in dementia: Same syndrome or different constructs? A critical review », *Arch. Gerontol. Geriatr.*, vol. 49, n° 2, p. 246-249, sept. 2009.
- [78] R. David *et al.*, « Place des nouvelles technologies en neuropsychiatrie », *NPG Neurol. - Psychiatr. - Gériatrie*, vol. 16, n° 96, p. 353-357, déc. 2016.
- [79] P. H. Robert *et al.*, « Recommendations for ICT use in Alzheimer's disease assessment: Monaco CTAD Expert Meeting », *J. Nutr. Health Aging*, vol. 17, n° 8, p. 653-660, 2013.
- [80] R. D. Gibbons, D. J. Weiss, E. Frank, et D. Kupfer, « Computerized Adaptive Diagnosis and Testing of Mental Health Disorders », *Annu. Rev. Clin. Psychol.*, vol. 12, n° 1, p. 83-104, mars 2016.
- [81] C. Botella *et al.*, « An Internet-based program for depression using activity and physiological sensors: efficacy, expectations, satisfaction, and ease of use », *Neuropsychiatr. Dis. Treat.*, p. 393, févr. 2016.
- [82] W. Fitts, L. Massimo, N. Lim, M. Grossman, et N. Dahodwala, « Computerized assessment of goal-directed behavior in Parkinson's disease. », *J. Clin. Exp. Neuropsychol.*, vol. 38, n° 9, p. 1015-1025, nov. 2016.
- [83] R. David *et al.*, « Decreased daytime motor activity associated with apathy in Alzheimer disease: an actigraphic study », *Am. J. Geriatr. Psychiatry Off. J. Am. Assoc. Geriatr. Psychiatry*, vol. 20, n° 9, p. 806-814, sept. 2012.
- [84] A. Kluge *et al.*, « Combining actigraphy, ecological momentary assessment and neuroimaging to study apathy in patients with schizophrenia. », *Schizophr. Res.*, vol. 195, p. 176-182, mai 2018.
- [85] S. Fedor, P. Chau, N. Bruno, R. W. Picard, et J. Camprodon, « Can We Predict Depression From the Asymmetry of Electrodermal Activity? », *Iproceedings*, vol. 2, n° 1, p. e23, 2016.
- [86] A. Greco, G. Valenza, A. Lanata, G. Rota, et E. P. Scilingo, « Electrodermal Activity in Bipolar Patients during Affective Elicitation », *IEEE J. Biomed. Health Inform.*, vol. 18, n° 6, p. 1865-1873, nov. 2014.
- [87] S. Park et K. Kim, « Physiological reactivity and facial expression to emotion-inducing films in patients with schizophrenia. », *Arch. Psychiatr. Nurs.*, vol. 25, n° 6, p. e37-47, déc. 2011.
- [88] A. Y. Kim *et al.*, « Automatic detection of major depressive disorder using electrodermal activity », *Sci. Rep.*, vol. 8, nov. 2018.
- [89] G. Perugia, D. Rodríguez-Martín, M. D. Boladeras, A. C. Mallofré, E. Barakova, et M. Rauterberg, « Electrodermal activity: Explorations in the psychophysiology of engagement with social robots in dementia », in *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2017, p. 1248-1254.
- [90] S. A. Chau, J. Chung, N. Herrmann, M. Eizenman, et K. L. Lancôt, « Apathy and Attentional Biases in Alzheimer's Disease », *J. Alzheimers Dis. JAD*, vol. 51, n° 3, p. 837-846, 2016.
- [91] A. König *et al.*, « A review of current information and communication technologies: can they be used to assess apathy? Current and new methods for the assessment of apathy », *Int. J. Geriatr. Psychiatry*, vol. 29, n° 4, p. 345-358, avr. 2014.
- [92] A. König *et al.*, « Ecological Assessment of Autonomy in Instrumental Activities of Daily Living in Dementia Patients by the Means of an Automatic Video Monitoring System », *Front. Aging Neurosci.*, vol. 7, juin 2015.

- [93] A. König *et al.*, « The role of information and communication technologies in clinical trials with patients with Alzheimer's disease and related disorders », *Front. Aging Neurosci.*, vol. 7, juin 2015.
- [94] I. R. Murray et J. L. Arnott, « Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion », *J. Acoust. Soc. Am.*, vol. 93, n° 2, p. 1097-1108, 1993.
- [95] A. König *et al.*, « Automatic speech analysis for the assessment of patients with predementia and Alzheimer's disease », *Alzheimers Dement. Diagn. Assess. Dis. Monit.*, vol. 1, n° 1, p. 112-124, mars 2015.
- [96] N. Cummins, S. Scherer, J. Krajewski, S. Schnieder, J. Epps, et T. F. Quatieri, « A review of depression and suicide risk assessment using speech analysis », *Speech Commun.*, vol. 71, p. 10-49, juill. 2015.
- [97] T. Taguchi *et al.*, « Major depressive disorder discrimination using vocal acoustic features », *J. Affect. Disord.*, vol. 225, p. 214-220, janv. 2018.
- [98] M. Goudbeek et K. Scherer, « Beyond arousal: valence and potency/control cues in the vocal expression of emotion », *J. Acoust. Soc. Am.*, vol. 128, n° 3, p. 1322-1336, sept. 2010.
- [99] M. Tahon et L. Devillers, « Towards a Small Set of Robust Acoustic Features for Emotion Recognition: Challenges », *IEEEACM Trans. Audio Speech Lang. Process.*, vol. 24, n° 1, p. 16-28, janv. 2016.
- [100] A. Batliner *et al.*, « Whodunnit – Searching for the most important feature types signalling emotion-related user states in speech », *Comput. Speech Lang.*, vol. 25, n° 1, p. 4-28, janv. 2011.
- [101] B. Schuller, A. Batliner, S. Steidl, et D. Seppi, « Recognising realistic emotions and affect in speech: State of the art and lessons learnt from the first challenge », *Speech Commun.*, vol. 53, n° 9-10, p. 1062-1087, 2011.
- [102] B. Schuller *et al.*, « Paralinguistics in speech and language—State-of-the-art and the challenge », *Comput. Speech Lang.*, vol. 27, n° 1, p. 4-39, janv. 2013.
- [103] F. Eyben, M. Wöllmer, et B. Schuller, « Opensmile: the munich versatile and fast open-source audio feature extractor », in *Proceedings of the international conference on Multimedia - MM '10*, Firenze, Italy, 2010, p. 1459.
- [104] F. Eyben, F. Wenginger, F. Gross, et B. Schuller, « Recent Developments in openSMILE, the Munich Open-source Multimedia Feature Extractor », in *Proceedings of the 21st ACM International Conference on Multimedia*, New York, NY, USA, 2013, p. 835–838.
- [105] F. Eyben *et al.*, « The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for Voice Research and Affective Computing », *IEEE Trans. Affect. Comput.*, vol. 7, n° 2, p. 190-202, avr. 2016.
- [106] L. Devillers, M. Tahon, M. A. Sehili, et A. Delaborde, « Détection des états affectifs lors d'interactions parlées : robustesse des indices non verbaux », *Trait. Autom. Lang. TAL*, vol. Volume 55, n° n°2 (pages 123 à 149), p. 27, 2014.
- [107] E. Cambria, « Affective Computing and Sentiment Analysis », *IEEE Intell. Syst.*, vol. 31, n° 2, p. 102-107, mars 2016.
- [108] I. R. Titze, « Physiologic and acoustic differences between male and female voices », *J. Acoust. Soc. Am.*, vol. 85, n° 4, p. 1699-1707, avr. 1989.
- [109] S. Nolen-Hoeksema, « Sex differences in unipolar depression: Evidence and theory », *Psychol. Bull.*, vol. 101, n° 2, p. 259-282, 1987.
- [110] M. F. Folstein, P. R. McHugh, et S. E. Folstein, « Mini-mental state (MMSE) Journal of Psychiatric Research », 1975.
- [111] C. Lortie, « Le vieillissement de la voix: de la production à l'évaluation », Laval, Laval, Canada, 2017.
- [112] M. Alpert, E. R. Pouget, et R. R. Silva, « Reflections of depression in acoustic measures of the patient's speech », *J. Affect. Disord.*, vol. 66, n° 1, p. 59-69, sept. 2001.
- [113] K. R. Scherer, T. Johnstone, et G. Klasmeyer, « Vocal expression of emotion », in *Handbook of affective sciences*, New York, NY, US: Oxford University Press, 2003, p. 433-456.
- [114] H. Ellgring et K. R. Scherer, « Vocal indicators of mood change in depression », *J. Nonverbal Behav.*, vol. 20, n° 2, p. 83-110, 1996.
- [115] E. Goeleven, R. De Raedt, S. Baert, et E. H. W. Koster, « Deficient inhibition of emotional information in depression », *J. Affect. Disord.*, vol. 93, n° 1-3, p. 149-157, juill. 2006.
- [116] A. Abdaoui, J. Azé, S. Bringay, et P. Poncelet, « FEEL: a French Expanded Emotion Lexicon », *Lang. Resour. Eval.*, vol. 51, n° 3, p. 833-855, sept. 2017.
- [117] S. M. Mohammad et P. D. Turney, « Crowdsourcing a word–emotion association lexicon », *Comput. Intell.*, vol. 29, n° 3, p. 436-465, 2013.

- [118] A. Abdaoui, J. Azé, S. Bringay, N. Grabar, et P. Poncelet, « Expertise in French health forums », *Health Informatics J.*, vol. 25, n° 1, p. 17-26, mars 2019.
- [119] P. Boersma, « Praat, a system for doing phonetics by computer », *Glott Int.*, vol. 5, 2002.
- [120] A. Mahdhaoui, « Analysis of Social Signals for modelling Face-to-Face Interaction », Theses, Université Pierre et Marie Curie - Paris VI, 2010.
- [121] N. De Jong et T. Wempe, « Praat script to detect syllable nuclei and measure speech rate automatically. », *Behav. Res. Methods*, vol. 41, n° 2, p. 385-390, mai 2009.
- [122] D. Bennabi, P. Vandel, C. Papaxanthis, T. Pozzo, et E. Haffen, « Psychomotor Retardation in Depression: A Systematic Review of Diagnostic, Pathophysiologic, and Therapeutic Implications », *BioMed Res. Int.*, vol. 2013, 2013.
- [123] G. Fant, *Acoustic Theory of Speech Production: With Calculations Based on X-Ray Studies of Russian Articulations*. Walter de Gruyter, 1970.
- [124] K. N. Stevens, *Acoustic Phonetics*. MIT Press, 2000.
- [125] C. Gendrot, « Influence de la qualité de la voix sur la perception de quatre émotions simulées : une étude perceptive et physiologique », *Rev. PAROLE*, vol. 1, n° 13, p. 1-18, nov. 2004.
- [126] M. Tahon, G. Degottex, et L. Devillers, « Usual voice quality features and glottal features for emotional valence detection », présenté à Speech Prosody, 2012.
- [127] D. Widlöcher et G. Liébert, *Les logiques de la dépression*. 1983.
- [128] J. A. Harrigan, K. S. Lucic, et R. Rosenthal, « Retelling Anxious Events: Effects on Trait and State Anxiety », *Personal. Individ. Differ.*, vol. 12, n° 9, p. 917-927, 1991.

CONCLUSIONS ET PERSPECTIVES

Notre étude a montré l'existence de marqueurs vocaux potentiels pour le diagnostic différentiel de l'apathie et de la dépression : l'étude des paramètres prosodiques de fréquence (périodicité du signal) et de qualité du signal (moyennes des deuxièmes formants) lors du stimulus négatif révèle des profils vocaux caractéristiques avec des intonations préservées (autocorrélation négative) s'accompagnant d'un timbre vocalique marqué par une certaine résonance buccale (valeurs élevées des formants 2) dans la voix « apathique » ; et une monotonie vocale (autocorrélation positive) sans résonance particulière au niveau des cavités supra-laryngées dans la voix « dépressive ». D'autres paramètres prosodiques d'intensité et de qualité du signal semblent intéressants pour détecter l'apathie seule : l'étude des points d'énergie les plus bas du signal et de sa qualité (moyennes hautes des deuxièmes formants en condition négative et forte variance des troisièmes formants en condition positive) permettent de différencier une voix « apathique » d'une voix « témoin ». Enfin, l'étude de la périodicité du signal et de l'énergie qu'il contient semble permettre de détecter la dépression caractérisée par une monotonie vocale (autocorrélation positive) et une forte diminution de l'intensité vocale (point d'énergie du signal relevés à zéro). Les résultats confirment aussi la validité de la plupart des descripteurs acoustiques constituant notre panel ainsi que l'intérêt des tâches vocales semi-spontanées déclenchées par stimulus à valence émotionnelle pour la reconnaissance des émotions.

Néanmoins, ni l'analyse des paramètres sémantiques (analyse des sentiments) ni l'analyse des paramètres prosodiques de rythme n'ont pu être menées puisque l'analyse statistique les a exclus du champ de notre étude. Les hypothèses relatives à une organisation temporelle spécifique des signaux de parole dans l'apathie ou la dépression n'ont donc pas pu être investiguées. Une étude portant sur un plus grand échantillon de voix serait nécessaire pour analyser la temporalité des signaux de parole, confirmer nos résultats préliminaires et renforcer notre panel de descripteurs, en ajoutant notamment les coefficients cepstraux (MFCC) qui capturent très précisément les fréquences de résonance du conduit vocal.

A l'issue de ce travail de recherche, exploratoire et résolument transversal, les perspectives pour la clinique orthophonique se déclinent différemment selon qu'on envisage un exercice salarié ou libéral. Pour un établissement spécialisé dans le diagnostic et la prise en charge des maladies neurodégénératives ou psychiatriques, un outil d'aide au diagnostic basé sur la reconnaissance automatique de la parole permettrait de procéder à une partie des évaluations pluridisciplinaires. Ce gain de temps présenterait un intérêt majeur pour les équipes. Pour l'orthophoniste en particulier, certaines tâches linguistiques d'expression orale telles que la fluence phonétique/sémantique, la

description d'image, les répétitions de mots/phrases pourraient être enregistrées et préciser un état cognitif et émotionnel (temps de réponse, latences, « affect burst ») selon des données quantitatives et qualitatives analysées en temps réel. Dans le cadre d'un exercice libéral comprenant des interventions en EHPAD ou à domicile, ce type d'outil peut aussi s'avérer utile pour le dépistage et/ou le monitoring des résidents. D'autres applications à visée rééducative peuvent aussi être ajoutées (jeux sérieux, stimulation cognitive, communication améliorée et alternative, musique et chansons). En effet, l'orthophoniste peut être confronté au profil vocal typique de la dépression à l'occasion d'une séance. Un outil portable de reconnaissance et d'analyse de la parole lui fournirait une mesure objective lui permettant d'orienter au plus vite le patient vers un médecin.

Pour conclure cette recherche dédiée aux seuls signaux sonores émis sous l'influence de stimuli émotionnels, il nous paraît important d'évoquer l'existence de nombreuses études s'intéressant au traitement multimodal de l'émotion qui inclue des signaux biologiques tels que le son mais aussi les mouvements (expression faciale, gestes, écriture) et qui s'appuient sur l'intelligence artificielle et ses algorithmes pour parvenir à les comprendre et, à terme, les reproduire. Ces avancées concernent aussi les orthophonistes qui prennent en charge divers troubles de la communication (pragmatique) qui intègrent bien souvent une composante émotionnelle.

Résumé / Abstract

Rainouard Aurore

Titre du mémoire : L'analyse automatisée de la parole et du langage peut-elle contribuer au diagnostic différentiel de dépression et d'apathie chez une population âgée présentant une déficience cognitive légère ?

Résumé : Le diagnostic différentiel entre apathie et dépression est compliqué par le chevauchement des symptômes et les évaluations chronophages et subjectives. Le signal vocal informe sur l'état interne du locuteur et se prélève facilement grâce aux nouvelles technologies ce qui en fait une mesure objective utilisable par les cliniciens, dont les orthophonistes, soucieux du dépistage, diagnostic ou monitoring des troubles du comportement et/ou de l'humeur de leurs patients. En utilisant la reconnaissance automatique de la parole nous avons cherché à identifier des marqueurs vocaux qui permettraient de différencier une population « avec apathie » d'une population « avec dépression » ou de différencier des sujets « avec apathie » ou « avec dépression » de sujets « témoins ».

Trois groupes de patients « apathiques », « dépressifs » et « témoins » (n = 24) présentant un trouble cognitif léger ont été enregistrés lors de deux tâches narratives déclenchées par stimulation émotionnelle. La prosodie émotionnelle et la sémantique sentimentale des signaux ont été automatiquement extraites et examinées entre les groupes.

Nos résultats confirment l'intérêt des tâches vocales déclenchées par stimulation émotionnelle pour la démarche diagnostique des troubles de l'humeur, plaident en faveur de la contribution de l'analyse automatisée de la parole et identifient plusieurs marqueurs vocaux potentiels: la périodicité et la qualité (deuxième formant) du signal sonore lors de la tâche à valence négative (diagnostic différentiel apathie-dépression) ; l'intensité et la qualité (deuxième et troisième formants) du signal sonore (diagnostic apathie) ; la périodicité et l'intensité du signal sonore (diagnostic dépression).

Mots-clés : TIC, Voix, Apathie, Dépression, Diagnostic, Personnes âgées

Title: Can automated speech and language analysis contribute to the differential diagnosis of depression and apathy in an elderly population with mild cognitive impairment?

Abstract: The differential diagnosis between apathy and depression is complicated by overlapping symptoms and time-consuming and subjective assessments. The voice signal provides information on the internal state of the speaker and can be easily extracted using new technologies, making it an objective measure that can be used by clinicians, including speech-language pathologists, concerned about screening, diagnosis or monitoring of behavioural and/or mood disorders in their patients. Using automatic speech recognition, we sought to identify speech markers that would differentiate a population "with apathy" from a population "with depression" or would differentiate subjects "with apathy" or "with depression" from "control" subjects.

Three groups of "apathetic", "depressive" and "control" patients (n = 24) with mild neurocognitive disorders were recorded during two narrative tasks emotionally triggered. The emotional prosody and sentimental semantics of the signals were automatically extracted and examined between the groups.

Our results confirm the value of emotionally triggered speech tasks for the diagnosis of mood disorders, argue in favour of the contribution of automated speech analysis and identify several potential speech markers: the periodicity and quality (second formant) of the speech signal during the negative-valent task (apathy-depression differential diagnosis); the intensity and quality (second and third formants) of the speech signal (apathy diagnosis); the frequency and intensity of the speech signal (depression diagnosis).

Keywords: ICT, Voice, Apathy, Depression, Diagnostic, Elderly