

Université de Montréal

**Marqueurs discursifs de neurodégénérescence liée à la pathologie Alzheimer**

*Par*

Antoine Slegers

Département de Psychologie, Faculté des Arts et des Sciences

Thèse présentée en vue de l'obtention du grade de Ph.D en Psychologie,  
option neuropsychologie clinique, profil recherche et intervention

Dépôt initial : juin 2021

© Antoine Slegers, 2021



Université de Montréal

Département de Psychologie, Faculté des Arts et des Sciences

---

*Cette thèse intitulée*

**Marqueurs discursifs de neurodégénérescence liée à la pathologie Alzheimer**

*Présenté par*

**Antoine Slegers**

*A été évalué(e) par un jury composé des personnes suivantes*

**Nadia Gosselin**

Présidente-rapporteuse

**Simona Maria Brambati**

Directrice de recherche

**Elaine de Guise**

Membre du jury

**Vanessa Taler**

Examinatrice externe



## Résumé

La maladie d'Alzheimer (MA) et les aphasies progressives primaires (APP) s'accompagnent de perturbations du langage expressif parfois subtiles, mais précoces dans l'évolution de ces maladies neurodégénératives. Considérés dans une approche automatisée, ces changements pourraient constituer des marqueurs de dégénérescence identifiés de façon non invasive et peu onéreuse. À ce titre, ils font l'objet d'études visant à automatiser leur utilisation clinique. Cependant, l'intégration des marqueurs langagiers à une approche diagnostique centrée sur les biomarqueurs reste à faire. À cette fin, la présente thèse a deux objectifs. D'abord, recenser systématiquement les marqueurs du discours qui distinguent le mieux les personnes avec une MA de témoins en santé. Ensuite, appliquer une approche automatisée et à un large éventail de marqueurs de discours pour identifier, dans un groupe hétérogène de patients avec une APP, lesquels ont une pathologie Alzheimer sous-jacente. Afin de mettre en contexte ces deux objectifs, nous proposons une introduction générale comprenant les éléments suivants : la pathophysiologie de la MA et des APP, le rôle croissant des biomarqueurs dans la prise de décision clinique dans les maladies neurodégénératives, les études pionnières du discours en neurodégénérescence, ainsi que de récentes études computationnelles sur les marqueurs de discours dans la MA et les APP.

Nos résultats font émerger un patron multidimensionnel (acoustique, lexical, syntaxique, sémantique et pragmatique) de changements langagiers qui distinguent les personnes avec une MA de témoins en santé, avec une prépondérance des marqueurs lexicosémantiques. Dans le groupe de patients avec une APP avec une imagerie amyloïde positive ou négative, nous mesurons ensuite le pouvoir de classification d'un court échantillon de discours et montrons qu'il peut être avantageusement comparé à d'autres biomarqueurs. Nous discutons du patron spécifique de

marqueurs discriminants pour ce sous-groupe de patients, notamment l'importance des marqueurs psycholinguistiques pour prédire le résultat de l'imagerie amyloïde à partir du discours.

**Mots-clés** : maladie d'Alzheimer, discours, aphasie progressive primaire, amyloïde, traitement automatique du langage naturel, apprentissage automatique, psycholinguistique

## Abstract

Alzheimer's disease (AD) and primary progressive aphasia (PPA) feature changes in expressive language that appear early in the course of the disease. Within an automated analysis framework, these language changes could offer a non-invasive and inexpensive alternative to the collection of biomarkers which are not readily available in most settings. Current research is thus focused on the automated analysis of language data for clinical use. The usefulness of connected speech (CS) markers has not yet been established in a diagnostic perspective focused on biomarkers. To this aim, the present thesis contains two phases. First, we systematically review the CS markers that best differentiate persons with AD from healthy controls. Second, we automatically extract a wide array of CS markers in a heterogeneous group of PPA patients by combining expert knowledge and the latest natural language processing software. A machine-learning classification approach identifies PPA patients for the presence of underlying AD pathology. The most discriminant CS features are identified. To integrate the two phases of the thesis, we provide a general introduction with the following sections: the pathophysiology of AD and PPAs, the growing importance of biomarkers in clinical decision-making for neurodegenerative diseases, the seminal studies of CS in neurodegenerative diseases, and the latest computational studies of CS markers in AD and PPA. Our results bring forth a multidimensional pattern (acoustic, lexical, syntactic, semantic, pragmatic) of language changes that distinguish people with AD from healthy controls, with an emphasis on lexical-semantic features. In the group of PPA patients with either positive or negative amyloid imaging, we then describe the classificatory power of a short sample of CS and show that it compares favorably to other biomarkers. We discuss the specific pattern of discriminant markers for this subgroup of patients, in particular the role of psycholinguistics.

**Keywords:** Alzheimer's disease, connected speech, primary progressive aphasia, amyloid, natural language processing, machine learning, psycholinguistics



## Table des matières

Résumé .....	5
Abstract .....	7
Table des matières .....	9
Liste des tableaux .....	11
Liste des figures .....	13
Liste des sigles et abréviations .....	15
Remerciements .....	19
Chapitre 1 – Contexte théorique.....	21
1. Introduction générale.....	21
1.1 Manifestations cliniques de la maladie d’Alzheimer et troubles de langage associés.....	21
1.2 Pathophysiologie et biomarqueurs dans les classifications récentes de la maladie d’Alzheimer.....	24
1.3 Pathophysiologie et biomarqueurs dans les classifications récentes des aphasies progressives primaires.....	27
1.4 Rôle des biomarqueurs dans le diagnostic et la prise en charge clinique des maladies neurodégénératives.....	30
1.5 Analyse automatique du discours comme marqueur non invasif et peu cher de dépistage clinique .....	32
1.5.1 Littérature clinique existante sur les différences de discours entre MA et sujets témoins	35
1.5.2 Littérature clinique existante sur les différences de discours entre les variantes de l’APP	36
1.5.3 Approches computationnelles de l’analyse automatique du discours en MA et en APP	36

1.6 Objectifs .....	39
Chapitre 2 – Méthodologie et résultats .....	41
Article 1 : Connected speech features from picture description in Alzheimer’s disease : A systematic review .....	42
Article 2 : Connected speech markers of amyloid burden in primary progressive aphasia .....	92
Chapitre 3 – Discussion générale .....	126
3.1 Sommaire et interprétation des résultats .....	127
3.1.1 Objectifs et résultats de l’Article 1 .....	127
3.1.2 Limites de l’article 1 .....	129
3.1.3 Contributions de l’article 1 .....	132
3.1.4 Objectifs et résultats de l’article 2 .....	135
Efficacité comparée du discours par rapport au bilan cognitif/diagnostic clinique et aux biomarqueurs plasmatiques .....	138
Marqueurs discriminants de pathologie Alzheimer dans l’APP .....	139
Marqueurs non discriminants de pathologie Alzheimer dans l’APP .....	139
3.1.5 Limites de l’article 2 .....	141
3.1.6 Contributions de l’article 2 .....	142
3.2 Contributions scientifiques générales des deux articles .....	142
Avenues futures .....	145
Conclusion .....	147
Références bibliographiques .....	149

## Liste des tableaux

Tableau 1. –	Search terms and descriptors used for electronic database search .....	80
Tableau 2. –	Connected speech variables in the 44 reviewed articles.....	88
Tableau 3. –	Table 1: Average sociodemographic information and severity of impairment by group.	116
Tableau 4. –	Table 2: Averaged classification outcomes for leave-one-out cross-validation.	116
Tableau 5. –	Table 3: Group averages of statistically significant features.....	117
Tableau 6. –	Table A1: list of all extracted features.....	122
Tableau 7. –	Table A2: distribution of clinical diagnoses by amyloid status.....	123
Tableau 9. –	Table A3: Features selected in at least 10% of folds during cross-validation....	124



## Liste des figures

Figure 1. –	Flowchart of the review process.....	89
Figure 2. –	Summary of tested variables by domain of connected speech.....	90
Figure 3. –	Most commonly reported discriminant variables from AD picture description .....	91
Figure 4. –	Figure 1: Group distributions of statistically significant features. ....	118



## Liste des sigles et abréviations

ACP :	atrophie corticale postérieure
APP :	aphasie progressive primaire
DLFT :	dégénérescence lobaire frontotemporale
IA :	intelligence artificielle
LCR :	liquide céphalorachidien
MA :	maladie d'Alzheimer
PiB :	imagerie par émission de positron amyloïde au Pittsburgh compound B
TALN :	traitement automatique du langage naturel
TCL :	trouble cognitif léger
TEP :	tomographie par émission de positrons
vlAPP :	variante logopénique de l'aphasie progressive primaire
vsAPP :	variante sémantique de l'aphasie progressive primaire
vnfAPP :	variante non fluente de l'aphasie progressive primaire





*À Emma, Paul et Diane*



## Remerciements

Mes plus sincères remerciements vont à ma directrice de recherche la professeure Simona Maria Brambati. Pour son support indéfectible, son mentorat scientifique rigoureux et les opportunités extraordinaires offertes tout au long de mon parcours, je lui serai éternellement reconnaissant. Professeure Brambati m'a inculqué l'audace d'écrire et le courage de penser. Pour la supervision par une chercheuse infatigable promouvant de hauts standards scientifiques, mais incarnant toujours patience et humanité dans ses interventions, ma gratitude est difficile à décrire. Je dirai donc simplement : merci pour tout.

Je veux remercier toute la brigade du labo LINGUA, en particulier Christophe et Bérengère pour le code et les révisions patientes, Johémie pour les nombreux projets et son support, Georges pour les données psycholinguistiques, Maxime pour l'imagerie, les projets et les conseils. Merci aussi à Audrey, Jadziah et Sarah pour la collecte des données. Julien, Marianne et Mariem, ainsi que plus tard Melody et Tina, pour vos conseils et d'avoir fait du labo ce qu'il est.

Je désire remercier les professeurs Antonio Zadra et Sven Joubert dont les encouragements m'ont convaincu que j'avais ma place comme praticien-chercheur. Un merci particulier à professeure Isabelle Rouleau qui m'a accueilli généreusement et n'a jamais cessé de croire en nos projets. Merci à Gyslain Giguère du Département pour les conseils et à Ahmed Semine et l'équipe du Département pour le soutien administratif.

Ceux qui viennent de l'extérieur du labo : merci à Jessica qui m'a soutenu et encouragé à mes débuts, à Geneviève pour la contre-vérification patiente de mon spaghetti-code et à Jean-Frédéric pour les *crash course* nocturnes de *deep learning*. Je tiens également à remercier les membres du Groupe d'Intérêt en Neuropsychologie (GIN), en particulier Chanel, Olivier et Patricia pour les discussions et le soutien existentiel.

Mes remerciements vont également aux volontaires qui ont donné leur temps pour contribuer à nos recherches au CRIUGM, ainsi qu'aux patients et à leur famille dont la confiance accordée permet de faire avancer les connaissances sur les maladies qui les touchent.

Je remercie les organisations qui ont supporté mes travaux et mes voyages doctoraux : en ordre chronologique les IRSC, les FRQS, la FAÉCUM, le RQRV, le CRIUGM, le CRSNG.

Merci finalement aux personnes qui ont accepté de faire partie du comité de révision de la présente thèse, les professeures Nadia Gosselin, Éline De Guise et Vanessa Taler.

## Chapitre 1 – Contexte théorique

### 1. Introduction générale

#### 1.1 Manifestations cliniques de la maladie d'Alzheimer et troubles de langage associés

En 2016, la maladie d'Alzheimer (MA) touchait 546 000 Canadiens et coûtait annuellement 10,4 milliards de dollars. On prévoit que les maladies cognitives, dont la MA est de loin la forme la plus répandue, affecteront près d'un million de personnes dès 2031 (Canada, 2016). En l'absence d'un traitement efficace de la MA, d'importants efforts de recherche sont actuellement dédiés à la découverte de marqueurs précoces de la maladie et à l'amélioration des conditions de vie des personnes atteintes et de leurs proches.

La plupart des cas de MA se déclarent après 65 ans et leur prévalence augmente avec l'âge : 2-3% chez les 65-74 ans, 10-11% chez les 75-84 ans, et 29-37% chez les 85 ans et plus (Canada, 2016). Certains cas de démence se déclarent avant l'âge de 65 ans, auquel cas on parle de démence à début précoce. La MA demeure la cause plus fréquente de démence à début précoce dans une majorité d'études, bien que les formes de démence non Alzheimer (frontotemporale, à corps de Lewy, etc.) ont une plus haute prévalence parmi les démences à début précoce que parmi les démences à début tardif (Vieira et al., 2013).

En bref, la MA dite sporadique à début tardif compte pour environ 90 % des cas de MA, alors que les formes précoces de la maladie constituent 1 à 6 % des cas. Il existe des formes familiales génétiques liées à des mutations du gène précurseur de la protéine amyloïde, du gène de la préséniline-1 ou du gène de la préséniline-2. Les formes familiales de la MA représentent toutefois moins de 1% des cas (Piaceri et al., 2013). L'allèle 4 du gène de l'apolipoprotéine E (ApoE) constitue le plus grand facteur de susceptibilité à la MA à début tardif et demeure un facteur de

risque significatif pour le développement d'une MA à début précoce (Liu et al., 2013). Néanmoins, environ 60% des cas de MA sporadique ne sont pas liés à l'ApoE, ce qui suggère un rôle de d'autres déterminants génétiques et de facteurs environnementaux dans l'apparition et l'évolution de la plupart des cas de MA.

La plupart des patients avec une MA présentent la forme amnésique de la maladie, dans laquelle ils éprouvent une difficulté marquée à apprendre et à rappeler les informations apprises récemment. Ce syndrome amnésique progressif et prédominant lié à l'atrophie du cortex entorhinal et des hippocampes est la présentation initiale de plus de 85% des cas de MA (Dubois et Michon, 2015). Au-delà de ces troubles de mémoire caractéristiques, pratiquement chaque cas de MA s'accompagne de difficultés de langage expressif à un moment ou à un autre de son évolution.

Les troubles du langage sont repérés très fréquemment par les proches aidants de la personne atteinte : 85% d'entre eux ont noté la présence d'un manque du mot, 83% ont remarqué une difficulté à nommer les objets, 81% ont vu une difficulté à écrire une lettre et 76% ont constaté une baisse dans la compréhension des instructions (Bayles et Tomoeda, 1991). En fait, les perturbations du langage et les bris de communication associés sont parmi les symptômes les plus dévastateurs pour le patient et ses proches-aidants et les plus coûteux pour le système de santé (Murray et al., 1999; Orange et Colton-Hudson, 1998). De plus, des changements langagiers subtils sont souvent détectables très tôt dans l'évolution de la maladie (Taler et Phillips, 2008), possiblement une décennie avant le diagnostic (Amieva et al., 2005). Néanmoins, la prévalence et le décours temporel des troubles du langage dans l'évolution d'un cas de MA « typique » demeure difficile à estimer avec précision, notamment en raison de l'hétérogénéité des déficits de langage. Néanmoins, comme nous l'avons indiqué précédemment, certains ont soutenu que la majorité, sinon *tous* les patients avec une démence de type Alzheimer présentent une perturbation des fonctions langagières à un

stade ou à un autre de la maladie (Murdoch et al., 1987; Smith et al., 1987). De manière intéressante, la patiente originale d'Aloïs Alzheimer présentait probablement une variante atypique de la MA avec un trouble du langage à l'avant-plan (Möller et Graeber, 1998).

Un nombre croissant d'études se penchent sur la détection clinique de troubles langagiers au stade du trouble cognitif léger (TCL), un état jugé intermédiaire entre la santé et la démence où la personne manifeste des troubles cognitifs, mais alors que ceux-ci ne causent pas de déficits fonctionnels dans les activités de la vie quotidienne (Petersen et al., 2009). Deux tiers des patients qui ont par la suite développé une MA confirmée par autopsie présentaient des anomalies hétérogènes du langage expressif au stade du TCL (Ahmed, Haigh, et al., 2013). Cependant, seulement 70% des personnes rencontrant les critères du TCL présentent une neuropathologie de type Alzheimer (Dubois et Michon, 2015), ce qui souligne l'importance de bien caractériser les cohortes de patients pour étudier d'éventuels marqueurs langagiers de la MA au stade du TCL.

La perturbation du langage qui apparaît le plus fréquemment dans la MA est le manque du mot, lors duquel la personne ne peut accéder à la représentation lexicale (le mot) en temps voulu (Balthazar et al., 2008). Ces troubles sont typiquement mis en évidence à l'examen clinique par une tâche de dénomination sous confrontation telle que le Boston Naming Test (BNT) dans lequel la personne doit nommer des images d'objets ou d'animaux qui lui sont présentées. Une faible performance au BNT prédit un déclin plus rapide des fonctions cognitives globales (Boller et al., 1991), un phénomène qui a aussi été relié à une dégradation du système sémantique (Mickes et al., 2007).

L'échec en dénomination peut correspondre dans certains cas à un trouble d'accès lexical - comme le phénomène du mot « sur le bout de la langue » où la personne sait ce dont il s'agit mais ne peut accéder à sa représentation lexicale - ou dans d'autres cas à cause d'un trouble sémantique en vertu

duquel la personne n'arrive pas à identifier avec précision ce qu'on lui présente en raison d'une dégradation de ses connaissances sémantiques. Dans d'autres cas, des erreurs perceptives peuvent également découler de déficits visuoperceptuels. Le mécanisme cognitif sous-jacent (accès lexical, trouble sémantique, trouble visuoperceptuel) peut être mis en lumière par une classification des erreurs, une démarche d'analyse qualitative dont l'importance est bien établie en neuropsychologie clinique (Hodges et al., 1991; Mitrushina et al., 2005; Nicholas et al., 1996; Slegers, Cole, et al., 2018).

Dans plusieurs cas de MA, les troubles d'accès lexical et les troubles sémantiques coexistent dans des proportions variables pour expliquer les difficultés marquées en dénomination d'image qui accompagnent la progression de la maladie (Montembeault et al., 2017). Bien que ce débat sur la part relative des deux déficits ne soit pas encore tranché définitivement, une épreuve de dénomination comme le BNT démontre une bonne sensibilité et spécificité (sensibilité : 99%, spécificité : 74%) pour identifier les personnes atteintes d'une MA (Belleville et al., 2017). Ce test est également utile pour identifier les personnes avec un TCL qui développeront plus tard une MA (Eckerström et al., 2013). L'association entre les scores au BNT et les perturbations du discours quotidien ont reçu des confirmations récentes (Fraser, Lundholm Fors, et al., 2019; Kavé et Dassa, 2018; Kavé et Goral, 2016). En somme, les changements dans le langage expressif, qui peuvent être détectés à l'aide de tâches assez sensibles, revêtent un intérêt particulier pour le dépistage précoce de la MA dans des visées de recherche et de diagnostic clinique.

## **1.2 Pathophysiologie et biomarqueurs dans les classifications récentes de la maladie d'Alzheimer**

La pathophysiologie de la MA, c'est-à-dire le phénomène biologique à l'origine des symptômes, consiste en la présence concomitante de dégénérescences neurofibrillaires associées à



l'hyperphosphorylation de la protéine tau et d'accumulation extracellulaire de la protéine bêta-amyloïde (A $\beta$ ). Ce sont ces deux processus qui causent des pertes neuronales et l'atrophie cérébrale observée dans la MA (Braak et Braak, 1991). Alors que jusqu'à récemment, elle ne pouvait être établie que *post mortem*, l'accumulation de protéine A $\beta$  peut désormais être détectée *in vivo* par imagerie amyloïde : la tomographie par émission de positron (TEP). Cette technique consiste à injecter un traceur radioactif comme le Pittsburgh compound B (PiB) qui, en se liant préférentiellement aux plaques amyloïdes, permet d'estimer la présence de ces dernières via un indice de fixation normalisé du traceur.

La confirmation par les biomarqueurs est importante puisque d'une part, ce ne sont pas tous les patients avec un syndrome amnésique qui présentent le phénomène pathophysiologique de la MA et que d'autre part, ce ne sont pas toutes les personnes chez qui on détecte une pathophysiologie MA qui sont symptomatiques. La disponibilité d'un examen des biomarqueurs permettrait de supporter le diagnostic différentiel d'autres affections du SNC causant un syndrome amnésique progressif, mais qui ne serait pas dû à l'installation d'une pathologie MA, par exemple une maladie des grains argyrophiles, une sclérose hippocampique ou un syndrome de Wernicke-Korsakoff. Les patients atteints de ces affections, bien qu'ils éprouvent des troubles de mémoire dont l'ampleur peut être similaire à une MA, n'en présentent généralement pas la pathophysiologie. À l'inverse, jusqu'à 80% des patients avec une démence à corps de Lewy présentent une TEP amyloïde positive (Dubois et Michon, 2015), sans nécessairement que des troubles de la mémoire épisodique soient à l'avant-plan du tableau clinique, qui est dans leur cas le plus souvent composé d'hallucinations, de troubles de l'humeur et de problèmes visuoperceptifs.

Selon une revue systématique récente, la sensibilité et la spécificité de la TEP amyloïde atteignent respectivement 91% (79 à 98%) et 92% (76 à 100%) pour prédire le diagnostic histopathologique

post mortem de MA (Fink et al., 2020). Néanmoins, on estime que 20 à 30% des personnes âgées asymptomatiques obtiennent un résultat d'imagerie amyloïde positif (Rodrigue et al., 2012), ce qui met en doute la spécificité de la charge amyloïde mesurée par TEP pour prédire le déclin cognitif. Les études d'autopsie ont aussi montré qu'un fardeau amyloïde significatif peut être détecté chez 25-35% des adultes âgés en santé et dans 50-60% des cas de TCL (Bahar-Fuchs et al., 2013), ce qui est concordant avec les proportions estimées par TEP. Un fardeau amyloïde estimé par TEP est associé à de plus faibles performances aux tâches cognitives (Villemagne et al., 2018), en particulier aux tâches de mémoire épisodique. En dépit des limites du fardeau amyloïde comme cause directe des troubles cognitifs, il est vraisemblable de croire que les individus décédés « asymptomatiques » avec un fardeau amyloïde avéré à l'autopsie auraient éventuellement développé des troubles cognitifs s'ils avaient vécu plus longtemps (Wolk et al., 2009).

En 2011, les critères de recherche du National Institute on Aging pour la maladie d'Alzheimer ont été mis à jour pour refléter les nouvelles connaissances sur la MA et d'autres maladies neurodégénératives, dans le but explicite d'augmenter la spécificité du diagnostic de MA (McKhann et al., 2011). Ces critères révisés ont été rédigés afin de tenir compte de plusieurs nouveautés dans la recherche et la prise de décision clinique, dont quatre aspects principaux en particulier.

Premièrement, la présence de pathologie Alzheimer peut être détectée dans un large spectre de populations cliniques (personnes en santé ou avec une démence). Ainsi, la démence de type Alzheimer est comprise comme la manifestation clinique causée par des changements pathophysiologiques spécifiques qui peuvent avoir précédé de plusieurs années le diagnostic clinique. Deuxièmement, des maladies apparentées dont la pathophysiologie est différente de la MA ont été caractérisées relativement récemment, comme la démence frontotemporale et certaines

aphasies primaires progressives (APP). Troisièmement, l'accessibilité croissante des biomarqueurs non seulement à des fins de recherche mais également dans l'investigation clinique. Quatrièmement, le constat que la mémoire n'est pas toujours le symptôme dominant de la MA et qu'il existe des formes de MA où sont atteints d'abord les processus visuospatiaux (atrophie corticale postérieure, ACP) ou langagiers (variante logopénique de l'aphasie primaire progressive, vlAPP). En outre, la révision des critères pour la MA comprend désormais un niveau de *probabilité*, c'est-à-dire que les critères stipulent que la présence des biomarqueurs positifs augmente la probabilité que la présentation clinique d'un cas de MA « probable » ou « possible » soit effectivement due à une pathologie Alzheimer sous-jacente et non à une autre maladie.

### **1.3 Pathophysiologie et biomarqueurs dans les classifications récentes des aphasies progressives primaires**

Reflétant les deuxième et quatrième points des critères de McKhann et al. (2011) relevés plus haut, la dernière itération des critères de classification des APP réserve elle aussi une place importante aux biomarqueurs (Gorno-Tempini et al., 2011). Les APP sont un ensemble syndromes cliniques caractérisés par l'apparition progressive et la prédominance de déficits langagiers en l'absence d'autres symptômes neurocognitifs (Mesulam, 1982, 1987; Mesulam, 2001). Ces déficits touchent la production verbale, la dénomination d'objets, la syntaxe ou la compréhension de mots et sont observables en conversation ou à l'occasion d'une évaluation clinique du langage. Les APP sont un phénomène comparativement plus rare que la MA et qui fait partie des « troubles neurocognitifs à début précoce », un groupe de maladies qui composent entre 2 et 8% des cas de troubles neurocognitifs (Canada, 2016).

La dernière publication des critères diagnostiques des APP (Gorno-Tempini et al., 2011) distingue trois variantes : non fluente/agrammatique (vnfAPP), sémantique (vsAPP) et logopénique (vlAPP).

Sur le plan nosographique, la dernière variante à être identifiée est la variante logopénique (Gorno-Tempini et al., 2008).

La vnfAPP est caractérisée par un discours laborieux et un agrammatisme (Montembeault et al., 2018). Le débit verbal est ralenti par une apraxie du discours, c'est-à-dire un déficit de planification moteur. Des distorsions, suppressions et substitutions de phonèmes peuvent résulter d'un trouble de la sélection phonémique ou d'un trouble moteur. La prosodie est souvent affectée et une dysarthrie peut accompagner l'apraxie du discours. Les patients avec une vnfAPP utilisent proportionnellement un plus petit nombre de verbes dans la construction de leurs phrases. Ces manifestations cliniques correspondent à l'atrophie cérébrale de la région du gyrus frontal inférieur. Les critères diagnostiques de la vnfAPP requièrent l'un ou l'autre de deux éléments centraux entre 1) agrammatisme du langage expressif et 2) discours laborieux/hésitant comportant des paraphasies phonémiques/distorsions irrégulières ainsi qu'au moins deux des trois caractéristiques suivantes : 1) déficit de compréhension des phrases complexes ; 2) compréhension intacte des mots isolés ; 3) connaissance préservée des objets.

La vsAPP est caractérisée par une dégradation progressive et multimodale des connaissances sémantiques se manifestant à l'examen par une anomie sévère et un déficit marqué de compréhension de mots isolés (Montembeault et al., 2018). Des perturbations comportementales apparaissent aussi souvent dans les stades intermédiaires et avancés de la vsAPP, typiquement la désinhibition, la rigidité mentale et les compulsions. Le langage expressif des patients avec une vsAPP est généralement grammaticalement correct mais peu informatif. Les paraphasies sémantiques et le recours à la catégorie superordonnée sont fréquentes dans les tâches de dénomination. Les mots rares ou non prototypiques sont remplacés par des termes plus fréquents. La lésion cérébrale caractéristique est une atrophie bilatérale, plus souvent prédominante à gauche,

des lobes temporaux antérieurs. Les deux caractéristiques essentielles au diagnostic sont la présence d'une anomie (déficit de dénomination sous confrontation) et d'un déficit de la compréhension des mots isolés. Trois des quatre caractéristiques suivantes sont aussi nécessaires : 1) agnosie des objets ; 2) dyslexie de surface ou dysgraphie ; 3) préservation de la répétition ; 4) grammaire et aspects moteurs du langage expressif préservés.

La vlAPP est caractérisée par un trouble d'accès lexical et des déficits en répétition de phrases (et éventuellement en compréhension de phrases complexes). Le mécanisme cognitif sous-jacent suspecté est une défaillance de la mémoire auditivoverbale à court-terme (Montembeault et al., 2018). La répétition et la compréhension de mots isolés est préservée. Le déclin cognitif dans la vlAPP serait plus global (touchant progressivement la mémoire, les habiletés visuospatiales et l'attention) et rapide que dans les deux autres variantes de l'APP. L'atrophie cérébrale associée est souvent située près de la jonction temporopariétale. Les critères diagnostiques de la vlAPP requièrent les deux caractéristiques centrales de trouble d'accès lexical apparent dans le discours et de trouble de répétition de phrases complexes. Trois des quatre caractéristiques suivantes doivent également être présentes : 1) paraphasies phonologiques en discours et dénomination d'image ; 2) compréhension de mots isolés et connaissance des objets préservées ; 3) préservation des aspects moteurs du langage ; 4) absence d'agrammatisme franc.

Chacune des trois variantes présente une pathophysiologie cérébrale sous-jacente plus typique. Dans le cas de la vnfAPP, on observe le plus souvent une tauopathie (FTD-4R tau) alors que la vsAPP comporte habituellement une pathologie de la protéine TDP-43. Dans le cas de la vlAPP, une pathologie amyloïde similaire à celle observée dans la MA amnésique est plus souvent constatée à l'autopsie (Montembeault et al., 2018; Ruksenaite et al., 2021). Cependant, la correspondance entre le phénotype clinique (variante de l'APP) et la pathologie sous-jacente

(Alzheimer ou autre) n'est pas systématique. Ainsi, jusqu'à 55% de tous les patients avec une APP présentent un dépôt anormal de protéine bêta-amyloïde (Villarejo-Galende et al., 2017). Lorsqu'on isole les cas de vAPP, ce taux grimpe à 86%, variant entre 57 et 100% selon le centre d'expertise duquel proviennent les données (Bergeron et al., 2018). Suivant la nécessité de développer des traitements agissant au niveau pathophysiologique plutôt que phénotypique, les plus récentes conceptualisations des APP, à l'instar des critères révisés de la MA, sont elles aussi articulées autour des biomarqueurs (Ruksenaite et al., 2021).

#### **1.4 Rôle des biomarqueurs dans le diagnostic et la prise en charge clinique des maladies neurodégénératives**

Une autre révision récente des critères de diagnostic pour la maladie d'Alzheimer (MA) propose une simplification des critères de McKhann et al. (2011). Cette révision se résume à deux étapes du processus diagnostic, à savoir l'établissement d'abord d'un phénotype clinique et ensuite d'un biomarqueur pathophysiologique (Dubois et al., 2014; Morris et al., 2014). Le phénotype clinique peut correspondre à la variante amnésique typique de la MA (anciennement démence de type Alzheimer) ou aux variantes atypiques comme la variante frontale, la variante langagière et la variante visuelle.

L'identification de la pathologie sous-jacente par l'intermédiaire des biomarqueurs influence le traitement et la prise en charge des patients et constitue donc un enjeu important (Bensaïdane et al., 2016; Rabinovici et al., 2008; Rogalski et al., 2016; Tippett, 2020). De nombreux patients chez qui on a confirmé post mortem une pathologie associée à la dégénérescence lobaire frontotemporale (DLFT) avaient reçu erronément un diagnostic de MA durant la vie. De façon analogue, 10-40% de patients avec une DLFT diagnostiquée présentaient post mortem une histopathologie MA (Rabinovici et al., 2011). Dans le contexte d'émergence de traitements ciblant

la pathophysiologie, il importe d'identifier correctement les patients qui seraient susceptibles de bénéficier de ce type de traitement, y compris les personnes présentant une variante atypique de la MA. Dans l'histoire récente, l'approbation de l'aducanumab, un anticorps monoclonal par la Federal Drug Administration (FDA) a suscité la controverse (Karlawish et Grill, 2021) dans un revirement de situation après résultats initiaux décevants en phase III. En dépit des doutes et incertitudes quant à leur efficacité, on anticipe que le développement d'essais cliniques pour d'autres molécules agissant au niveau de la pathologie amyloïde ou tau constitueront un champ de recherche actif dans les années à venir (Gandy et al., 2021).

Au Canada, les biomarqueurs reconnus comme les plus spécifiques de la pathologie Alzheimer sont ceux obtenus par ponction lombaire (liquide céphalorachidien, LCR) et par TEP amyloïde. Cependant, plusieurs patients se présentant avec une DLFT diagnostiquée sur la base de critères cliniques présentent une imagerie amyloïde positive (Tan et al., 2017), près du tiers lors selon l'une des premières études sur le sujet (Rabinovici et al., 2007). L'histopathologie post mortem demeure le seul indicateur parfaitement fiable de la présence pathologie Alzheimer et la sensibilité et la spécificité de la TEP amyloïde sont estimées à respectivement 91 et 92%.

Les biomarqueurs sont en phase de devenir des composantes de plus en plus centrales au diagnostic des maladies neurodégénératives dont font partie la MA et les APP (Laforce et al., 2016; Witte et al., 2015). Toutefois, il persiste des barrières liées à la disponibilité des infrastructures et des centres d'expertise. Les lignes de pratique actuelles stipulent que l'imagerie amyloïde par TEP ne devrait pas être considéré comme une procédure de routine et est réservée aux patients présentant un diagnostic incertain *après* une évaluation experte *et seulement* lorsque le résultat de l'imagerie est susceptible de modifier la prise de décision clinique (Laforce et al., 2016). À ce sujet, la conséquence la plus fréquente de l'obtention de résultats d'imagerie amyloïde est un changement

de diagnostic, bien que les effets sur la qualité de vie du patient doivent encore être étudiés systématiquement (Rabinovici et al., 2019).

### **1.5 Analyse automatique du discours comme marqueur non invasif et peu cher de dépistage clinique**

Comme nous venons de le décrire, les procédures actuelles pour recueillir les biomarqueurs ont le désavantage d'être coûteuses et invasives, ainsi que de requérir une visite du patient dans un centre hospitalier muni d'infrastructures spécialisées. En raison de ces limites, un vaste et effervescent champ de recherche concerne la découverte de biomarqueurs non invasifs, peu chers (Villemagne et al., 2018) et accessibles en télémédecine (Goldsack et al., 2020). Parmi les marqueurs candidats, l'analyse de courts échantillons de discours fait l'objet d'un nombre croissant d'études (Battista et al., 2020; Clarke et al., 2020; Myszczyńska et al., 2020). Effectivement, les changements langagiers ont l'avantage de survenir tôt dans l'évolution des APP et de la MA et d'être détectables à partir de tâches cliniques simples.

Le discours continu (ou discours, *connected speech*) se définit comme « une production langagière utilisée de façon continue, comme dans une situation de communication naturelle » c'est-à-dire en phrases complètes telles qu'elles sont produites dans la vie courante (Crystal, 2011). Une analyse dite « multidimensionnelle » d'un échantillon de discours peut s'effectuer aux niveaux phonologique/articulatoire (prononciation, débit), lexical (les mots employés), morphosyntaxique (l'articulation de l'énoncé, p.ex. les clauses subordonnées et la longueur des phrases), de la fluidité et de ses bris (pauses de recherche de mot, phrases avortées, répétitions), de la sémantique (éléments de contenu significatif) et de la pragmatique (intention et efficacité communicative, par exemple). Le discours contient donc plusieurs variables qui pourraient constituer des marqueurs d'un processus de dégénération lorsque l'une ou l'autre de ses composantes sont modifiées par la



maladie. Ainsi, l'analyse computationnelle de ces échantillons de discours dans une perspective d'automatisation pourrait servir de dépistage par des marqueurs précoces, peu chers et non invasifs de la présence d'un processus neurodégénératif spécifique, à savoir la pathologie Alzheimer.

La promesse de techniques automatisées repose notamment sur des approches d'apprentissage statistique (*machine learning*). Celles-ci consistent, pour un ordinateur, à apprendre « à partir de données constituant [son] expérience et de généraliser les connaissances acquises à de nouveaux paramètres, au lieu de s'appuyer sur des règles explicitement programmées » (Canada, 2021). Dans le cas d'échantillons de discours, le traitement automatique du langage naturel (TALN) concerne la traduction « entre les langages informatiques et les langues humaines. Il s'agit d'une méthode visant à permettre à un ordinateur de lire et de comprendre une ligne de texte sans lui fournir un indice ou un calcul quelconque. Autrement dit, le TALN automatise le processus de traduction entre les ordinateurs et les humains », traduction qui peut ensuite être traitée par un algorithme d'apprentissage automatique.

En apprentissage automatique, il existe des approches dites supervisées et des approches dites non supervisées. Les approches non supervisées sont utilisées pour déterminer des regroupements de données (par exemple des groupes de patients) sans savoir à l'avance ce que ces regroupements seront. À l'inverse, la technique de l'apprentissage supervisé (par un expert) permet à un algorithme d'apprentissage de ranger les intrants dans des classes spécifiques connues (Canada, 2021). Les variables indépendantes (caractéristiques, ou prédicteurs) du modèle sont sélectionnées et combinées pour prédire une variable dépendante (*output*, ou étiquette) numérique/catégorique, comme il est possible de le faire par exemple par régression linéaire/logistique (Hastie et al., 2009). Pour ce faire, le jeu de données est divisé en trois ensembles. D'abord, un ensemble d'entraînement sur lequel l'algorithme « apprend » et un ensemble de validation sur lequel il peut tester ses

prédictions de manière itérative. Ensuite un ensemble de test, qui n'a jamais été « vu » par l'algorithme, est utilisé pour estimer la précision du modèle et sa plus ou moins bonne capacité à généraliser des prédictions à des données « nouvelles » à partir des caractéristiques pertinentes (*features* ou prédicteurs) retenues au cours de la phase d'entraînement sur les ensembles d'entraînement et de validation. Plus concrètement, dans le cas des échantillons de discours, il s'agirait pour l'algorithme d'apprentissage de classer correctement des échantillons, à partir de leurs caractéristiques (par exemple leur longueur, le vocabulaire employé, etc.) comme ayant été produit ou bien par une personne présentant une pathologie MA ou bien par une personne en santé/ne présentant pas de pathologie MA.

En amont de cette procédure de classification automatisée, il est aussi possible de maximiser l'extraction de caractéristiques de l'échantillon (de discours) par différentes techniques dites d'intelligence artificielle, qui se définit comme le « domaine de l'informatique consacré à la résolution de problèmes cognitifs communément associés à l'intelligence humaine, comme l'apprentissage, la résolution de problèmes, la perception visuelle et la reconnaissance de la parole et des formes » (Canada, 2021). Cette extraction de caractéristiques peut concerner par exemple les propriétés sémantiques des mots utilisés par la personne via l'apprentissage profond, un « sous-ensemble de l'apprentissage automatique qui imite le fonctionnement du cerveau humain en matière de traitement de données [...] En règle générale, il s'agit d'un algorithme à niveaux multiples qui identifie graduellement les [mots] à des niveaux plus élevés d'abstraction » (Canada, 2021). Ce type d'algorithme requiert des données massives dont le traitement est devenu possible grâce à des innovations relativement récentes.

Dans la littérature actuelle, l'analyse computationnelle du discours et l'apprentissage automatique a permis de différencier les patients avec une MA amnésique de témoins en santé (Fraser, Linz, et

al., 2019; Fraser et al., 2016), ainsi que de distinguer les trois variantes cliniques de l'APP (Fraser et al., 2014; Themistocleous, Webster, et al., 2020; Zimmerer et al., 2020). Cependant, l'importance relative des différentes caractéristiques du discours n'est pas fermement établie pour guider la prise de décision clinique. Par ailleurs, aucune étude n'a fait à notre connaissance le lien direct entre un large éventail de caractéristiques du discours et la pathophysiologie sous-jacente (Alzheimer vs non-Alzheimer).

#### 1.5.1 Littérature clinique existante sur les différences de discours entre MA et sujets témoins

Une riche littérature clinique témoigne de différences de groupe entre des patients avec une MA typique et des témoins dans plusieurs dimensions du discours. Les premières études expérimentales sur le sujet ont relevé notamment des changements dans le débit de la voix (Cherney et Canter, 1993; Illes, 1989) et le contenu lexical (Ehrlich et al., 1997; Hier et al., 1985), une simplification de la syntaxe (Croisile et al., 1996; Cummings et al., 1985; Ska et Duong, 2005), des bris dans la fluidité du discours (Ehrlich et al., 1997; Hier et al., 1985; Nicholas et al., 1985), un appauvrissement du contenu sémantique (Giles et al., 1996; Groves-Wright et al., 2004) et une diminution de l'efficacité du discours (Shimada et al., 1998; Smith et al., 1989). Ces études ont été suivies de nombreux travaux ayant donné des résultats parfois contradictoires et globalement difficiles à synthétiser. De plus, elles reposent sur des critères cliniques datant de bientôt quarante ans, en conséquence de quoi elles sont basées sur le phénotype clinique correspondant à la MA amnésique classique, qui, sauf rares exceptions (Ahmed, de Jager, et al., 2013), ne comprennent pas l'apport des biomarqueurs pour supporter le diagnostic. Finalement, l'abondance et l'hétérogénéité des études publiées créent le besoin d'une approche quantitative des preuves

existantes concernant les différentes dimensions du discours les plus discriminantes pour identifier les signes langagiers de la MA.

### 1.5.2 Littérature clinique existante sur les différences de discours entre les variantes de l'APP

Les trois variantes cliniques de l'APP ont été décrites plus récemment que la MA. Chacune est caractérisée par un profil neuropsychologique (Kamath et al., 2020) et langagier (Haley et al., 2021) particulier. On croit en outre que ces profils laissent des traces distinctives dans de courts échantillons de discours, conformément aux corrélats neuroanatomiques de chaque variante de l'APP (Cho et al., 2021). En résumé, les patients avec un vnfAPP forment des phrases courtes, simples et souvent agrammaticales, avec un débit lent et une haute proportion de noms, parfois avec des distorsions (Ash et al., 2013; Wilson et al., 2010). Les patients avec une vsAPP, à l'inverse, ont un débit relativement normal, forment des phrases complètes et complexes, mais produisent une faible proportion de noms et de contenu informatif en utilisant des mots imprécis et plus fréquents (Ash et al., 2013; Boschi et al., 2017; Wilson et al., 2010). Les patients avec une vlAPP produisent des anomalies syntaxiques relatives à leurs difficultés d'accès lexical et de mémoire de travail, tels que des répétitions et des révisions de phrases, produisent plus de pauses remplies (« euh ») et utilisent proportionnellement plus de verbes (Ash et al., 2013; Boschi et al., 2017; Wilson et al., 2010). Comme dans le cas des études sur les anomalies de discours dans la MA, ces études sont centrées sur le phénotype clinique et non sur les biomarqueurs.

### 1.5.3 Approches computationnelles de l'analyse automatique du discours en MA et en APP

La linguistique computationnelle et l'apprentissage machine ont initialement été appliquées à l'étude du discours dans les maladies neurodégénératives par le biais d'expériences de

classification comme celles de Jarrold et al. (2010) ou de Roark et al. (2011) avec des patients atteints d'un trouble cognitif léger (TCL). Depuis, des travaux subséquents dans le domaine de la MA ont montré qu'il est possible de distinguer avec une précision significative les patients avec une MA de témoins en santé (Hernández-Domínguez et al., 2018; Zhou et al., 2016). D'autres travaux se penchent également, à partir d'échantillons de discours, sur la détection précoce et sur la prédiction de la conversion du TCL vers la MA (Sanborn et al., 2020; Themistocleous, Eckerström, et al., 2020). D'autres approches semi-automatisées ont servi la classification des APP (Cho et al., 2021; Themistocleous, Webster, et al., 2020) suivant les premiers efforts de Peintner et al. (2008). En raison de l'engouement actuel pour l'intelligence artificielle (IA) et l'apprentissage automatique, on assiste à la prolifération d'études - de qualité variable - appliquant différentes déclinaisons d'un algorithme à un jeu de données partagé pour obtenir la meilleure performance de classification (p. ex. Land Jr et Schaffer (2019); Mittal et al. (2020); Padhee et al. (2020); Zhu et Liang (2020)). Fait important, les travaux les plus récents issus de groupes de développeurs logiciel confirment la pertinence d'une approche guidée par l'expertise clinique comparativement à une approche basée uniquement sur l'IA (Balagopalan et al., 2020).

En résumé, la recherche existante a démontré qu'un court échantillon de discours peut différencier un patient qui aurait une MA typique d'un témoin en santé ou encore déterminer la variante clinique de l'APP dont il est atteint. L'expertise clinique rassemblée dans cette littérature (sections 1.4.1 et 1.4.2) suggère que les approches les plus efficaces sont multidimensionnelles, c'est-à-dire qu'elles considèrent plusieurs domaines langagiers dont les propriétés acoustiques, la fluidité, le lexique, la syntaxe, la sémantique et la pragmatique (Boschi et al., 2017; Filiou et al., 2019; Mueller et al., 2018). La tâche clinique la plus courante qui permet de faire émerger ces multiples dimensions du discours est la description d'image comme le vol du biscuit de la *Boston Diagnostic Aphasia*

*Examination* (Goodglass et al., 2000) ou la scène du pique-nique de la *Western Aphasia Battery* (Kertesz, 1982). Dans ce type de tâche, le patient doit décrire en phrases complètes une scène visuelle complexe. On peut procéder ensuite à l'analyse du signal acoustique ou des transcriptions afin d'extraire des caractéristiques indicatrices d'un processus pathologique. Pour paver la voie à l'automatisation de l'analyse du discours dans ce but d'identification, il serait d'abord utile de compiler, dans une approche quantitative systématique, les meilleurs marqueurs de description d'image permettant d'identifier les personnes présentant une MA typique, la forme de démence qui représente plus de 70% des cas documentés. Ensuite, en raison de l'existence et du chevauchement des phénotypes cliniques (présentations amnésiques vs langagières, notamment) et de la concordance imparfaite entre le phénotype clinique et la pathologie sous-jacente, il faudrait vérifier si le discours permet également de déceler spécifiquement la pathologie Alzheimer dans les cas de maladies neurodégénératives moins typiques comme les APP, où les troubles du langage sont à l'avant-plan. Il serait donc utile de valider les marqueurs du discours identifiés comme les plus typiques de la MA lorsqu'on est face à un groupe hétérogène de patients avec une APP.

En résumé, nous identifions donc deux lacunes dans la littérature actuelle que nous positionnons dans le contexte de l'émergence de techniques computationnelles d'analyse du discours. Premièrement, un survol de la littérature clinique existante au sujet des variables de description d'image révèle qu'elle est abondante et hétérogène. Nous ne disposons donc pas d'indicateurs quantitatifs et exhaustifs qui permettraient de recenser, classer et synthétiser les variables d'intérêt afin d'obtenir un portrait global de la littérature. De plus, il nous semble utile de guider et de concentrer les efforts de recherche sur des variables cliniquement pertinentes dans un contexte où la recherche transdisciplinaire est menée par un nombre croissant de non-spécialistes de la MA et de non-cliniciens. Deuxièmement, considérant le chevauchement des présentations cliniques de la

MA mentionné plus haut, ainsi que l'existence de variantes atypiques de la MA qui se présentent avec une symptomatologie similaire (trouble de langage) mais une pathologie différente (pour chaque variante de l'APP), il faudrait vérifier la capacité d'une approche computationnelle du discours à guider l'obtention des biomarqueurs. À cet effet, nous considérons que dans un premier temps, une revue de littérature du discours de description d'image dans la MA - telle qu'elle était définie dans la littérature neurologique des 40 années précédant les critères de 2011 - permettrait d'extraire un large éventail de variables candidates et d'évaluer quantitativement le potentiel discriminant de chacune. Ensuite, cet ensemble de variables pourrait être transféré et bonifié par une littérature spécifique aux APP dans le but d'identifier les variables les plus pertinentes pour repérer un cas d'APP qui présente un résultat positif à l'imagerie amyloïde.

## **1.6 Objectifs**

La présente thèse comporte donc deux objectifs constitués chacun d'un article. Notre premier objectif consiste à mener une revue systématique de la littérature portant sur les caractéristiques du discours de description d'image qui distinguent le plus fréquemment les patients avec le phénotype clinique typique de la démence de type Alzheimer de témoins en santé. Notre second objectif consiste à déterminer les meilleurs marqueurs de la présence d'une pathologie Alzheimer validée par biomarqueur dans un groupe hétérogène de patients avec une APP. En utilisant un large éventail de caractéristiques du discours extraites automatiquement, nous élaborons une expérience de classification de ces patients qui permet de prédire la pathologie sous-jacente telle qu'objectivée par l'agrégation de protéine A $\beta$  mesurée par TEP-PiB.





## **Chapitre 2 – Méthodologie et résultats**

**Article 1 : Connected speech features from picture description in Alzheimer’s disease : A systematic review**

Antoine Slegers<sup>a,b</sup>, Renée-Pier Filiou<sup>a,b</sup>, Maxime Montembeault<sup>a,b</sup>, Simona Maria Brambati<sup>a,b</sup>

<sup>a</sup> Département de Psychologie, Université de Montréal, Montréal, QC, Canada

<sup>b</sup> Centre de Recherche de l’Institut Universitaire de Gériatrie de Montréal (CRIUGM), CIUSSS du Centre-Sud-de-l’Île-de-Montréal, Montréal, QC, Canada

Accepted 2 July 2018

**Handling Associate Editor:** Raffaella Migliaccio

**Correspondence to:** Simona Maria Brambati, Centre de Recherche de l’Institut Universitaire de Gériatrie de Montréal (CRIUGM), CIUSSS du Centre-Sud-de-l’Île-de-Montréal, 4565 Queen-Mary, Montréal, QC, H3W 1W5, Canada. E-mail: [simona.maria.brambati@umontreal.ca](mailto:simona.maria.brambati@umontreal.ca)

**Référence complète**

Slegers, A., Filiou, R. P., Montembeault, M., & Brambati, S. M. (2018). Connected speech features from picture description in Alzheimer’s disease: A systematic review. *Journal of Alzheimer's disease*, 65(2), 519-542.

## **Abstract**

The language changes that occur over the course of Alzheimer's disease (AD) can impact communication abilities and have profound functional consequences. Picture description tasks can be used to approximate everyday communication abilities of AD patients. As various methods and variables have been studied over the years, current knowledge about the most affected features of AD discourse in the context of picture descriptions is difficult to summarize. This systematic review aims to provide researchers with an overview of the most common areas of impairment in AD discourse as they appear in picture description tasks. Based on the 44 articles fulfilling inclusion criteria, our findings reflect a multidimensional pattern of changes in the production (speech rate), syntactic (length of utterance), lexical (word-frequency and use of pronouns), fluency (repetitions and word-finding difficulties), semantic (information units), and discourse (efficiency) domains. We discuss our findings in the light of current research and point to potential scientific and clinical uses of picture description tasks in the context of AD.

**Keywords:** Alzheimer's disease, language tests, psycholinguistics, systematic review

## INTRODUCTION

The most commonly diagnosed form of dementia is Alzheimer's disease (AD). In the majority of cases, AD patients present with an amnesic syndrome, in which learning and recall of recently learned information are impaired. AD patients also develop nonamnesic features such as deficits in language, visuospatial abilities, and executive functions [1]. Language is impacted at some level in most cases of AD, especially language production [2]. Language disturbances from one AD patient to another are reported to be quite heterogeneous [3, 4] and jeopardize AD patients' ability to interact with their environment and verbally communicate [5, 6]. Importantly, the breakdown of communication has been found to be the most difficult consequence of AD for caregivers to cope with [7] and is accompanied by more distress in their supporting role [8].

Language changes occurs in the earliest stages of the disease, including in the pre-AD stage of mild cognitive impairment (MCI) [9]. Most of studies aimed at characterizing the language profile in AD have employed language tasks capable of selectively assessing specific language functions, such as naming [10], syntax [11], or semantic processing [12]. In this type of study, difficulties with picture naming tasks [13-16] represent one of the most frequently reported language impairments [17, 18]. Some evidence indicates that this deficit can appear in early phases of the disease, even at pre-dementia stages, such as in individuals with MCI [15, 19] (for a comparative review of language differences in AD and MCI, see [9]). The nature of naming difficulties in AD is still matter of debate. On the one hand, naming difficulties could derive from the breakdown of semantic cognition, as was evidenced in both implicit (e.g., semantic priming effect [20-24]) and explicit (semantic categorization [25] and semantic knowledge tasks [15]) semantic abilities [26]. On the other hand, naming difficulties can be at least in part due to lexical access difficulties [27-29]. Indeed, AD patients can manifest naming difficulties for stimuli for

which the semantic representation is intact [16]. These lexical access difficulties may also contribute to the consistent impairment of AD patients verbal fluency tasks [30]. Syntax is another facet of language that may be affected by AD, both in comprehension [31] and in expression [32]. With a focus on temporal and phonological features, Szatloczki and colleagues recently reviewed the evolution of language changes in tasks such as reading and naming tasks at different stages of AD. They concluded that more work needs to be done to validate new assessment methods for language function in AD [33].

Interestingly, low scores on standard language tests (such as confrontation naming and verbal fluency) do not fully reflect the actual performance of patients in normal conversation, as they tend to leave out the social and psychological context of language use [34]. Consequently, the assessment of isolated language functions as in naming, fluency, or syntax tasks might not capture the magnitude of problems encountered in everyday communication contexts [34-36]. One way to obtain an ecological approximation of spontaneous discourse abilities in patients is through the connected speech sample (i.e., spoken language production used in a spontaneous and continuous manner) yielded by a picture description task, a narrative task, or an interview [3, 37, 38].

Studies on connected speech in AD have measured different dimensions of connected speech and produced conflicting results. Some studies have found no differences in connected speech characteristics between AD and controls groups [39], while other studies report important differences. Some studies have reported deficits in speech production characteristics (such as melodic line and acoustic features) [40, 41], syntactic complexity (mean length of utterance) [3], lexical content (percentage of pronouns, type-token ratio, mean frequency of words) [42, 43], fluency (revisions and repetitions) [44, 45], and the semantic/discourse aspects of the speech

(“emptiness” of speech) [46]. Many reasons could account for the conflicting results. One reason could be the fact that small samples are usually employed in these studies. This seems especially true regarding the longitudinal data available [3, 47, 48]. Moreover, it has been suggested that language deficits in AD can be heterogeneous [4, 49, 50] and not necessarily apparent in group analyses [39]. Another reason may reside in the methodological approaches of these studies. More specifically, the choice of the tasks and variables used to characterize connected speech varies from one study to another and can therefore yield different results [51]. The study of connected speech would greatly benefit from a comprehensive synthesis of the variables used to analyze different aspects of connected speech in AD and an overview of the main results. This would be helpful for a possible harmonization of connected speech analyses in AD.

An attempt at a comprehensive review of the literature on connected speech has been recently done. More specifically, using an unsystematic narrative review, Boschi, et al. [52] report a series of studies focusing on the analysis of linguistic characteristics of connected speech in the most prevalent neurodegenerative diseases, including AD. This work provides an overview of connected speech impairment elicited by a variety of tasks, including picture descriptions, narrative tasks, and interviews [52]. Their results point to a pattern of deficits on a wide range of variables, including speech rate and hesitations, increased use of pronouns, word finding difficulties, repetitions, revisions, neologisms, inflectional errors, use of discourse markers, low efficiency and cohesion, and uninformative speech that could be part of the signature of AD. Although the article by Boschi et al. (2017) provides a very useful overview of the literature in this field and include a description of databases and search terms, it suffers of some methodological limitations due to the use of an unsystematic narrative review approach. In fact, unsystematic narrative reviews are the traditional approach to summarize the literature on a specific research

topic. However, they are usually not based on a clear and objective method for the search and selection of the articles in the review [53]. In other words, unsystematic narrative reviews are likely to include only research selected by the authors and not all articles available [54]. This can lead to a subjective article selection bias that can affect authors' conclusions and interpretations. Systematic reviews can help overcoming these methodological limitations. Indeed, systematic reviews employ (and explicitly describe) methodological strategies to identify and select all the available publications on a specific research topic [55, 56]. Evidence shows that systematic reviews improve the reliability and the accuracy of the conclusions [57]. The systematic review has therefore become the reference standard for synthesizing evidence in health care because of its methodological rigor and is used to “support the development of clinical practice guidelines and inform clinical decision-making” [55, 56, 58].

A systematic review with meta-analyses has been conducted to synthesize the available data on the word retrieval aspect of connected speech in AD patients [59]. More specifically, Kavé and Goral demonstrated that lexical access—or word-finding—difficulties usually observed in confrontation naming, were also apparent in connected speech in AD elicited by picture descriptions, interviews, picture sequence or film description, and other descriptions [59]. The study by Kavé and Goral exclusively focused on word retrieval, not considering variables of other linguistic domains such as pragmatics and syntax. A systematic review of the characteristic in different linguistic domains is thus necessary in order to better define the global portrait of the connected speech profile in AD, and to complement the findings by Kavé and Goral. In addition, since the constraints of the task used to elicit the connected speech sample are known to influence the most salient variables in AD connected speech [52], a systematic review should probably focus on a single task. To this effect, Mueller et al. recently published a systematic review of connected

speech elicited by picture description tasks in both MCI and AD patients, concluding that the evidence of impairment in the former is inconclusive [60]. Picture-supported narratives have the advantage of providing a relatively constrained discourse task with expected topics, which is not the case for other ecological approximations of spontaneous speech such as open-ended autobiographical questions or conversations [61]. The expected topics allow a more standardized analysis of the lexicosemantic content of speech. Contrary to story retelling tasks, picture description affords the patient with pictorial support, helping persons with significant attentional, executive or memory deficits to produce a sample of connected speech [37]. This approach also has the advantage of capturing multiple aspects of language production using a single task, but the transcription and analysis processes has proven too time consuming for clinical use. As these steps become automated with emerging computational approaches, however, picture description tasks could represent a valuable tool for a rapid screening of language production abilities that can be implemented in the routine neuropsychological battery routinely used with these patients.

In current clinical practice, picture description tasks are administered as part of language batteries such as the Boston Diagnostic Aphasia Examination, in which the clinician counts and rates utterances and clauses [62]. However, this simple analysis does not exploit the richness of the discourse sample. Indeed, as stated by Ahmed and co-authors (2013), connected speech samples “provide a multitude of analytical dimensions” and can be used to extract variables from many different dimensions of connected speech. Unfortunately, the transcription of a verbal sample to a *verbatim* record and the analysis techniques used in research contexts are prohibitively time consuming and labor intensive, making multidimensional analyses difficult to import in clinical settings [63]. This limitation may prevent the wide use of picture description in dementia screening or assessment, despite convincing evidence that a combination of connected speech variables from



different domains can discriminate AD patients from the healthy elderly [40, 46, 51, 64-66] and that different changes occurs at different stages of the disease [47]. The clinical and scientific relevance of picture description tasks in AD patients hinges on specific knowledge of the most relevant variables and on affected language dimensions. Because heterogeneity across patients has been reported [4], it remains unclear if a definite and reliable pattern of language changes occur in AD patients describing a picture. The specific characteristics of connected speech in different linguistic domains remain to be assessed in a systematic way for picture description tasks. The connected speech features of AD in the context of picture description tasks is a question of adequate breadth to warrant a systematic review [67]. A systematic review of connected speech changes in AD as assessed by picture description tasks would provide a characterization, from a large sample, of the most affected dimensions and variables in this group of patients and could help clinicians and researchers choose relevant picture description tasks and develop guidelines for further therapies and studies based on synthesized evidence [56].

The main goal of the study is to systematically review the literature on the connected speech features that characterize AD patients specifically in picture description tasks and gain an overview of the most often affected language dimensions. We review and discuss the most often reported discriminant variables, complementing recent work with a summarization and quantitative appraisal of the available data concerning a specific task and patient group [52, 60]. The multidimensional nature of connected speech analysis and the great number of different variables reported makes summarizing results challenging. There is thus a need to group variables for a clear summarization. In the context of primary progressive aphasia, a progressive neurodegenerative disease characterized by relatively isolated language deterioration, Wilson et al. [68] used a classification model of connected speech adapted from the *Quantitative Production Analysis* [69,

70] encompassing the following dimensions: 1) speech rate and speech errors (such as phonological paraphasias); 2) other disruptions to fluency (such as repetitions and revisions); 3) lexical content (such as number of nouns, pronouns, etc.); and 4) syntactic structure and complexity (such as length of utterances, number of dependent clauses, etc.). An augmented version (including semantic and discourse dimensions) of this framework will be used in the current study. The semantic and discourse dimensions appear important additions to the framework because picture description tasks allow a more standardized assessment of the semantic content and its efficiency and organization compared to interviews.

## **METHODS**

### *Review protocol*

A comprehensive search was conducted in the electronic databases Medline (1946-2016), PubMed, Embase (1974-2016), and PsycInfo using 1) natural language in the title and abstract of references as well as 2) each database's specific descriptors as major topics to retrieve relevant studies (Table 1). We sought help from a professional librarian from the Institut Universitaire de Gériatrie de Montréal. Our last search was run on January 20, 2018. Reference lists of included articles were thoroughly searched for additional references relevant to the review. Additional references were obtained through a search on Google Scholar and Research Net, using the same natural language used in all databases. We followed the PRISMA-P statement [56] for the conduct of this review (Fig. 1 for PRISMA flow-diagram). Given that our goal is to identify the most commonly studied aspects of connected speech proven to be affected in AD patients and the

dimension in which they belong, the systematic review appeared to be the appropriate methodology to match the breadth of our investigation [67].

[insert Figure 1 here]

### *Eligibility criteria*

Inclusion criteria were the following: 1) experimental studies published in peer-reviewed journals, providing quantitative data from a picture description task; 2) presence of a control group; 3) AD is a focus of the study when more than one clinical population is studied; 4) detailed methodology is presented and verbal connected speech samples were collected; 5) no apparent conflict of interest is reported; and 6) article written in English or French. Thus, exclusion criteria were 1) absence of a control group; 2) AD not being the focus of study; 3) apparent conflict of interest between authors and the sponsor; 4) article not written in English or French; and 5) a study unpublished or published after January 20, 2018.

### *Extraction of language features and data summarization*

In each article, we individually extracted the connected speech features that were statistically tested. For summarization and clarity, we categorized the extracted features under six language dimensions following Wilson et al. [68] and Ahmed et al. [3] (adding the semantic and discourse domains as distinct fifth and sixth dimensions). This six-class system is based on the abnormal discourse classification by Saffran et al. [69] and the quantitative production analysis (QPA) of Berndt [70]. The six dimensions in our review are 1) speech production and speech sound errors; 2) other disruptions to fluency; 3) lexical content (lexical features of the words used); 4) syntactic structure and complexity, 5) semantic content (semantic features of the information

content provided), and 6) discourse/pragmatics. We added “Other variables” to accommodate variables that could not readily be assigned to one of the six categories, such as visual paraphasias (replacement of the target word by a word that shares visual features with the target, such as *umbrella* instead of *mushroom*), gestures, and response to errors.

## RESULTS

### *Study characteristics*

#### Stimuli

A total number of 3,075 articles were retrieved after the literature search process. Subsequently, 2,003 duplicates were eliminated, of which 499 were assessed for eligibility. Moreover, 457 did not meet inclusion criteria, for a final total of 44 studies that focused on picture description by AD patients (see Fig. 1 for a detailed flow-chart). Of these, 27 studies (61%) made use of the Cookie Theft Picture from the Boston Diagnostic Aphasia Examination (BDAE) [62]. Nine of those studies [41, 45, 47, 66, 71-75] used the Cookie Theft Picture in conjunction with other pictures/stimuli. Other single-picture stimuli used include various Norman Rockwell pictures [48, 76-79], the Picnic Scene [41, 80] from the Western Aphasia Battery [81], the Tripping Woman Picture [47, 66, 82] from Semenza and Cipolotti [83], and the Bank Robbery picture [4, 6, 84] from the *Protocole Montréal-Toulouse d'examen linguistique de l'aphasie* [85]. The remainder of articles used pictures from a children's book [86] or depictions of various domestic/everyday scenes [45, 51, 87]. In five articles, sequences of multiple pictures were presented, illustrating either the adoption of a dog [44], the chain of events leading to a traffic accident [6, 65, 88], or four sets of different daily life scenes [89].

## Languages

The language spoken by participants was English in 26 of the studies. Other languages were French [4, 6, 84, 88, 90], Brazilian Portuguese [44, 65, 91], Hebrew [43, 92, 93], Chinese [72, 73], Japanese [87, 94], Finnish [95], Italian [74] and German [96].

## Diagnostic criteria and disease severity

Fifteen of the articles surveyed reported results from patients diagnosed with “probable AD”, 26 included patients with “mild AD” and 24 included persons with “moderate AD”. Two studies included patients with “severe AD” [93, 96]. The most commonly reported diagnostic guidelines are those of the NINCDS-ADRDA [97] in 26 studies (59%). Seven studies reported Global Deterioration Scale scores [98] and seven reported Clinical Dementia Rating Scale scores [99]. MMSE [100] scores were reported in 25 of the 44 articles surveyed (57%).

## Aims of studies included in the systematic review

Ten studies (23%) compared AD patients to other clinical syndromes, such as semantic dementia [51], the logopenic variant of primary progressive aphasia [39], fluent aphasia [74, 78, 101], right-brain damage [71], vascular dementia [42, 95], and Parkinson’s disease [102]. Seventeen studies (39%) statistically tested an effect of disease severity. Thirteen studies (30%) statistically contrasted results from more than one task. Twelve of the articles (27%) were concerned strictly with comparing AD patients to a control group.

## Transcription rules and analysis

Twenty-five articles (57%) did not specify if one or more transcribers validated the transcriptions used for analysis. Three studies (7%) explicitly report using automated techniques

for the analysis of transcripts [40, 43, 51]. The most often-used multidimensional analysis guidelines is a variant of the QPA [69]. A recent, augmented version of the QPA is presented in Wilson et al. [68], building on work by Berndt [70], which is used in three articles and the current review. Other popular grids of analysis are the methods described in Croisile et al. [90] and in Tomoeda and Bayles [48], which were used in four and three articles, respectively.

### *Connected speech variables*

The variables tested in the reviewed articles are reported in Table 2 and are organized by their respective domain of connected speech (Fig. 2), totaling 412 statistical tests. Some of the variables found to be significantly different in AD patients are reported in more than one article. To identify the most-often discriminant variables (Fig. 3) across different articles, variables bearing different names but measuring the same connected speech features (e.g., “information units” and “content units”) were grouped together. The eight most often reported significant variables were selected for discussion with a heuristically defined cutoff based on the proportion of significant statistical tests: for inclusion in Fig. 3, a variable had to be tested at least four times and show a significant difference between AD patients and controls on more than 50% of tests. It must be noted that some variables that could be sensitive in distinguishing AD from controls may be underrepresented in this review because they have not been consistently measured across studies. Conversely, some variables that are less sensitive may be overrepresented because they are very routinely assessed in the analysis of connected speech although they are not specifically conceived to differentiate speech characteristics between AD and controls.

[insert Figures 2&3 here]

### Speech rate (speech production)

Speech rate is defined as the number of words divided by the duration of the speech sample [68]. The following variables were also considered a measure of speech rate and included Syllables per minute, Number of words per minute, Phonation rate (Proportion of an utterance that is vocalized, versus silence), and Rate of speech. AD patients spoke slower in 78% of cases (7/9).

### Utterance length (syntactic complexity)

Various measures of syntactic complexity exist, the most common of which is Mean length of utterance (MLU), i.e., the average number of words per utterance [68]. An utterance is defined as a sentence or any effort to express a thought that is terminated by a pause with a falling inflection [62]. Although not technically identical, we counted the following measures under Utterance length: Phrase length, Mean clause length, C-Unit length, Number of words per C-Unit, and Words per clause. AD patients produced shorter utterances in 56% of cases (5/9).

### Pronoun use (lexical content)

We define pronoun use as quantitative differences in usage of pronouns. In pronoun use, we included: Pronoun-to-noun ratio, Number of noun phrases with a pronoun, Anomia index (noun/(noun+pronoun)), % of pronouns of all words, and Pronoun use. We did not include “referential cohesion” or “pronouns without referents (anaphora)”, because these measures are qualitatively different from the unequivocal quantitative measure of the number of pronouns used. We also excluded “person deixis”, as it is not expressed exclusively using personal pronouns and authors treated it separately from anaphoric pronoun use [75]. AD patients used more pronouns in 88% of cases (7/8).

### Word frequency (lexical content)

We refer to Word frequency as a measure of the average “rarity” of the words used by the speaker. A high-frequency word is one that is more common in a corpus of reference for a given language. We included the following variables: Verb frequency, Mean frequency of all words, and Mean frequency of nouns. AD patients used words with higher frequencies in 100% of cases (5/5).

#### Repetitions (disruptions to fluency)

Different authors have used varying definitions of repetition. It can be understood as the immediate, contiguous repetition of the same word [44, 90], repetition of a single word in the same clause [65], cosine distance between clauses (the average amount of identical words in any two utterances, as represented in the vector space) [40] or the inappropriate repetition of an idea [77]. These various definitions were all considered and grouped: Ideational repetitions, Average cosine distance, Cosine cut-off (number of pairs of utterances whose cosine distance is less than 0.5, normalized by total number of unique utterance pairwise comparisons), and Repeated words. AD patients repeated themselves significantly more than controls in 88% of cases (7/8).

#### Word-finding difficulty (disruptions to fluency)

Word-finding difficulties (WFD) are described in Croisile et al. [90] as the absence of production of the target item, indicated by a pause or the production of an indeterminate term. AD patients presented more of these difficulties in 100% of studies (6/6).

#### Information units (semantic content)

An information unit is defined as a truthful, nonredundant piece of information about the stimulus picture [103]. We included the following variables under the umbrella term “information units”: Content units, Total semantic units, Subjects, Objects, Actions, Component measures, Quantity of essential material, Locations, Correct information units, Essential units, Information



conveyed, Information content, Number of content units, Repetition of expected ideas, Pictorial themes, Number of relevant descriptions, Key words, Places, Main concept score, and Localizations. AD patients provided fewer information units on 85% of the statistical comparisons (45/53).

### Efficiency and idea density (discourse)

Efficiency is the rate at which information is conveyed [46], and idea density (or conciseness) is the average number of ideas expressed per given number of words. Efficiency is based on speech duration and not on words spoken. We group idea density and efficiency because they both reflect the ability, at the discourse level, to produce relevant content efficiently. The following terms were included as corresponding to efficiency/idea density: Correct information units, Efficiency ratio, Narrative efficiency, Words per information units, Index of discourse effectiveness, Efficiency index, Information units per second, Conciseness, Conciseness ratio, Conciseness index, Number of accurate and complete concepts per minute, Efficiency of description, and Lexical index (ratio of the number of expected information units provided on the total number of words uttered). AD patients required more time or words to convey information in 86% of the measures (18/21).

## **DISCUSSION**

In this systematic review, we aimed to determine the different connected speech dimensions affected in AD patients in picture description tasks. Through an exhaustive review of 44 articles, we compiled a total of 412 statistical tests of a wide array of variables, from which we isolated the most often reported as discriminant between AD patients and controls. The following

eight variables belong in different dimensions of connected speech: speech production (rate of speech), syntactic complexity (MLU), lexical content (use of pronouns and word frequency), disruptions to fluency (repetitions and word-finding difficulties), semantic content (information units), and discourse (efficiency). These results highlight the importance of a multidimensional assessment of connected speech to aid in differential diagnosis of AD and for monitoring communicative abilities with disease progression.

### *Speech production*

Among the variables belonging to the category of speech production, rate of speech seems to be the variable showing a consistent difference between AD and controls. AD patients are reported to have, on average, a slower speech output (fewer words per minute) than the healthy elderly. While neural correlates of rate of speech have never been investigated in AD, it has been associated with damage to the left inferior frontal gyrus in primary progressive aphasia [104]. In individual AD patients, speech rate was not found to consistently decline with disease progression, and the measure has proven unstable on test-retest of patients [3]. These observations cast doubt on the reliability, and consequently on the clinical usefulness of this measure. However, acoustic features of speech (such as spectral characteristics of the voice signal) should not be discounted, as recent analyses have proven sensitive to articulatory changes associated with AD [40, 105].

### *Syntactic complexity*

Concerning the syntactic aspects of connected speech, the average length of utterance was the most often studied variable in AD patients. It was found to be shorter for AD patients, which

is interpreted as shorter and more simple sentences. A caveat of this measure is its sensitivity to the boundary placement in the transcription process, i.e., what is considered an utterance. It may thus be inconsistently calculated across different studies [68]. This is especially worrisome considering that most of the articles included in this systematic review did not specify transcription rules and guidelines nor explicitly state validation of the transcribed data. Our results are not consistent with the results presented in Boschi et al. for picture description tasks. Indeed, 5 out of 9 of the articles included in the present systematic review showed a statistically significant difference, whereas Boschi et al. report significant differences in only 2 out of 6 articles. It must be noted that this discrepancy is probably due to the fact that our systematic review included a greater number of articles with different datasets and transcription guidelines. Moreover, 3 of the 6 papers that did not show significant difference in Boschi et al. were based on the same connected speech dataset (DementiaBank). Although length of utterance may not be the most reliable index of syntactic complexity, its shortening appears to be detectable in picture descriptions of AD patients. Recently, Garrard et al. [106] have proposed a systematic tool for the automatic alignment of transcripts and automatic quantification of discrepancies. Such a system could be helpful to ensure reproducibility of studies and standardized data preparation for computerized analyses. In spite of these difficulties, syntax as a connected speech variable may not be discarded, as we know that AD patients exhibit changes in both the comprehension [31, 107] and expression of syntax [32, 108]. Comprehension of complex syntax is thought to place a heavy demand on working memory [107], which leaves AD patients at a disadvantage.

### *Lexical content*

Evidence of lexical content impairments was measured using two main variables, namely use of pronouns and frequency. AD patients showed increased reliance on pronouns compared to controls. This has been attributed to their semantic impairment and lexical access difficulties. The use of pronouns allows them to maintain relatively fluent speech in the face of lexico-semantic difficulties, substituting a pronoun in the place of a target noun they are unable to accurately name (*she* instead of *mother*, *this* instead of *kite*, etc.). An alternative hypothesis is that the use of pronouns is related to working memory deficits [109], an explanation known as the “working memory impairment hypothesis”. Almor et al. [109] argue that an increased use of pronouns in connected speech is linked to working memory problems but neither to dementia severity nor semantic impairment. These authors suggest that AD patients struggle to keep a fresh activation of semantic representations in working memory and thus rely on pronouns and very high-frequency words (see discussion below).

In their picture descriptions, AD patients tend to use more high-frequency words than controls. In the studies surveyed, this effect appeared in overall lexical content [93] and specifically in verbs [40] and nouns [43]. Word frequency has been tested five times in three articles in our review, and its effect on AD connected speech is not well documented. In confrontation naming tasks, however, various psycholinguistic variables have been shown to significantly impact the performance of AD patients: age-of-acquisition, name agreement, word frequency, and familiarity [110, 111]. How these variables impact the multiple dimensions of connected speech in a picture description task remains unclear, but their analysis is particularly well suited for computational analyses. In fact, automated algorithms could help extract psycholinguistic variables from connected speech transcriptions relying on open access databases [112]. Our results also highlight that the often-tested variables of TTR and open/closed-class word ratio are not sensitive to the

lexical and semantic impairment of AD patients. In other words, the supposed reduction in vocabulary and WFDs of AD patients cannot be reliably be measured using these common metrics in picture descriptions. However, vocabulary size can be investigated with other metrics that may be more sensitive to impairment. Recent work in natural language processing successfully enriched speech transcripts of MCI patients with semantic information from word embeddings and boosted classification accuracy [113]. Another example is how Hoffman et al. have applied latent semantic analysis [114] to extract the semantic diversity (the number of different contexts in which they appear) of words used by patients with semantic dementia [115]. Hence, vocabulary richness may be a valuable construct to study neurodegenerative diseases, despite disappointing results from previously used metrics such as TTR and open/closed-class word ratio.

### *Disruptions to fluency*

Disruptions of fluency in AD connected speech have been detected with repetitions and word-finding difficulties. Measures of repetitive content are operationalized in different ways between authors. When understood as the immediate, contiguous production of the same word [44, 90], we are inclined to interpret repetition as a consequence of WFD as in Forbes-McKay and Venneri [66]. Indeed, less anomic patients were reported to produce fewer repetitions [101]. The inappropriate repetition of an idea [77] and cosine distance between clauses (the extent to which two given utterances contain the same words) [40], however, could be attributed to memory deficits typically associated with AD. Hence, different measures of repetitiveness may reflect distinct cognitive/behavioral mechanisms, but it seems that all these distinct measures of repetitiveness are sensitive to AD in picture description tasks.

Even though word-finding difficulties increase with normal aging, AD patients experience more WFDs, or anomia, than healthy elderly people. In normal aging, this phenomenon is largely attributed to lexical access difficulties. As we noted in our introduction through discussion of picture naming and verbal fluency, the increased prevalence of WFD in AD could be linked to a combination of impaired lexical access and a degradation of semantic cognition. WFD does occur in picture description tasks, but one drawback of its use as an outcome variable in picture description tasks is that it may be difficult to assess in a standardized manner. It has been defined as “the absence of production of the target item, indicated by a pause or the production of an indefinite term” and as “indicated by a pause, an immediate repetition of a previous word or production of an indefinite term” [90]. Thus defined, this measure requires careful and time consuming manual examination of both the audio and transcribed speech data to be properly operationalized. One observation of interest is that the words retrieved after a word-finding pause tend to be of higher frequency for AD patients [116], a finding that highlights the importance of this variable when administering picture description tasks. In a recent review of word retrieval in connected speech, Kavé and Goral argue for the importance of assessing word retrieval in speech, and not only in single-words naming tasks [59].

### *Semantic content and discourse*

In the discourse and semantic domains, the most often reported significant variables are efficiency and number of information units, respectively. Speech of AD patients becomes noninformative and empty with disease progression [101]. The lower efficiency and the fewer number of information units conveyed by AD patients may reflect deficits in lexical access, semantic impairment, or both. The question of the extent to which each of these mechanisms is

shared by the naming difficulties of AD patients is still a matter of debate [16]. Alternatively, discourse efficiency has also been linked to executive function [117], which is known to be impaired in AD [118]. In the discourse domain, AD patients also have more trouble maintaining the theme, despite the pictorial support of picture description tasks [42, 86]. Information units and efficiency remain by far the most-often reported variables in picture description tasks, as they can serve as a quantitative measure of the so-called *emptiness* of AD discourse captured through picture description tasks. This emptiness is not solely attributable to the perceptual analysis of the picture, as deficits in information content were also observed in interviews [51] and informal conversation [119] with AD patients.

#### *Limitations and further study*

A better understanding of the relationship between connected speech and other language tests such as confrontation naming and verbal fluency could help delineate difficulties caused by impaired lexical access versus semantic degradation in AD discourse. Kavé and Goral [43] have argued that scores on confrontation naming were in fact associated with WFD in connected speech, whereas verbal fluency tasks were not as useful to predict lexical retrieval in a picture description task. This discrepancy is attributed to the more similar cognitive demands of picture naming and picture description, whereas the latter is less reliant on executive function compared to lexical fluency tasks. The relationship between connected speech variables and performance on confrontation naming of animals, objects, and especially unique entities such as famous people and buildings, however, needs further study [52].

From a methodological point of view, a clear majority of studies (60%) used the Cookie Theft Picture from the BDAE, which depicts an everyday scene that can be described in short and simple language, using very high-frequency words (e.g., girl, boy, mother, water, etc.). Some authors have emphasized the need for more complex pictures to shed light on early, subtle connected speech abnormalities [47, 65, 66, 82]. Not all variables, however, seem to be affected equally by heightened complexity, as is the case with deixis [75], meaning that results obtained with one picture description task are not necessarily generalizable to others.

A minority (30%) of the studies reported in this review were published from 2011 onwards, a year that saw the publication of the revised NINCDS-ADRDA criteria for AD [1]. Thus, we cannot exclude that some patients from earlier studies would receive a different diagnosis today, nor that the stages of disease and mixed-profile presentations correspond to present-day diagnostic criteria. Although connected speech disturbances caused in AD patients have been described as heterogeneous across patients [4], a consistent multidimensional pattern of connected speech impairment has successfully been extracted with machine learning techniques and a factor analysis [40]. These computational results overlap with much of the previous research summarized in this review (Fig. 1). Another critical issue that emerged in the research summarized in this review is the transcription process itself. As we stated in our discussion of syntactic content, most articles did not specify transcription guidelines, and this should be addressed in future research to ensure reproducibility of results [106].

Our systematic review included only connected speech studies elicited by picture description. A major advantage of picture description tasks is their ability to quickly capture a multidimensional sample of language variables [3]. The recent advances in computing techniques may enable short, automated analyses of discourse samples [63]. Hence, picture description tasks



are of obvious interest in clinical settings, where a simple three-minute, 150-word discourse sample offers a wealth of information about a patient's cognitive status and communicative abilities [51]. Moreover, picture description tasks provide an opportunity for cost-efficient multiple time-point testing in situations when one or more comprehensive language examinations from a speech-language pathologist are not feasible. For example, they could be used to routinely monitor the communicative skills of AD patients, as these critical abilities are known to decline with disease progression and are accompanied by various negative outcomes [8, 120, 121]. Additionally, picture description tasks could inform efficient communication strategies for caregivers and possible interventions with the patient that are tailored with its language profile.

Our results reveal that a wide array of language variables has proven useful to distinguish AD patients from the healthy elderly and to follow the course of disease progression, highlighting the need to go beyond tasks such as verbal fluency and confrontation naming and consider connected speech as provided by picture description tasks [37]. Current literature on the connected speech of AD patients favors a multidimensional approach [3, 40, 47], but the need for standardization of analytic procedures has been underlined [52]. A recent review of connected speech in neurodegenerative diseases has added a valuable contribution in this direction by synthesizing a great quantity of the available evidence [52].

### *Conclusion*

The present study represents the first attempt to systematically revise the literature on connected speech elicited by picture description in AD. The results give an overview of the multiplicity of variables studied in this field and the main results. Our review highlights the

importance of using a multidimensional analysis approach capable of extracting and measuring syntactic, lexical, fluency, and semantic features in spontaneous speech in AD. This approach leads to a comprehensive overview of the language production abilities of each patient. This information can be relevant not only for patient characterization and differential diagnosis but can also help caregivers and, eventually, contribute to refining intervention strategies. We also suggest the importance of developing automatic analysis tools to make the assessment of connected speech more suitable for clinical settings. Most of the analyses conducted in the articles surveyed rely on error-prone and time-consuming methods. This has recently been reported in a connected speech review by Boschi et al. [52] and is confirmed in our study. As programmers continue to meet computational challenges relevant to the study of normal and pathological discourse analysis, new technology stemming from these advances will enter hospitals and nursing homes, to the benefit of the patient, caregivers and the healthcare system (see Aluísio et al. [63] for one effort in this direction). We thus expect picture description tasks to become an important tool of speech-language-pathologists aiming to promote choice, dignity and engagement in meaningful activities through person-centered care [122]. Automated procedures have also been shown to produce reasonable accuracy in the classification of patients with AD [40, 64, 123] and primary-progressive aphasia [124]. With a focus on existing data, this review identified multidimensional variables that should become a target for the new computational tools that are to facilitate AD research and management.

## **ACKNOWLEDGMENTS**

We wish to extend our sincere gratitude to B ereng ere Houz e for her helpful comments on the manuscript and to Audrey Attia for her assistance with the conduct of the review protocol.

SMB is supported by a NSERC Discovery Grant and a Chercheur-boursier junior 2 award from the FRQS. MM is supported by a FRQS doctoral award. AS is supported by a Graduate scholarship award from the CIHR.

Authors' disclosures available online (<https://www.j-alz.com/manuscript-disclosures/17-0881r3>).

## REFERENCES

- [1] McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, Kawas CH, Klunk WE, Koroshetz WJ, Manly JJ, Mayeux R (2011) The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* **7**, 263-269.
- [2] Kempler D (1995) Language changes in dementia of the Alzheimer type. In *Dementia and Communication*, Lubinski R, ed. Singular Publishing Group, San Diego, CA, pp. 98-114.
- [3] Ahmed S, Haigh A-MF, de Jager CA, Garrard P (2013) Connected speech as a marker of disease progression in autopsy-proven Alzheimer's disease. *Brain* **136**, 3727-3737.
- [4] Duong A, Giroux F, Tardif A, Ska B (2005) The heterogeneity of picture-supported narratives in Alzheimer's disease. *Brain Lang* **93**, 173-184.
- [5] Lee H (2012) Thesis: Langage et maladie d'Alzheimer: analyse multidimensionnelle d'un discours pathologique. Université Paul Valéry-Montpellier III.
- [6] Ska B, Duong A (2005) Communication, discours et démence. *Psychol Neuropsychiatr Vieil* **3**, 125-133.
- [7] Murray J, Schneider J, Banerjee S, Mann A (1999) EURO CARE: a cross-national study of co-resident spouse carers for people with Alzheimer's disease: II—a qualitative analysis of the experience of caregiving. *Int J Geriatr Psychiatry* **14**, 662-667.
- [8] Orange JB, Colton-Hudson A (1998) Enhancing communication in dementia of the Alzheimer's type. *Top Geriatr Rehabil* **14**, 56-75.
- [9] Taler V, Phillips NA (2008) Language performance in Alzheimer's disease and mild cognitive impairment: A comparative review. *J Clin Exp Neuropsychol* **30**, 501-556.
- [10] Bayles KA, Trosset MW (1992) Confrontation naming in Alzheimer's patients: Relation to disease severity. *Psychol Aging* **7**, 197.
- [11] Kemper S, LaBarge E, Ferraro FR, Cheung H, Cheung H, Storandt M (1993) On the preservation of syntax in Alzheimer's disease: Evidence from written sentences. *Arch Neurol* **50**, 81-86.
- [12] Albert M, Milberg W (1989) Semantic processing in patients with Alzheimer's disease. *Brain Lang* **37**, 163-171.

- [13] Bowles NL, Obler LK, Albert ML (1987) Naming errors in healthy aging and dementia of the Alzheimer type. *Cortex* **23**, 519-524.
- [14] Laws KR, Adlington RL, Gale TM, Moreno-Martínez FJ, Sartori G (2007) A meta-analytic review of category naming in Alzheimer's disease. *Neuropsychologia* **45**, 2674-2682.
- [15] Joubert S, Brambati SM, Ansado J, Barbeau EJ, Felician O, Didic M, Lacombe J, Goldstein R, Chayer C, Kergoat M-J (2010) The cognitive and neural expression of semantic memory impairment in mild cognitive impairment and early Alzheimer's disease. *Neuropsychologia* **48**, 978-988.
- [16] Montembeault M, Brambati S, Joubert S, Boukadi M, Chapleau M, Laforce RJ, Wilson M, Macoir J, Rouleau I (2017) Naming unique entities in the semantic variant of primary progressive aphasia and Alzheimer's disease: Towards a better understanding of the semantic impairment. *Neuropsychologia* **95**, 11-20.
- [17] Hodges JR, Salmon DP, Butters N (1991) The nature of the naming deficit in Alzheimer's and Huntington's disease. *Brain* **114**, 1547-1558.
- [18] Bayles KA, Tomoeda CK (1983) Confrontation naming impairment in dementia. *Brain Lang* **19**, 98-114.
- [19] Adlam A-LR, Bozeat S, Arnold R, Watson P, Hodges JR (2006) Semantic knowledge in mild cognitive impairment and mild Alzheimer's disease. *Cortex* **42**, 675-684.
- [20] Chertkow H, Bub D (1990) Semantic memory loss in dementia of Alzheimer's type. *Brain* **113**, 397-417.
- [21] Predovan D, Gandini D, Montembeault M, Rouleau I, Bherer L, Joubert S, Brambati SM (2014) Loss of person-specific knowledge in Alzheimer's disease: Evidence from priming. *Neurocase* **20**, 263-268.
- [22] Giffard B, Laisney M, Mézenge F, De La Sayette V, Eustache F, Desgranges B (2008) The neural substrates of semantic memory deficits in early Alzheimer's disease: clues from semantic priming effects and FDG-PET. *Neuropsychologia* **46**, 1657-1666.
- [23] Giffard B, Laisney M, Desgranges B, Eustache F (2015) An exploration of the semantic network in Alzheimer's disease: Influence of emotion and concreteness of concepts. *Cortex* **69**, 201-211.

- [24] Loureiro IS, Lefebvre L (2016) Distinct progression of the deterioration of thematic and taxonomic links in natural and manufactured objects in Alzheimer's disease. *Neuropsychologia* **91**, 426-434.
- [25] Hodges JR, Patterson K (1995) Is semantic memory consistently impaired early in the course of Alzheimer's disease? Neuroanatomical and diagnostic implications. *Neuropsychologia* **33**, 441-459.
- [26] Verma M, Howard R (2012) Semantic memory and language dysfunction in early Alzheimer's disease: a review. *Int J Geriatr Psychiatry* **27**, 1209-1217.
- [27] Nebes RD, Martin DC, Horn LC (1984) Sparing of semantic memory in Alzheimer's disease. *J Abnorm Psychol* **93**, 321.
- [28] Ober BA, Shenaut GK (1995) Semantic priming in Alzheimer's disease: Meta-analysis and theoretical evaluation. *Adv Psychol* **110**, 247-271.
- [29] Rich JB, Park NW, Dopkins S, Brandt J (2002) What do Alzheimer's disease patients know about animals? It depends on task structure and presentation format. *J Int Neuropsychol Soc* **8**, 83-94.
- [30] Henry JD, Crawford JR, Phillips LH (2004) Verbal fluency performance in dementia of the Alzheimer's type: a meta-analysis. *Neuropsychologia* **42**, 1212-1222.
- [31] Giovannetti T, Hopkins MW, Crawford J, Bettcher BM, Schmidt KS, Libon DJ (2008) Syntactic comprehension deficits are associated with MRI white matter alterations in dementia. *J Int Neuropsychol Soc* **14**, 542-551.
- [32] Pakhomov S, Chacon D, Wicklund M, Gundel J (2011) Computerized assessment of syntactic complexity in Alzheimer's disease: a case study of Iris Murdoch's writing. *Behav Res Methods* **43**, 136-144.
- [33] Szatloczki G, Hoffmann I, Vincze V, Kalman J, Pakaski M (2015) Speaking in Alzheimer's disease, is that an early sign? importance of changes in language abilities in Alzheimer's disease. *Front Aging Neurosci* **7**, 195.
- [34] Sabat SR (1994) Language function in Alzheimer's disease: a critical review of selected literature. *Lang Commun* **14**, 331-351.
- [35] Mueller KD, Kosciak RL, Turkstra LS, Riedeman SK, LaRue A, Clark LR, Hermann B, Sager MA, Johnson SC (2016) Connected Language in Late Middle-Aged Adults at Risk for Alzheimer's Disease. *J Alzheimers Dis* **54**, 1539-1550.

- [36] Bucks RS, Singh S, Cuerden JM, Wilcock GK (2000) Analysis of spontaneous, conversational speech in dementia of Alzheimer type: Evaluation of an objective technique for analysing lexical performance. *Aphasiology* **14**, 71-91.
- [37] Giles E, Patterson K, Hodges JR (1996) Performance on the Boston Cookie Theft picture description task in patients with early dementia of the Alzheimer's type: missing information. *Aphasiology* **10**, 395-408.
- [38] Doyle PJ, Goda AJ, Spencer KA (1995) The communicative informativeness and efficiency of connected discourse by adults with aphasia under structured and conversational sampling conditions. *Am J Speech Lang Pathol* **4**, 130-134.
- [39] Ahmed S, de Jager CA, Haigh A-MF, Garrard P (2012) Logopenic aphasia in Alzheimer's disease: clinical variant or clinical feature? *J Neurol Neurosurg Psychiatry* **83**, 1056-1062.
- [40] Fraser KC, Meltzer JA, Rudzicz F (2016) Linguistic features identify Alzheimer's disease in narrative speech. *J Alzheimers Dis* **49**, 407-422.
- [41] Cummings JL, Benson DF, Hill MA, Read S (1985) Aphasia in dementia of the Alzheimer type. *Neurology* **35**, 394-394.
- [42] Hier DB, Hagenlocker K, Shindler AG (1985) Language disintegration in dementia: Effects of etiology and severity. *Brain Lang* **25**, 117-133.
- [43] Kavé G, Goral M (2016) Word retrieval in picture descriptions produced by individuals with Alzheimer's disease. *J Clin Exp Neuropsychol* **38**, 958-966.
- [44] de Lira JO, Ortiz KZ, Campanha AC, Bertolucci PHF, Minett TSC (2011) Microlinguistic aspects of the oral narrative in patients with Alzheimer's disease. *Int Psychogeriatr* **23**, 404-412.
- [45] Ehrlich JS, Obler LK, Clark L (1997) Ideational and semantic contributions to narrative production in adults with dementia of the Alzheimer's type. *J Commun Disord* **30**, 79-99.
- [46] Ahmed S, de Jager CA, Haigh A-M, Garrard P (2013) Semantic processing in connected speech at a uniformly early stage of autopsy-confirmed Alzheimer's disease. *Neuropsychology* **27**, 79.
- [47] Forbes-McKay K, Shanks MF, Venneri A (2013) Profiling spontaneous speech decline in Alzheimer's disease: a longitudinal study. *Acta Neuropsychiatr* **25**, 320-327.
- [48] Tomoeda CK, Bayles KA (1993) Longitudinal effects of Alzheimer disease on discourse production. *Alzheimer Dis Assoc Disord* **7**, 223-236.

- [49] Pravatà E, Tavernier J, Parker R, Vavro H, Mintzer JE, Spampinato MV (2016) The neural correlates of anomia in the conversion from mild cognitive impairment to Alzheimer's disease. *Neuroradiology* **58**, 59-67.
- [50] Frings L, Kloppel S, Teipel S, Peters O, Frolich L, Pantel J, Schroder J, Gertz H-J, Arlt S, Heuser I (2011) Left anterior temporal lobe sustains naming in Alzheimer's dementia and mild cognitive impairment. *Curr Alzheimer Res* **8**, 893-901.
- [51] Sajjadi SA, Patterson K, Tomek M, Nestor PJ (2012) Abnormalities of connected speech in semantic dementia vs Alzheimer's disease. *Aphasiology* **26**, 847-866.
- [52] Boschi V, Catricalà E, Consonni M, Chesi C, Moro A, Cappa SF (2017) Connected speech in neurodegenerative language disorders: a review. *Front Psychol* **8**, 269.
- [53] Mulrow CD (1987) The medical review article: state of the science. *Ann Intern Med* **106**, 485-488.
- [54] Montori VM, Swiontkowski MF, Cook DJ (2003) Methodologic issues in systematic reviews and meta-analyses. *Clin Orthop Relat Res* **413**, 43-54.
- [55] Mulrow CD (1994) Systematic reviews: rationale for systematic reviews. *BMJ* **309**, 597-599.
- [56] Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA (2015) Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* **4**, 1.
- [57] Cipriani A, Geddes J (2003) Comparison of systematic and narrative reviews: the example of the atypical antipsychotics. *Epidemiol Psychiatr Sci* **12**, 146-153.
- [58] Pae CU (2015) Why systematic review rather than narrative review? *Psychiatry Investig* **12**, 417-419.
- [59] Kavé G, Goral M (2018) Word retrieval in connected speech in Alzheimer's disease: a review with meta-analyses. *Aphasiology* **32**, 4-26.
- [60] Mueller KD, Hermann B, Mecollari J, Turkstra LS (2018) Connected speech and language in mild cognitive impairment and Alzheimer's disease: A review of picture description tasks. *J Clin Exp Neuropsychol*, doi: 10.1080/13803395.2018.1446513.
- [61] Chenery HJ, Murdoch BE (1994) The production of narrative discourse in response to animations in persons with dementia of the Alzheimer's type: preliminary findings. *Aphasiology* **8**, 159-171.



- [62] Goodglass H, Kaplan E, Barresi B (2000) *Boston Diagnostic Aphasia Examination-(BDAE-3)*. Psychological Corporation, San Antonio. TX.
- [63] Aluisio S, Cunha A, Toledo C, Scarton C (2016) A computational tool for automated language production analysis aimed at dementia diagnosis. In *International Conference on Computational Processing of the Portuguese Language, XII; Demonstration Session*, University of Lisbon.
- [64] Rentoumi V, Raoufian L, Ahmed S, de Jager CA, Garrard P (2014) Features and machine learning classification of connected speech samples from patients with autopsy proven Alzheimer's disease with and without additional vascular pathology. *J Alzheimers Dis* **42**, S3-S17.
- [65] Drummond C, Coutinho G, Fonseca RP, Assunção N, Teldeschi A, de Oliveira-Souza R, Moll J, Tovar-Moll F, Mattos P (2015) Deficits in narrative discourse elicited by visual stimuli are already present in patients with mild cognitive impairment. *Front Aging Neurosci* **7**, 96.
- [66] Forbes-McKay K, Venneri A (2005) Detecting subtle spontaneous language decline in early Alzheimer's disease with a picture description task. *Neurol Sci* **26**, 243-254.
- [67] Dijkers M (2015) What is a scoping review? *KT Update* **4**, 1-5.
- [68] Wilson SM, Henry ML, Besbris M, Ogar JM, Dronkers NF, Jarrold W, Miller BL, Gorno-Tempini ML (2010) Connected speech production in three variants of primary progressive aphasia. *Brain* **133**, 2069-2088.
- [69] Saffran EM, Berndt RS, Schwartz MF (1989) The quantitative analysis of agrammatic production: Procedure and data. *Brain Lang* **37**, 440-479.
- [70] Berndt RS (2000) *Quantitative production analysis a training manual for the analysis of aphasic sentence production*, Psychology Press.
- [71] Cherney LR, Canter GJ (1993) Informational content in the discourse of patients with probable Alzheimer's disease and patients with right brain damage. *Clin Aphasiol* **21**, 123-134.
- [72] Lai Y, Lin Y (2012) Discourse markers produced by Chinese-speaking seniors with and without Alzheimer's disease. *J Pragmatics* **44**, 1982-2003.

- [73] Lai Y, Pai H (2009) To be semantically-impaired or to be syntactically-impaired: Linguistic patterns in Chinese-speaking persons with or without dementia. *J Neurolinguistics* **22**, 465-475.
- [74] Carlomagno S, Santoro A, Menditti A, Pandolfi M, Marini A (2005) Referential communication in Alzheimer's type dementia. *Cortex* **41**, 520-534.
- [75] March EG, Wales R, Pattison P (2006) The uses of nouns and deixis in discourse production in Alzheimer's disease. *J Neurolinguistics* **19**, 311-340.
- [76] Bayles KA, Tomoeda CK, McKnight PE, Helm-Estabrooks N, Hawley JN (2004) Verbal perseveration in individuals with Alzheimer's disease. *Semin Speech Lang* **26**, 335-347.
- [77] Tomoeda CK, Bayles KA, Trosset MW, Azuma T, McGeagh A (1996) Cross-sectional analysis of Alzheimer disease effects on oral discourse in a picture description task. *Alzheimer Dis Assoc Disord* **10**, 204-215.
- [78] Chapman SB, Highley AP, Thompson JL (1998) Discourse in fluent aphasia and Alzheimer's disease: Linguistic and pragmatic considerations. *J Neurolinguistics* **11**, 55-78.
- [79] Chapman SB, Ulatowska HK, King K, Johnson JK, McIntire DD (1995) Discourse in early Alzheimer's disease versus normal advanced aging. *Am J Speech Lang Pathol* **4**, 124-129.
- [80] Smith SR, Chenery HJ, Murdoch BE (1989) Semantic abilities in dementia of the Alzheimer type. II. Grammatical semantics. *Brain Lang* **36**, 533-542.
- [81] Kertesz A (1982) *Western aphasia battery test manual*, Psychological Corp.
- [82] Forbes KE, Venneri A, Shanks MF (2002) Distinct patterns of spontaneous speech deterioration: an early predictor of Alzheimer's disease. *Brain Cogn* **48**, 356-361.
- [83] Semenza C, Cipolotti L (1989) *Neuropsicologia con carta e matita*. Cleup Editrice Padova, Padova.
- [84] Duong A, Tardif A, Ska B (2003) Discourse about discourse: what is it and how does it progress in Alzheimer's disease? *Brain Cogn* **53**, 177-180.
- [85] Nespoulous J, Lecours AR, Lafond D, Lemay A, Puel M, Joannette Y, Cot F, Rascol A (1992) *Protocole Montréal-Toulouse d'examen linguistique de l'aphasie (MT86)*. L'Ortho-Edition. Isbergues, France.
- [86] Ash S, Moore P, Vesely L, Grossman M (2007) The decline of narrative discourse in Alzheimer's disease. *Brain Lang* **103**, 181-182.

- [87] Shimada M, Meguro K, Yamazaki H, Horikawa A, Hayasaka C, Yamaguchi S, Yamaguchi K, Katsuyama N, Nakano M, Yamadori A (1998) Impaired verbal description ability assessed by the picture description task in Alzheimer's disease. *Arch Gerontol Geriatr* **27**, 57-65.
- [88] St-Pierre M-C, Ska B, Béland R (2005) Lack of coherence in the narrative discourse of patients with dementia of the Alzheimer's type. *J Multiling Commun Disord* **3**, 211-215.
- [89] Kong AP-H, Whiteside J, Bargmann P (2016) The Main Concept Analysis: Validation and sensitivity in differentiating discourse produced by unimpaired English speakers from individuals with aphasia and dementia of Alzheimer type. *Logoped Phoniatr Vocol* **41**, 129-141.
- [90] Croisile B, Ska B, Brabant M-J, Duchene A, Lepage Y, Aimard G, Trillet M (1996) Comparative study of oral and written picture description in patients with Alzheimer's disease. *Brain Lang* **53**, 1-19.
- [91] de Lira JO, Minett TSC, Bertolucci PHF, Ortiz KZ (2014) Analysis of word number and content in discourse of patients with mild to moderate Alzheimer's disease. *Dement Neuropsychol* **8**, 260-265.
- [92] Kavé G, Levy Y (2003) Morphology in picture descriptions provided by persons with Alzheimer's disease. *J Speech Lang Hear Res* **46**, 341-352.
- [93] Kavé G, Dassa A (2018) Severity of Alzheimer's disease and language features in picture descriptions. *Aphasiology* **32**, 27-40.
- [94] Choi H (2009) Performances in a picture description task in Japanese patients with Alzheimer's disease and with mild cognitive impairment. *Korean J Commun Disord* **14**, 326-337.
- [95] Vuorinen E, Laine M, Rinne J (2000) Common pattern of language impairment in vascular dementia and in Alzheimer disease. *Alzheimer Dis Assoc Disord* **14**, 81-86.
- [96] Bschor T, Kuhl K-P, Reischies FM (2001) Spontaneous speech of patients with dementia of the Alzheimer type and mild cognitive impairment. *Int Psychogeriatr* **13**, 289-298.
- [97] McKhann GM, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM (1984) Clinical diagnosis of Alzheimer's disease Report of the NINCDS-ADRDA Work Group\* under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* **34**, 939-939.

- [98] Reisberg B, Ferris SH, de Leon MJ, Crook T (1982) The Global Deterioration Scale for assessment of primary degenerative dementia. *Am J Psychiatry* **139**, 1136-1139.
- [99] Hughes CP, Berg L, Danziger WL, Coben LA, Martin R (1982) A new clinical scale for the staging of dementia. *Br J Psychiatry* **140**, 566-572.
- [100] Folstein MF, Folstein SE, McHugh PR (1975) "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* **12**, 189-198.
- [101] Nicholas M, Obler LK, Albert ML, Helm-Estabrooks N (1985) Empty speech in Alzheimer's disease and fluent aphasia. *J Speech Lang Hear Res* **28**, 405-410.
- [102] McNamara P, Obler LK, Au R, Durso R, Albert ML (1992) Speech monitoring skills in Alzheimer's disease, Parkinson's disease, and normal aging. *Brain Lang* **42**, 38-51.
- [103] Bayles KA, Boone DR, Tomoeda CK, Slauson TJ, Kaszniak AW (1989) Differentiating Alzheimer's patients from the normal elderly and stroke patients with aphasia. *J Speech Hear Disord* **54**, 74-87.
- [104] Amici S, Ogar J, Brambati SM, Miller BL, Neuhaus J, Dronkers NL, Gorno-Tempini ML (2007) Performance in specific language tasks correlates with regional volume changes in progressive aphasia. *Cogn Behav Neurol* **20**, 203-211.
- [105] Meilán JJ, Martínez-Sánchez F, Carro J, Sánchez JA, Pérez E (2012) Acoustic markers associated with impairment in language processing in Alzheimer's disease. *Span J Psychol* **15**, 487-494.
- [106] Garrard P, Haigh A-M, de Jager C (2011) Techniques for transcribers: assessing and improving consistency in transcripts of spoken language. *Lit Ling Comput* **26**, 389-405.
- [107] Small JA, Kemper S, Lyons K (1997) Sentence comprehension in Alzheimer's disease: Effects of grammatical complexity, speech rate, and repetition. *Psychol Aging* **12**, 3.
- [108] Altmann LJ, Kempler D, Andersen ES (2001) Speech errors in Alzheimer's disease: reevaluating morphosyntactic preservation. *J Speech Lang Hear Res* **44**, 1069-1082.
- [109] Almor A, Kempler D, MacDonald MC, Andersen ES, Tyler LK (1999) Why do Alzheimer patients have difficulty with pronouns? Working memory, semantics, and reference in comprehension and production in Alzheimer's disease. *Brain Lang* **67**, 202-227.
- [110] Tippett LJ, Meier SL, Blackwood K, Diaz-Asper C (2007) Category specific deficits in Alzheimer's disease: fact or artefact? *Cortex* **43**, 907-920.

- [111] Taylor R (1998) Effects of age of acquisition, word frequency, and familiarity on object recognition and naming in dementia. *Percept Mot Skills* **87**, 573-574.
- [112] Laske C, Sohrabi HR, Frost SM, López-de-Ipiña K, Garrard P, Buscema M, Dauwels J, Soekadar SR, Mueller S, Linnemann C (2015) Innovative diagnostic tools for early detection of Alzheimer's disease. *Alzheimers Dement* **11**, 561-578.
- [113] Santos LBd, Corrêa Jr EA, Oliveira Jr ON, Amancio DR, Mansur LL, Aluísio SM (2017) Enriching complex networks with word embeddings for detecting mild cognitive impairment from speech transcripts. *arXiv preprint arXiv:1704.08088*.
- [114] Landauer TK, Dumais ST (1997) A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychol Rev* **104**, 211.
- [115] Hoffman P, Meteyard L, Patterson K (2014) Broadly speaking: vocabulary in semantic dementia shifts towards general, semantically diverse words. *Cortex* **55**, 30-42.
- [116] Gayraud F, Lee H-R, Barkat-Defradas M (2011) Syntactic and lexical context of pauses and hesitations in the discourse of Alzheimer patients and healthy elderly subjects. *Clin Linguist Phon* **25**, 198-209.
- [117] Cannizzaro MS, Coelho CA (2013) Analysis of narrative discourse structure as an ecologically relevant measure of executive function in adults. *J Psycholinguist Res* **42**, 527-549.
- [118] Perry RJ, Hodges JR (1999) Attention and executive deficits in Alzheimer's disease: A critical review. *Brain* **122**, 383-404.
- [119] Dijkstra K, Bourgeois MS, Allen RS, Burgio LD (2004) Conversational coherence: Discourse analysis of older adults with and without dementia. *J Neurolinguistics* **17**, 263-283.
- [120] Potkins D, Myint P, Bannister C, Tadros G, Chithramohan R, Swann A, O'Brien J, Fossey J, George E, Ballard C (2003) Language impairment in dementia: impact on symptoms and care needs in residential homes. *Int J Geriatr Psychiatry* **18**, 1002-1006.
- [121] Savundranayagam MY, Hummert ML, Montgomery RJ (2005) Investigating the effects of communication problems on caregiver burden. *J Gerontol B Psychol Sci Soc Sci* **60**, S48-S55.

- [122] Bourgeois M, Brush J, Douglas N, Khayum R, Rogalski E (2016) Will you still need me when I'm 64, or 84, or 104? The importance of speech-language pathologists in promoting the quality of life of aging adults in the United States into the future. *Semin Speech Lang* **37**, 185-200.
- [123] Aluísio S, Cunha A, Scarton C (2016) Evaluating progression of Alzheimer's disease by regression and classification methods in a narrative language test in Portuguese. in *International Conference on Computational Processing of the Portuguese Language* Springer, pp. 109-114.
- [124] Fraser KC, Meltzer JA, Graham NL, Leonard C, Hirst G, Black SE, Rochon E (2014) Automated classification of primary progressive aphasia subtypes from narrative speech transcripts. *Cortex* **55**, 43-60.
- [125] Groves-Wright K, Neils-Strunjas J, Burnett R, O'Neill MJ (2004) A comparison of verbal and written language in Alzheimer's disease. *J Commun Disord* **37**, 109-130.
- [126] Zraick RI, Carr PB, Gregg BA, Smith-Olinde L, Ghormley C, Hutton TJ (2011) Information units produced by persons with mild Alzheimer's disease during a picture description task. *J Med Speech Lang Pathol* **19**, 37-45.

	Spontaneous speech (SS)	Mild cognitive impairment (MCI)	Alzheimer's disease (AD)
	<b>Natural language</b> (Title/Abstract) All databases		
	(connected OR spontaneous) AND (speech OR language OR discourse)	(mild cognitive impairment OR MCI)	Alzheimer*
<b>Databases</b>	<b>Descriptors</b>		
PubMed (MeSH Major Topic) and MEDLINE (MeSH Subject Heading)	Natural language processing Speech Speech acoustics Speech discrimination tests Speech disorders Speech language pathology Speech production measurement Verbal behavior	Mild cognitive impairment	Alzheimer disease
Embase (Subject Headings)	Connected speech Connected speech abnormality Conversation Conversation analysis Discourse analysis Language Disorders/et (Etiology) Language disability/di (Diagnosis) Narrative Narrative analysis Natural Language Processing Oral communication Speech Speech analysis Speech and Language Speech and Language Assessment Speech and Language disability	Mild cognitive impairment	Alzheimer disease

	Speech articulation Speech discrimination Speech disorder Speech disorders* Spontaneous speech Spontaneous language production Verbal behavior		
PsycINFO/PsycARTICLES (Index Terms)	conversation conversation analysis discourse discourse Analysis narratives natural Language natural language processing oral communication speech speech and language disorder speech characteristics speech disorders speech pauses	Cognitive impairment	Alzheimer disease

Tableau 1. – Search terms and descriptors used for electronic database search



First Author [ref]	Year	AD	Controls	Language	Production	Syntactic	Lexical	Fluency	Semantic	Discourse	Other variables
Ahmed [3]	2013	9	9	English	Distortions Phonological paraphasias *Speech rate	Embeddings *MLU *Words in sentences *Syntactic errors *Nouns with determiners *Verbs with inflections	Closed class words *Pronouns *Verbs	Incomplete sentences *False starts *Filled pauses *Repaired sequences	*Total semantic units *Subjects *Objects *Actions	*Idea density *Efficiency	
Ahmed [46]	2013	18	18	English					Total semantic units *Component measures (nouns and verbs)	Idea density *Efficiency	
Ahmed [39]	2012	18	18	English	Distortions Phonological paraphasias *Speech rate	MLU *Syntactic errors *Words in sentences *Embeddings *Verbs with inflections *Nouns with determiners	Open/Closed class words *Pronouns *Verbs	Repaired sequences Incomplete sentences *Filled pauses *False starts			
Ash [86]	2007	20	10	English	*Speech rate	*MLU		*Word finding difficulty	*Quantity of essential material reported	*Accuracy of content Global connectedness *Maintenance of the theme *Coherence *Local connectedness	
Bayles [76]	2004	30	40	English	Number of words				*Information units	Global connectedness	
Bschor [96]	2001	41	40	German	Number of words				Feat: no difference on number of features (adjectives) *Objects *Locations *Actions Features		
Carlomagno [74]	2005	21	18	Italian	*Number of words *Words per minute				*Correct information units *Paraphasias (lexical)	*Information (crucial, non-crucial, wrong) *Correct information units per minute *Percent correct information units	Informative gestures *Miniturns
Chapman [78]	1998	10	10	English						Pragmatics (communicative)	Linguistic level (not different from)

										intentions) *Pragmatics (drawing inferences)	controls); includes hesitations, circumlocutions, semantic paraphasic errors, paragrammatic errors, neologisms. *Cognitive level (memory, attention, problem solving)
Chapman [79]	1995	12	12	English			Reference: pronoun to referent ratio and referential errors			*Frame of interpretation *Propositions supporting frame of interpretation *Propositions disrupting frame of interpretation *Structure of information (fewer narrative for AD)	
Cherney [71]	1993	10	10	English	*Rate of speech *Total utterances				Essential units	Elaborations Irrelevancies Off-topic comments Incorrect utterances *Efficiency ratio	
Choi [94]	2009	27	20	Japanese		Total number of sentences Total number of phrases Number of phrases per sentence			*Information units	*Number of main concepts *Narrative efficiency	
Croisile [90]	1996	22	24	French	Phonemic paraphasias *Total number of words produced	Grammatical errors Number of independent clauses Number of incomplete clauses Total number of clauses *Number of subordinate clauses	*Nouns *Verbs *Adjectives/adverbs *Functors	Revisions Repetitions *Word-finding difficulties	Subjects Places *Objects *Actions *Total information units Semantic substitutions	*Implausible details Modalizations *Words per information unit	
Cummings [41]	1985	30	70	English	*Melodic line	*Grammatical competence *Phrase length			*Information content		
De Lira [91]	2014	37	26	Portuguese	*Number of words				*Number of information units		
De Lira [44]	2011	60	61	Portuguese	Phonemic paraphasias	Subordinated sentences		*Word finding difficulties	Semantic substitutions		

						*Coordinated sentences *Reduced sentences		*Repetitions *Revisions			
Drummond [65]	2015	14	41	Portuguese	Total number of words Narrative time		Number of open-class words *Number of closed-class words *Referential cohesion (adequate use of pronouns, explicit referents and no repeated-name-penalty phenomenon)	*Repeated words		*Type of discourse narrative versus descriptive *Narrative structure: complete versus incomplete Index of discourse effectiveness *Total macropropositions *Total micropropositions *Irrelevant micropropositions	
Duong [4]	2005	46	53	French		Syntactic index: number of complex clauses/ total clauses) Transitional markers	Referential index			Lexical index (IU/total words) Micropropositions Right shifts Macropropositions Macropropositions per element Narrative structure (complete versus incomplete)	
Duong [84]	2003	5	27	French					*Repetition of expected ideas	*M/R ratio: modalizing words/referential words	
Ehrlich [45]	1997	16	16	English	Total number of words	Number of C-Units Clauses per C-Unit *Number of words per C-Unit	Pronouns without reference (anaphora) *Deictic terms (indefinites)	*Fragment index (Fluency: include false starts, filled pauses, immediate repetitions, incomplete clauses)	*Information Units	*Correct target propositions (semantic) *Efficiency index (target propositions/total words) Self-referential and extraneous statements (tangentiality)	
Forbes [82]	2002	22	22	English	Phonemic paraphasias *Melodic line (intonational color)	Grammatical form		*Word finding delays/difficulties	*Semantic paraphasias *Information conveyed	*Information content	*Error monitoring to word-finding delay/difficulty *Visual paraphasias
Forbes-McKay [66]	2005	30	240	English	Melodic line Articulation	Grammatical form Phrase length		*Word finding delays	*Pictorial themes *Semantic paraphasias	*Information content	*Error monitoring to word finding

					Phonological paraphasias						delays
Forbes-McKay [47]	2013	31	30	English	Articulation Phonological paraphasias	*Grammatical form Phrase length			*Semantic paraphasias *Pictorial themes	*Information content	*Visual paraphasias *Response to word finding delays *Error monitoring
Fraser [40]	2016	167	97	English	*Mel-frequency cepstral coefficients (MFCC): skewness (MFCC 1), skewness (MFCC 2), kurtosis (MFCC 5), kurtosis (VEL(MFCC 3)), skewness (MFCC 8), skewness (MFCC 12) *Phonation rate *Words Not-In-dictionary (including phonological paraphasias, distortions and unrecognizable words)	*Sentence fragments (ROOT->FRAG) *Adverbs with deictic function (ADVP->RB) *Prepositional phrase rate (PP) *Verb phrase rate *Verb phrase with auxiliary *Verb phrase with gerund *Verb phrase with gerund and prepositional phrase *Various verb phrase structures (VP->VBG_PP, VB->IN_S, VB->AUX_ADJP, VB->AUX, VB->VBD_NP)	Brunet's index (vocabulary richness) Type-Token Ratio (vocabulary richness) Moving-average type-token ratio (MATTR, vocabulary richness) *Pronoun to noun ratio *Number of noun phrases with a pronoun *Frequency (use of frequent words) *Verb frequency *Nouns *Word length *Noun phrases with a determiner (NP->DT_NN) *Honoré's statistic (vocabulary richness) *Inflected verbs *Verbs	*Average cosine distance (index of repetitive content) *Cosine cutoff: 0.5 (repetitive content) *Interjections (INTJ->UH)	*Key word: window, sink, cookie, curtain, counter, stool, mother *Info unit: window, curtain, cookie, sink, girl, girl's action, dish, stool, woman		
Giles [37]	1996	48	18	English	Total time				*Information units	*Information units per second	
Groves-Wright [125]	2004	28	14	English					*Information units	*Main concepts *Efficiency *Conciseness ratio	
Hier [42]	1985	26	15	English	Palitlatia *Total words	Fragments (missing word but semantically correct) Subordinate clauses Mean clause length Prepositional phrases *Errors in prepositions	*Unique words *Anomia index (nouns/(nouns+pronouns)) *Empty words	*Aposiopesis (abrupt termination of sentence)		*Relevant observations *Conciseness index Gratuitous comments	Perseverations

Kavé [93]	2018	35	35	Hebrew	*Total word number		*% of Content words of all words *Pronoun ratio *TTR *Mean word frequency % Verbs of all words % Verbs in PAAL form % Verbs in present tense % Prepositions of all words % Subordination markers		*Information units		
Kavé [43]	2016	20	20	Hebrew	Total number of words		Type-token ratio (all words) Type-token ratio (nouns) *% Content words of all words *% Nouns of all words *% Pronouns of all words *Mean frequency of all words *Mean frequency of nouns *Mean word length in letters				
Kavé [92]	2003	14	48	Hebrew		Total number of sentences Syntactic errors *Words per clause *Clause type (independent, dependent, incomplete) *Sentence type (simple declarative, head-complement, existential, relative, conjoined, and impersonal)	Proportion of verbs out of nouns and verbs Proportion of inflected verbs Proportion of verb roots out of verb forms *Pronoun use		Actions Places *Objects *Actions *Total units	*Semantic errors	*Circumlocutionary comments
Kong [89]	2016	13	20	English					*Main Concept Score	Number of Absent concepts *Number of Accurate and complete	

										concepts Number of Accurate but incomplete concepts Number of Inaccurate concepts *Number of Accurate and complete concepts per minute	
Lai [72]	2012	30	30	Chinese		Conceptual-epistemic Conceptual- evaluative Non-conceptual: inferential *Non-conceptual: contrastive *Non-conceptual: elaborative					
Lai [73]	2009	30	30	Chinese		Incomplete clauses Simple declarative sentence Head complement sentence Relative sentence Question sentence Syntactic errors: functors, tense confusion, ambiguous classifiers, unintelligible sentences *Independent clauses *Dependent clauses *Existential sentence *Conjoined sentence *Impersonal sentence *Unintelligible sentences	*Word-finding difficulties	*Actors *Places *Objects *Actions *Total units	*Incoherence	*Circumlocutionary comments (external)	
March [75]	2006	26	26	English			*Noun use *Person deixis *Spatial deixis				
McNamara [102]	1992	15	141	English							*Undetected errors *Reformulation *Lemma
Nicholas [101]	1985	19	30	English	Number of total words Literal paraphasias Verbal-	"Ands"	*Indefinite terms *Conjunctions *Deictic terms *Pronouns	*Repetitions	*Thematic elements (8) *Semantic paraphasias	*Empty phrases	Comments Judgments

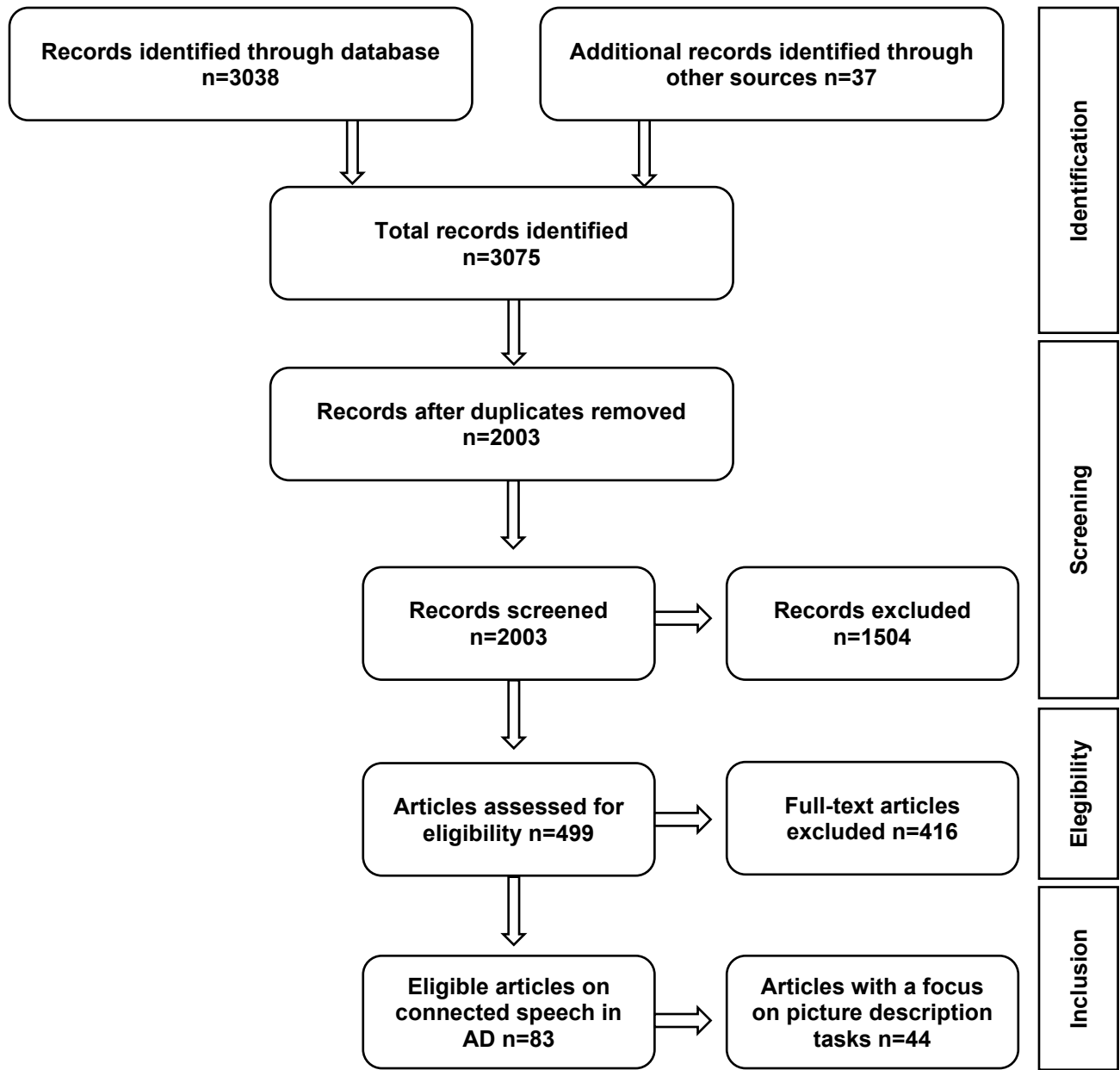
					phonological paraphasias Unrelated verbal paraphasias Neologisms		without antecedents				
St-Pierre [88]	2005	29	29	French						Related utterance *Relevant utterance *Irrelevant utterance	
Sajjadi [51]	2012	20	30	English	*Speech rate *Unit length *Combined phonological errors	Discourse markers Number of clauses per T-unit *Complete T-units *Elliptical T-Units *Abandoned T-units *Complex units *Number of arguments per verb *Verb agreement errors	Open-and closed-class word errors *Number of open-and closed-class words	Spontaneity *Hesitation markers *Editing breaks	Circumlocutions *Semantic errors *Information content	*Pictorial themes	
Shimada [87]	1998	23	17	Japanese					*Amount of information conveyed	Number of relevant and irrelevant descriptions (total) *Efficiency of description *Number of relevant descriptions	
Ska [6]	2005	46	53	French		*Syntactic index (complex phrases/total phrases) *Transition markers				Number of micropropositions Number of shifts in list of micropropositions *Lexical index (expected information/total) *Number of expected macropsositions *Narrative scheme elements	
Smith [80]	1989	18	18	English	Total syllables Syllables per minute Total words *Total time	Total number of clauses Elements per clause	Specifications per element Deictics per element Unique words Anomia index		Total content units	Elements Relevant observations Conciseness index *Idiosyncratic versus appropriate utterances *Content units per minute *Syllables per content units *Elements per clause	

Tomoeda [77]	1996	63	52	English	*Total words			*Aborted phrases *Ideational repetitions *Revisions	Circumlocutions *Information units	*Conciseness	Frustrations
Tomoeda [48]	1993	3	3	English	*Total words			Revisions Aborted phrases *Ideational repetitions	Circumlocutions *Information units	*Conciseness	Frustrations
Vuorinen [95]	2000	13	20	Finnish	Number of words per minute				*Eight central themes		
Zraick [126]	2011	8	21	English	*Syllables produced *Words spoken				Objects Localizations Actions Figures *Sum of total information units		

Tableau 2. – Connected speech variables in the 44 reviewed articles

(\*significant difference for AD patients)





(Image insérée ici)

Figure 1. – Flowchart of the review process

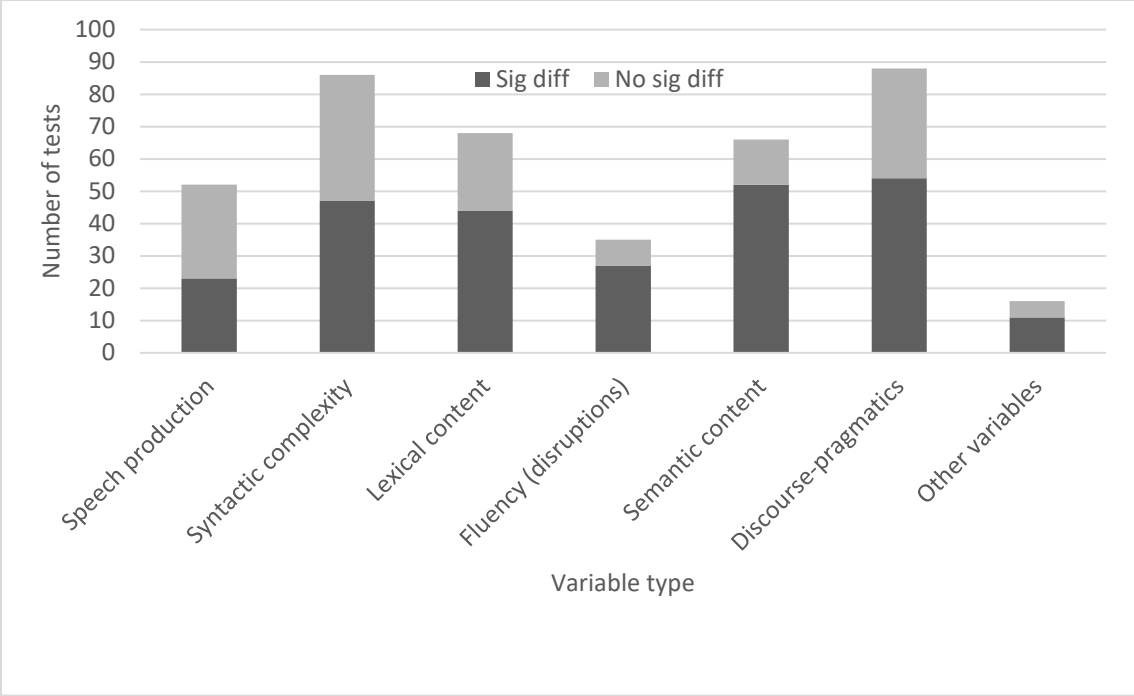


Figure 2. – Summary of tested variables by domain of connected speech

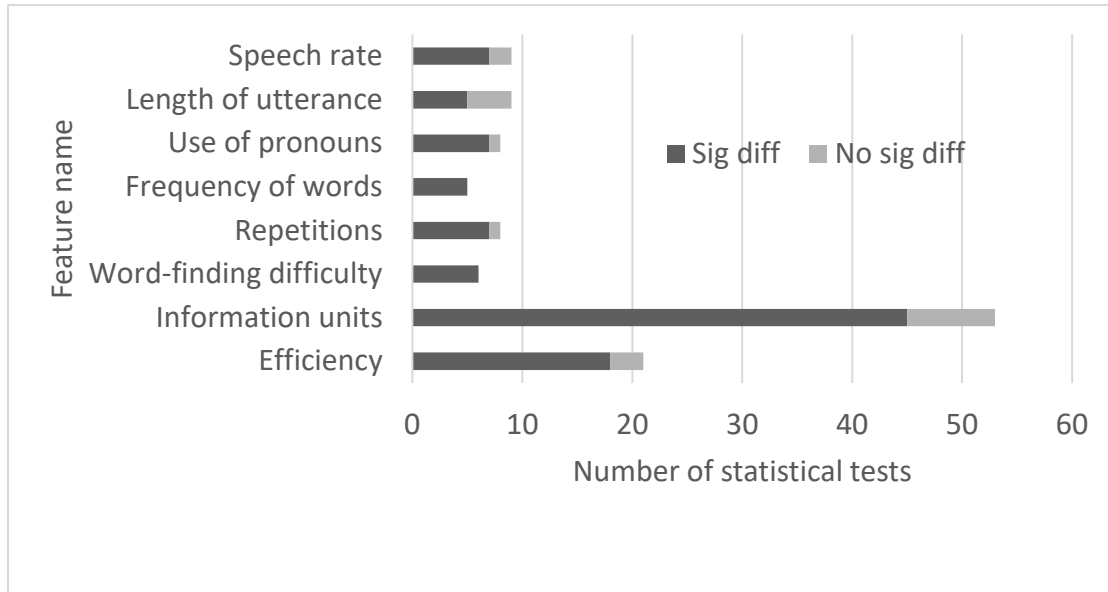


Figure 3. – Most commonly reported discriminant variables from AD picture description

## **Article 2 : Connected speech markers of amyloid burden in primary progressive aphasia**

Antoine Slegers<sup>a, b</sup>, Geneviève Chafouleas<sup>c</sup>, Maxime Montembeault<sup>d</sup>, Christophe Bedetti<sup>b</sup>, Ariane E. Welch<sup>d</sup>, Gil D. Rabinovici<sup>d</sup>, Philippe Langlais<sup>c</sup>, Maria Luisa Gorno-Tempini<sup>d</sup>, Simona Maria Brambati<sup>\* a b</sup>

- a. Department of Psychology, Université de Montréal, Pavillon Marie-Victorin C. P. 6128, succursale Centre-ville, H3C 3J7, Canada.
- b. Centre de Recherche de l'Institut Universitaire de Gériatrie de Montréal (CRIUGM), 4565 Queen-Mary, Montréal, H3W 1W5, Canada.
- c. Department of Computer Science and Operational Research, Université de Montréal, Pavillon André-Aisenstadt, DIRO, CP 6128 succ. Centre-ville, Montréal, H3C 3J7, Canada.
- d. Memory and Aging Center, University of California, San Francisco, Box 1207, 675 Nelson Rising Lane, Suite 190, San Francisco, CA 94143, USA.

\*Corresponding author:

Simona Maria Brambati [simona.maria.brambati@umontreal.ca](mailto:simona.maria.brambati@umontreal.ca)

Centre de Recherche de l'Institut Universitaire de Gériatrie de Montréal (CRIUGM), 4565 chemin Queen-Mary, Montréal, Canada, H3W 1W5

Tel: (514) 340-3540 #4147

Declarations of interest: none.

### **Référence complète :**

Slegers, Antoine, Geneviève Chafouleas, Maxime Montembeault, Christophe Bedetti, Ariane E. Welch, Gil D. Rabinovici, Philippe Langlais, Maria L. Gorno-Tempini, and Simona M. Brambati. "Connected speech markers of amyloid burden in primary progressive aphasia." *Cortex* 145 (2021): 160-168.

## **Abstract**

**INTRODUCTION:** Positron emission tomography (PET) amyloid imaging has become an important part of the diagnostic workup for patients with primary progressive aphasia (PPA) and uncertain underlying pathology. Here, we employ a semi-automated analysis of connected speech (CS) with a twofold objective. First, to determine if quantitative CS features can help select primary progressive aphasia (PPA) patients with a higher probability of a positive PET amyloid imaging result. Second, to examine the relevant group differences from a clinical perspective. **METHODS:** 117 CS samples from a well-characterised cohort of PPA patients who underwent PET amyloid imaging were collected. Expert consensus established PET amyloid status for each patient, and 40% of the sample was amyloid positive. **RESULTS:** leave-one-out cross-validation yields 77% classification accuracy (sensitivity: 74%, specificity: 79%). **DISCUSSION:** Our results confirm the potential of CS analysis as a screening tool. Discriminant CS features from lexical, syntactic, pragmatic, and semantic domains are discussed.

**Keywords:** primary progressive aphasia, biomarkers, connected speech, natural language processing, Alzheimer's disease, differential diagnosis, telemedicine

## **Highlights**

Predicting biomarkers of AD pathology in PPA patients

Natural language processing enables automatic feature extraction and classification

Distinct lexical, syntactic, pragmatic, and semantic features of PiB+ speech

Tool could be cost-effective for frontline screening in telemedicine

## 1. Introduction

Primary progressive aphasia (PPA) are a group of neurodegenerative clinical syndromes characterized by the insidious and gradual onset of predominantly language-related impairments, of which three clinical variants are recognized : the non-fluent (nfvPPA), semantic (svPPA), and logopenic (lvPPA) variants (Gorno-Tempini et al., 2011). Among all patients who meet the criteria for PPA, up to 55% present with amyloidosis (Villarejo-Galende et al., 2017). AD pathology is observed in the majority of lvPPA cases (Bergeron et al., 2018; Montembeault et al., 2018), a clinical syndrome characterized by profound word retrieval difficulties in spontaneous speech and sentence repetition deficits (Gorno-Tempini et al., 2011; McKhann et al., 2011). The prevalence of AD pathology in lvPPA cases is variable across different expert centers, ranging from 57% to 100% (Bergeron et al., 2018). The detection of amyloidosis in PPA samples represents an important and current issue for treatment and enrollment in clinical trials (Rabinovici et al., 2008; Rogalski et al., 2016).

Amyloid imaging positivity is established in vivo with positron emission tomography (PET) by the standardized uptake value ratios (SUVR) of a radioligand (e.g. Pittsburgh Compound B, Florbetapir, Florbetaben, or Flutemetamol). Appropriate use criteria state that amyloid imaging is indicated for atypical presentations of AD such as lvPPA (Dubois et al., 2010; Johnson et al., 2013) and in vivo amyloid PET measures have shown high correlation with post-mortem histopathological reports (Fink et al., 2020; Santos-Santos et al., 2018; Tan et al., 2017). However, amyloid PET scanning is 1) invasive, 2) remains prohibitively expensive for most care settings and is 3) incompatible with telemedicine. There is thus a pressing need for cost-effective methods to select patients for PET amyloid imaging and novel disease-modifying treatments.

Current research focuses on the non-invasive detection of biomarkers in individuals most likely to benefit from a full clinical and biomarker assessment (Villemagne et al., 2018). The potential of automated approaches to connected speech (CS) analysis in the search for early and reliable markers of neurodegeneration has been underlined (Battista et al., 2020; Beltrami et al., 2018; Clarke et al., 2020; Laske et al., 2015; Myszczyńska et al., 2020).

Previous studies on CS in PPA have shown that the three clinical variants of PPA present specific CS profiles expressing their characteristic cognitive and speech impairments (Ash et al., 2013). In patients with lvPPA, CS is marked by false starts, hesitations, filled pauses and repairs, as well as a different rate of verbs, nouns and pronouns, syntactic errors and embeddings compared to svPPA or to nfvPPA groups (Ash et al., 2013; Boschi et al., 2017; Wilson et al., 2010). The fact that lvPPA patients produce different CS features from other PPA variants and that they are disproportionately more likely to be amyloid-positive raises the possibility that CS could be used to screen for amyloid positivity. However, it is still unknown whether such a classification is possible and, if so, which CS feature combination enables accurate screening of PPA patients for amyloidosis without first relying on a syndromic variant classification, a complex diagnostic procedure that requires access to clinical expertise, batteries of cognitive/language tests and complementary instrumental measures that may not be available in all clinical and research settings.

The general hypothesis of the study is that CS features can discriminate PPA with or without amyloidosis and that the most discriminant CS features would be related to the core problem affecting spontaneous speech in these patients, namely word retrieval difficulties, i.e. (Gorno-Tempini et al., 2011; McKhann et al., 2011). To address this issue, we investigated 117 transcripts of picture descriptions by a cohort of PPA patients who also underwent carbon 11-labeled Pittsburgh Compound-B PET (PiB). From the transcripts, we extracted a large set of features



covering lexical, syntactic, semantic, and pragmatic aspects of CS. Using the currently growing approach in the field (Clarke et al., 2020), we combined CS features available in advanced natural language processing tools with other selected CS features based on recent clinical evidence on patients with PPA and/or AD pathology (Boschi et al., 2017; Fraser et al., 2014; Kamath et al., 2020; Slegers, Filiou, et al., 2018). First, we based a classification experiment on this set of features to identify which patients have PET-confirmed amyloid deposits and plaques. Second, we identified the set of features that differentiate PiB+ from PiB- transcripts at the group level.

## **2. Methods**

### **2.1 PPA participant database**

Data used in the preparation of this article were obtained from the UCSF Memory and Aging Center database. We retrospectively identified all participants who received a diagnosis of PPA, according to current diagnostic criteria (Gorno-Tempini et al., 2011). To be included in the study, participants must have a CS sample from the picture description task from the Western Aphasia Battery (Kertesz, 1982) and carbon 11-labeled Pittsburgh Compound-B PET imaging (PiB). The general exclusion criteria for all participants were as follows: native language other than English, developmental learning disabilities, history of a psychiatric disorder, history of traumatic brain injury, and uncorrected hearing or vision problems.

All participants provided written informed consent, which was approved by the Internal Review Board of the University of California, San Francisco.

### **2.2 Amyloid imaging: positron emission tomography with Pittsburgh Compound-B (PiB)**

Carbon 11-labeled Pittsburgh Compound-B PET (PiB) was performed at Lawrence Berkeley National Laboratory as previously described (Villeneuve et al., 2015). Native space SUVRs were created by normalizing mean images (at 50-70 minutes postinjection) by mean activity in cerebellum gray matter. Visual reads of native space SUVR images were performed by experienced investigators blinded to clinical data using published criteria (Rabinovici et al., 2011). Visual inspection based on these criteria has been validated previously as reproducible and reliable estimate of increased tracer uptake when compared with quantitative analysis. Every patient was thus assigned an amyloid positive (PiB+) or amyloid negative (PiB-) status.

### **2.3 Cognitive and sociodemographic data**

As part of their assessment, all patients underwent a multidisciplinary evaluation including a neurological examination, neuropsychological assessment, caregiver interview, screening laboratory tests, and high-resolution structural T1-weighted magnetic resonance imaging (MRI). Neuropsychological assessment included a previously described battery of standardized tests assessing memory, visuospatial abilities, and executive functions, global cognition and functional status (Gorno-Tempini et al., 2004).

### **2.4 Connected speech sample**

A total of 117 CS transcripts were included in the study, including 47 PiB+ and 70 PiB- PPA patients. Table 1 shows the characterization of the sample by amyloid status (PiB).

[insert Table 1 about here]

CS samples were obtained from a total of 113 PPA patients. We included two transcripts for four patients who produced picture descriptions on two separate visits. The present study aims at

predicting the amyloid status (PiB). We purposefully bypass clinical diagnosis in our classification experiment. However, for sample characterization, we report that the PPA group included 20 svPPA with predominantly left anterior lobe atrophy (svPPA-l; n=20), 12 svPPA with predominantly right anterior lobe atrophy (svPPA-r; n=12), 38 nvPPA, 43 lvPPA.

CS was elicited through the “picnic scene” picture description task from the Western Aphasia Battery (Kertesz, 1982). Patients were instructed the following: “take a look at this picture, tell me what you see, and try to talk in sentences”. The audio was imported into Audacity (<http://audacity.sourceforge.net>) and manually transcribed in English orthography. Unintelligible words were removed in preprocessing. Previous reports have shown highly reliable transcriptions using similar guidelines (Wilson et al., 2010).

## **2.5 CS feature extraction**

We contribute to the growing trend in CS research (Clarke et al., 2020) with a hybrid approach similar to Fraser et al. (2014); Garrard et al. (2014). We first extract a large array of features focusing exclusively on variables that can be computed automatically with natural language processing tools and add features from experimental (Ash et al., 2013; Fraser et al., 2014; Wilson et al., 2010), review (Auclair-Ouellet, 2015; Boschi et al., 2017; Slegers, Filiou, et al., 2018) and meta-analytic (Kamath et al., 2020) studies comparing groups of PPA subtypes to one another and/or amnesic AD to controls. We then perform a priori feature selection. This framework is common in studies deploying AI on language samples of AD patients and ensures that the model includes features describing clinically relevant CS characteristics (cf. de la Fuente Garcia et al. (2020b) for a systematic review).

Text pre-processing was performed using spaCy (<https://spacy.io/>) version 2.3.2 and the “en\_core\_web\_lg” model with standard pre-processing steps of tokenization, Named Entity Recognition (NER), Part of Speech tagging (POS) and dependency parsing (DEP). Variables computed with spaCy covered many lexical and syntactic features typical for the analysis of CS in neurodegenerative diseases. We supplemented psycholinguistic, semantic, and pragmatic features that were found to be crucial for the study of patients with possible amyloid pathology (or more generally PPA) and that are not already implemented in standard NLP libraries. In total, 261 features were computed: lexical=53; syntactic=157; semantic=37; pragmatic=14. Table A1 provides the list of extracted features with their definition and operationalisation. Analysis code is available (<https://osf.io/uf3zb/>).

### *2.5.1 Lexical features*

Part-of-speech tags (POS) were extracted. We examined accuracy of the tags within a predefined (random seed set at 42) random 10% subset of transcripts. Upon close inspection of the spacy POS-tags, we identified two anomalies that required correction. First, 70% of gerunds (i.e. fishing, sailing, drinking, swimming, flying, reading) were wrongfully tagged as “nouns” and were manually changed to “verbs”. Second, 2.4% of the hesitation markers “uhm” were wrongfully tagged and was manually assigned an “interjection” tag.

Type-token ratio (TTR), Honoré (R) and Brunet (W) measures were also extracted as in the study of Bucks et al. (2000). Brunet’s W is an inverted scale: a higher Brunet’s W indicates lower lexical diversity.

The absolute number of inflected verbs and their proportion relative to sample length were computed. Multiple psycholinguistic databases were queried for ratings of imageability and

familiarity (Scott et al., 2019), concreteness (Brysbaert et al., 2014), age of acquisition (Kuperman et al., 2012), frequency (Brysbaert et New, 2009), valence, arousal, and dominance (Warriner et al., 2013). Averages of these psycholinguistic variables were computed for 1) all words and for 2) nouns, 3) verbs, 4) adverbs, and 5) adjectives separately.

### 2.5.2 Syntactic features

Universal dependency<sup>1</sup> relationships (DEP) were extracted and proportions were established relative to both the number of words and the number of sentence roots in the transcript. The number of children, i.e. the direct dependents of a word, which are connected to it by a single arc to the left or to the right in the dependency tree, was averaged. Mean sentence length and “noun chunks” were extracted using spaCy’s *sentence* and *noun\_chunk* boundaries, respectively.

### 2.5.3 Semantic features

Information content units (ICU) were computed based on the list proposed by Jensen et al. (2006). The ICU list includes 36 elements, which denote subjects, objects, places, and actions in the picture. ICUs are extracted as 36 separate Boolean features i.e. as they were mentioned (True) or not (False) in the patient’s description. The total number of ICUs mentioned was also tallied.

### 2.5.4 Pragmatic features

Features for pragmatic elements of speech were inspired by the findings of Garrard et al. (2014); Rentoumi et al. (2014) were counted as follows, where words in brackets denote one instance of the pragmatic event: filled pauses (“uhm”), word-finding difficulties (“know”, “remember”, “unable”), deictic words (“this”, “that”, “here”), and uncertainty (“think”, “look”, “like”, “kind”,

---

<sup>1</sup> <https://universaldependencies.org/>

“see”, “maybe”, “can”, “something”). We implemented a measure of semantic idea density by computing the average cosine distance between all pairwise combinations of words within a sliding window (Ivensky, 2019). The word embeddings to average cosine distances were extracted from the same spacy “*en\_core\_web\_lg*” model that supported grammatical dependency and POS tagging. Sliding windows of 3, 10, 20, and 50 words with a step size of half the window length were implemented, outputting four idea density scores per participant.

## **2.6 Classification and significant group differences**

### *2.6.1 Classification with leave-one-out cross-validation*

In total, 261 features were extracted (Table A1, appendix). To reduce dimensionality of the classification model, feature selection is performed within each of the 117 folds of the cross-validation scheme. For our first objective of classification, feature selection operates as follows: 1) perform Welch’s t-tests across groups; 2) correct the associated *P values* for multiple comparisons by the Benjamini-Hochberg method for False Discovery Rate (FDR) at .05; 3) subset only significant features at  $p < .05$  after correction for FDR; 4) starting with the largest Welch’s coefficient, enter the features in the model in a stepwise manner at the condition that the added feature is not correlated  $> .75$  to a previously entered feature with a higher *t* statistic. The classifier model is built with the selected features and then used to predict the unseen, “left-out” sample. This procedure is repeated within each fold of the leave-one-out cross-validation. As no dedicated test set is established for final model evaluation, we report performance metrics (accuracy, sensitivity, specificity) as averages of the leave-one-out cross-validation on the whole dataset (117 folds).

Data analysis was performed using Python with scikit-learn (version 0.21.2) implementations of two classifiers: Logistic Regression with “liblinear” solver and Support Vector Machine (SVM)

with linear kernel. The classifiers and kernels were chosen for three reasons. First, using two classifiers demonstrates robustness and stability of classification accuracy (Fraser, Lundholm Fors, et al., 2019), i.e. ensuring that results are not specific to a single classifier. Second, linear kernels yield readily interpretable results (Fraser, Linz, et al., 2019). Third, linear kernels tend to perform well when the number of features is relatively high compared to the number of data points (Hsu et al., 2003), which is arguably the case in this experiment.

### *2.6.2 Significant group differences*

For our second objective of discussing group-level differences, we considered the whole sample instead of selecting features within each fold as in section 2.6.1. We identify significant group differences with a similar procedure: a series of Welch's tests corrected for multiple comparisons with the Benjamini-Hochberg method for False Discovery Rate (FDR) at .05 and stepwise removal of multicollinearity, i.e. removal of features when correlated  $>.75$  with any feature with a higher  $t$  statistic.

## **3. Results**

### *3.1 Classification of amyloid burden (PiB) from CS features*

We selected a model that boosted sensitivity (detect more of the PiB+ cases at the cost of more false positives) by adding 50% weight to the PiB+ class. The purpose of the a posteriori modification in class weights is to facilitate comparison to other biomarkers of AD pathology. SVM returned an average classification accuracy of 77% with a sensitivity of 74% and specificity 79%. As Table 2 shows, Logistic Regression returned similar, but slightly weaker results.

[insert Table 2 about here]

### 3.2 Group differences: discriminant CS features for PiB+ PPA patients

Figure 1 shows distributions of the significant features that differentiate between PiB+ and PiB- groups on the whole dataset. These twelve features form an array that spans multiple linguistic domains: eight lexical features, one syntactic feature, one semantic feature, and two pragmatic features. In the lexical domain, four features were lower in the PiB+ group: average imageability of all words, average imageability of nouns, number of nouns per sentence root, and average concreteness of verbs. Conversely, adjective dominance and frequency of verbs and adjectives, as well as Brunet's W index were higher in the PiB+ group. In the syntactic domain, PiB+ patients had a lower ratio of the average number of children to the total number of sentence roots. In the semantic domain, PiB+ patients were less likely to precisely name the drink being poured by the woman in the picture. In the pragmatic domain, PiB+ patients produced more words denoting uncertainty (e.g. "know", "remember", "looks like") relative to their total output in number of words (proportion of uncertainties). The PiB+ group also exhibited lower semantic idea density. To ensure that the difference in years of education between the groups does not skew results, post-hoc regression analyses were carried out and showed that all variables retained a significant association to PiB status beyond the effect of education. Figure 1 and Table 3 summarize these findings.

[insert Figure 1 about here]

[insert Table 3 about here]

## 4. Discussion



The present study confirms our hypotheses by providing two main results. First, we showed that a short, non-invasive, and inexpensive sample of connected speech (CS) can identify PiB+ patients with 77% accuracy, sensitivity of 74% and specificity of 79%, confirming the potential of CS in detecting the presence of amyloid in the brain of patients with PPA. Although methodological and sampling differences preclude a direct comparison of results with previous research, our analysis compares favorably to other approaches to identification of underlying pathophysiology such as clinical diagnosis and cognitive profiling (sensitivity: 78%; specificity: 80%) (Giannini et al., 2017) and plasma biomarkers (sensitivity: 78%; specificity: 77%) (Ashton et al., 2019).

Second, we demonstrated that an array of lexical, semantic, and pragmatic CS features, mostly attributable to word retrieval difficulties, play a central role in determining the CS profile characteristic of PiB+ patients. The set of discriminant features indicates that PiB+ patients present lower idea density, struggle to produce concrete nouns (imageability, frequency and concreteness), and that the nouns they manage to access are more abstract and are interspersed among many other words (nouns proportion), which tend to be less varied on average (Brunet W index). In addition, PiB+ PPA patients present more uncertainties (e.g. think, maybe, looks like, etc.), which likely reflect CS attempts at circumventing word retrieval problems. Thus, this aspect of the laborious speech production of PiB+ PPA patients can be captured by analyzing lexical features of CS transcripts as in Garrard et al. (2014); Rentoumi et al. (2014). The groups did not differ in the number of word-finding pauses, filled pauses, or other disruptions to fluency typically associated with word retrieval difficulties (Ash et al., 2013; Wilson et al., 2010). This latter set of features may not be sensitive markers since all PPA patients experience - albeit for different pathophysiological and neuropsychological reasons - a significant degree of pragmatic difficulties and disruptions to fluency. Beyond word retrieval difficulties, these features could reflect semantic

processing struggles, as in svPPA, or hesitation in grammatical construction, as in nfPPA. The pragmatic features included in the analysis were crafted from observations garnered in related CS studies of svPPA (Garrard et al., 2014) and multifactorial AD (Rentoumi et al., 2014). These features are more subjective than the remainder of the feature set.

Future studies should include more automatic and fewer task-dependent features describing the characteristics (filled and unfilled) and durations of the pauses. In addition, the inclusion of features related to distortions and phonological paraphasia, which are known to occur in lvPPA, should also be considered in future studies to better capture word retrieval difficulties in PiB+ PPA (Wilson et al., 2010). The inclusion of more sensitive variables could help improve classification accuracy in future studies, considering that the present experiment achieves 77%, from 60% chance accuracy.

So far, the set of discussed features can be traced to word retrieval difficulties previously described in PiB+ PPA patients but can also be observable in patients with semantic deterioration (PiB-svPPA patients). Interestingly, the analysis also revealed the presence of a discriminant syntactic feature that is not shared by svPPA patients, namely a higher ratio of the number of average left “children” to the total number of sentence roots. Short sentences as a marker of PiB+ speech could appear surprising considering they do not usually present with agrammatism. However, consistent with the absence of agrammatism, PiB+ patients do not produce fewer verbs. In this framework, the lower ratio of the average number of children to the total number of sentence roots suggests that PiB+ patients start many sentences with root verbs but that these have few dependents. This simplification of syntax in the absence of agrammatism is likely to reflect verbal working memory deficits that are well-documented in PiB+ PPA but also described in AD patients. The pattern of CS features is generally compatible with the neuroanatomical correlates of lexical retrieval and verbal working memory deficits, namely atrophy in the posterior superior and middle temporal gyri

and inferior parietal lobule, and damage to the posterior inferior longitudinal fasciculus, anterior inferior longitudinal fasciculus, uncinate fasciculus, and superior longitudinal fasciculus (Montembeault et al., 2018).

In addition, PiB+ PPA patients also displayed higher average dominance of adjectives that is driven by the presence of emotional words such as “happy”, “good”, “strong”, “active”, and “nice”. PiB+ patients tended to include more vivid adjectives related to characters’ feelings and general atmosphere of the picnic scene. This finding is compatible with the better processing of emotional concepts by left anterior insular regions in lvPPA compared to nfPPA patients (Piguet et al., 2015). Finally, PiB+ patients were significantly less likely to identify the drink being poured by the woman in the picture. Indeed, 60% of the PiB- group specified what the drink could be, whereas only 26% of the PiB+ group did. Patients of the latter group often substituted the information with uncertainty, e.g. “*not sure what she is drinking*” (not specifying the drink) instead of “*she may be pouring coffee or tea*” (specifying the drink). A drawback of this marker is its specificity to the picnic picture description task.

## **5. Conclusion**

For patients presenting with language symptoms suggestive of PPA, a full clinical examination is indicated before scanning per current recommendations on clinical amyloid imaging (Laforce et al., 2016; Witte et al., 2015). When extensive testing is not feasible or delayed, an advantageous approach could be the collection of a short sample of CS. We showed that CS analysis has good sensitivity and specificity to AD pathology.

### *Acknowledgments*

We would like to thank all study participants and their families for their generous support to our research.

#### *Declaration of interest*

Role of the funding source: The funders were not involved in the study design, collection, analysis or interpretation of data, nor were they involved in writing the paper or the decision to submit this article for publication.

To our knowledge, no part of the study analyses was pre-registered in a time-stamped, institutional registry prior to the research being conducted. Here, we report how we determined our sample size (maximum data available), all data exclusions (N/A), all inclusion/exclusion criteria (cf. Methods section), whether inclusion/exclusion criteria were established prior to data analysis (yes), all manipulations, and all measures in the study. As speech data cannot be fully anonymized, the raw data cannot be made publicly available. Dr Gorno-Tempini is responsible for granting access to raw data.

#### *Financial support*

The study was supported by grants from the Alfonso Martin Escudero Foundation, National Institutes of Health (National Institute of Neurological Disorders and Stroke [grant R01 NS050915] and National Institute on Aging [grant numbers P50 AG03006, P50 AG023501, P01AG019724, R01 AG045611, R01 AG027859, and K24 DC015544-01]), grant DHS04-35516 from the State of California, grant 03-75271 DHS/ADP/ARCC from the Alzheimer's Disease Research Centre of California; Alzheimer's Association, Larry L. Hillblom Foundation, John Douglas French Alzheimer's Foundation, Koret Family Foundation, Consortium for Frontotemporal Dementia Research, Tau Consortium, McBean Family Foundation, Career Scientist Award from the US

Department of Veterans Affairs Clinical Sciences R&D Program, and Avid Radiopharmaceuticals. G.C. was supported by a UNIQUE-IVADO scholarship. A.S. was supported by an Alexander Graham Bell Canada Graduate Scholarship from the Natural Sciences and Engineering Research Council of Canada and S.M.B. received funding from the Société Alzheimer du Canada.

## REFERENCES

- Ash, S., Evans, E., O'shea, J., Powers, J., Boller, A., Weinberg, D., . . . Rascovsky, K. (2013). Differentiating primary progressive aphasia in a brief sample of connected speech. *Neurology*, *81*(4), 329-336.
- Ashton, N. J., Nevado-Holgado, A. J., Barber, I. S., Lynham, S., Gupta, V., Chatterjee, P., . . . Blennow, K. (2019). A plasma protein classifier for predicting amyloid burden for preclinical Alzheimer's disease. *Science advances*, *5*(2), eaau7220.
- Auclair-Ouellet, N. (2015). Inflectional morphology in primary progressive aphasia and Alzheimer's disease: A systematic review. *Journal of Neurolinguistics*, *34*, 41-64.
- Battista, P., Salvatore, C., Berlinger, M., Cerasa, A., & Castiglioni, I. (2020). Artificial Intelligence And Neuropsychological Measures: The Case Of Alzheimer's Disease. *Neuroscience & Biobehavioral Reviews*.
- Beltrami, D., Gagliardi, G., Rossini Favretti, R., Ghidoni, E., Tamburini, F., & Calzà, L. (2018). Speech analysis by natural language processing techniques: A possible tool for very early detection of cognitive decline? *Frontiers in aging neuroscience*, *10*, 369.
- Bergeron, D., Gorno-Tempini, M. L., Rabinovici, G. D., Santos-Santos, M. A., Seeley, W., Miller, B. L., . . . van Berckel, B. N. (2018). Prevalence of amyloid- $\beta$  pathology in distinct variants of primary progressive aphasia. *Annals of neurology*, *84*(5), 729-740.
- Boschi, V., Catricalà, E., Consonni, M., Chesi, C., Moro, A., & Cappa, S. F. (2017). Connected Speech in Neurodegenerative Language Disorders: A Review. *Frontiers in Psychology*, *8*.
- Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. [journal article]. *Behavior Research Methods*, *41*(4), 977-990. doi: 10.3758/brm.41.4.977
- Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior research methods*, *46*(3), 904-911.

Bucks, R. S., Singh, S., Cuerden, J. M., & Wilcock, G. K. (2000). Analysis of spontaneous, conversational speech in dementia of Alzheimer type: Evaluation of an objective technique for analysing lexical performance. *Aphasiology*, *14*(1), 71-91.

Clarke, N., Foltz, P., & Garrard, P. (2020). How to do things with (thousands of) words: Computational approaches to discourse analysis in Alzheimer's disease. *Cortex*.

de la Fuente Garcia, S., Ritchie, C., & Luz, S. J. J. o. A. s. D. (2020). Artificial Intelligence, Speech, and Language Processing Approaches to Monitoring Alzheimer's Disease: A Systematic Review. (Preprint), 1-27.

Dubois, B., Feldman, H. H., Jacova, C., Cummings, J. L., DeKosky, S. T., Barberger-Gateau, P., . . . Galasko, D. (2010). Revising the definition of Alzheimer's disease: a new lexicon. *The Lancet Neurology*, *9*(11), 1118-1127.

Fink, H. A., Hemmy, L. S., Linskens, E. J., Silverman, P. C., MacDonald, R., McCarten, J. R., . . . Miller, M. A. (2020). Diagnosis and Treatment of Clinical Alzheimer's-Type Dementia: A Systematic Review.

Fraser, K. C., Linz, N., Li, B., Fors, K. L., Rudzicz, F., König, A., . . . Kokkinakis, D. (2019). *Multilingual prediction of Alzheimer's disease through domain adaptation and concept-based language modelling*. Paper presented at the Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers).

Fraser, K. C., Lundholm Fors, K., Eckerström, M., Öhman, F., & Kokkinakis, D. J. F. i. a. n. (2019). Predicting MCI status from multimodal language data using cascaded classifiers. *11*, 205.

Fraser, K. C., Meltzer, J. A., Graham, N. L., Leonard, C., Hirst, G., Black, S. E., & Rochon, E. (2014). Automated classification of primary progressive aphasia subtypes from narrative speech transcripts. *Cortex*, *55*, 43-60.

Garrard, P., Rentoumi, V., Gesierich, B., Miller, B., & Gorno-Tempini, M. L. (2014). Machine learning approaches to diagnosis and laterality effects in semantic dementia discourse. *Cortex*, *55*, 122-129.

Giannini, L. A., Irwin, D. J., McMillan, C. T., Ash, S., Rascovsky, K., Wolk, D. A., . . . Grossman, M. (2017). Clinical marker for Alzheimer disease pathology in logopenic primary progressive aphasia. *Neurology*, *88*(24), 2276-2284.

Gorno-Tempini, M., Hillis, A., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. e., . . . Boeve, B. F. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, *76*(11), 1006-1014.

Gorno-Tempini, M. L., Dronkers, N. F., Rankin, K. P., Ogar, J. M., Phengrasamy, L., Rosen, H. J., . . . Miller, B. L. (2004). Cognition and anatomy in three variants of primary progressive aphasia. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, *55*(3), 335-346.

Hsu, C.-W., Chang, C.-C., & Lin, C.-J. (2003). A practical guide to support vector classification: Taipei.

Ivensky, I. (2019). *Automatic detection of semantic dementia from scene description*. M.Sc., Université de Montréal, Montréal.

Jensen, A. M., Chenery, H. J., & Copland, D. A. (2006). A comparison of picture description abilities in individuals with vascular subcortical lesions and Huntington's disease. *Journal of communication disorders*, *39*(1), 62-77.

Johnson, K. A., Minoshima, S., Bohnen, N. I., Donohoe, K. J., Foster, N. L., Herscovitch, P., . . . Hartley, D. M. (2013). Appropriate use criteria for amyloid PET: a report of the Amyloid Imaging Task Force, the Society of Nuclear Medicine and Molecular Imaging, and the Alzheimer's Association. *Journal of Nuclear Medicine*, *54*(3), 476-490.

Kamath, V., Sutherland, E. R., & Chaney, G.-A. (2020). A Meta-Analysis of Neuropsychological Functioning in the Logopenic Variant of Primary Progressive Aphasia: Comparison with the Semantic and Non-Fluent Variants. *Journal of the International Neuropsychological Society*, *26*(3), 322-330.

Kertesz, A. (1982). *Western aphasia battery test manual*: Psychological Corp.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior research methods*, *44*(4), 978-990.



- Laforce, R., Rosa-Neto, P., Soucy, J.-P., Rabinovici, G. D., Dubois, B., & Gauthier, S. (2016). Canadian consensus guidelines on use of amyloid imaging in Canada: update and future directions from the specialized task force on amyloid imaging in Canada. *Canadian Journal of Neurological Sciences*, 43(4), 503-512.
- Laske, C., Sohrabi, H. R., Frost, S. M., López-de-Ipiña, K., Garrard, P., Buscema, M., . . . Linnemann, C. (2015). Innovative diagnostic tools for early detection of Alzheimer's disease. *Alzheimer's & Dementia*, 11(5), 561-578.
- McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Kawas, C. H., . . . Mayeux, R. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & dementia*, 7(3), 263-269.
- Montembeault, M., Brambati, S. M., Gorno-Tempini, M. L., & Migliaccio, R. (2018). Clinical, anatomical, and pathological features in the three variants of primary progressive aphasia: a review. *Frontiers in neurology*, 9, 692.
- Myszczyńska, M. A., Ojamies, P. N., Lacoste, A. M. B., Neil, D., Saffari, A., Mead, R., . . . Ferraiuolo, L. (2020). Applications of machine learning to diagnosis and treatment of neurodegenerative diseases. *Nature Reviews Neurology*. doi: 10.1038/s41582-020-0377-8
- Piguet, O., Leyton, C. E., Gleeson, L. D., Hoon, C., & Hodges, J. R. J. J. o. A. s. D. (2015). Memory and emotion processing performance contributes to the diagnosis of non-semantic primary progressive aphasia syndromes. 44(2), 541-547.
- Rabinovici, G., Rosen, H., Alkalay, A., Kornak, J., Furst, A., Agarwal, N., . . . Karydas, A. (2011). Amyloid vs FDG-PET in the differential diagnosis of AD and FTLD. *Neurology*, 77(23), 2034-2042.
- Rabinovici, G. D., Jagust, W. J., Furst, A. J., Ogar, J. M., Racine, C. A., Mormino, E. C., . . . Miller, B. L. (2008). A $\beta$  amyloid and glucose metabolism in three variants of primary progressive aphasia. *Annals of neurology*, 64(4), 388-401.
- Rentoumi, V., Raoufian, L., Ahmed, S., de Jager, C. A., & Garrard, P. (2014). Features and machine learning classification of connected speech samples from patients with autopsy proven

Alzheimer's disease with and without additional vascular pathology. *Journal of Alzheimer's Disease*, 42(s3), S3-S17.

Rogalski, E., Sridhar, J., Rader, B., Martersteck, A., Chen, K., Cobia, D., . . . Mesulam, M.-M. (2016). Aphasic variant of Alzheimer disease: clinical, anatomic, and genetic features. *Neurology*, 87(13), 1337-1343.

Santos-Santos, M. A., Rabinovici, G. D., Iaccarino, L., Ayakta, N., Tammewar, G., Lobach, I., . . . Spinelli, E. (2018). Rates of amyloid imaging positivity in patients with primary progressive aphasia. *JAMA neurology*, 75(3), 342-352.

Scott, G. G., Keitel, A., Becirspahic, M., Yao, B., & Sereno, S. C. (2019). The Glasgow Norms: Ratings of 5,500 words on nine scales. *Behavior research methods*, 51(3), 1258-1270.

Slegers, A., Filiou, R., Montembeault, M., & Brambati, S. (2018). Connected Speech Features from Picture Description in Alzheimer's Disease: A Systematic Review. *Journal of Alzheimer's disease: JAD*, 65(2), 519.

Tan, R. H., Kril, J. J., Yang, Y., Tom, N., Hodges, J. R., Villemagne, V. L., . . . Ittner, L. M. (2017). Assessment of amyloid  $\beta$  in pathologically confirmed frontotemporal dementia syndromes. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 9, 10-20.

Villarejo-Galende, A., Llamas-Velasco, S., Gómez-Grande, A., Puertas-Martin, V., Contador, I., Sarandeses, P., . . . Ruiz-Solis, S. (2017). Amyloid pet in primary progressive aphasia: case series and systematic review of the literature. *Journal of neurology*, 264(1), 121-130.

Villemagne, V. L., Doré, V., Burnham, S. C., Masters, C. L., & Rowe, C. C. (2018). Imaging tau and amyloid- $\beta$  proteinopathies in Alzheimer disease and other conditions. *Nature Reviews Neurology*, 14(4), 225-236.

Villeneuve, S., Rabinovici, G. D., Cohn-Sheehy, B. I., Madison, C., Ayakta, N., Ghosh, P. M., . . . Marks, S. M. (2015). Existing Pittsburgh Compound-B positron emission tomography thresholds are too high: statistical and pathological evaluation. *Brain*, 138(7), 2020-2033.

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior research methods*, 45(4), 1191-1207.

Wilson, S. M., Henry, M. L., Besbris, M., Ogar, J. M., Dronkers, N. F., Jarrold, W., . . . Gorno-Tempini, M. L. (2010). Connected speech production in three variants of primary progressive aphasia. *Brain*, *133*(7), 2069-2088.

Witte, M. M., Foster, N. L., Fleisher, A. S., Williams, M. M., Quaid, K., Wasserman, M., . . . Levenson, J. L. (2015). Clinical use of amyloid-positron emission tomography neuroimaging: practical and bioethical considerations. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, *1*(3), 358-367.

<b>Sociodemographic information and severity</b>	<b>PiB negative N=70</b>	<b>PiB positive N=47</b>	<b>Welch's t-test</b>	<b>P value</b>
<b>Age</b>	65.83	63.71	-1.38	.17
<b>Handedness (% right-handed)</b>	77	77	0.07	.97
<b>Gender (% female)</b>	53	60	0.69	.49
<b>Education (years)</b>	15.72	16.91	2.05	.04
<b>MMSE</b>	24.03	22.28	-1.38	.17
<b>CDR</b>	0.59	0.6	0.14	.89
<b>CDR Box Score</b>	2.98	3.33	0.73	.46

Tableau 3. – Table 1: Average sociodemographic information and severity of impairment by group.

Welch's t-test with significance level, uncorrected. MMSE: Mini-Mental State Evaluation; CDR: Clinical Dementia Rating scale.

<b>Classifier</b>	<b>Accuracy</b>	<b>Sensitivity</b>	<b>Specificity</b>
<b>Logistic Regression</b>	0.74	0.72	0.74
<b>SVM</b>	<b>0.77</b>	<b>0.74</b>	<b>0.79</b>

Tableau 4. – Table 2: Averaged classification outcomes for leave-one-out cross-validation

<b>Connected feature</b>	<b>speech</b>	<b>PiB- mean</b>	<b>PiB+ mean</b>	<b>Welch's t-test</b>	<b>P value</b>	<b>corr. value</b>	<b>P</b>
<b>Imageability (all word classes)</b>		5.62	5.19	-5.05	.000002	.0004	
<b>Imageability (nouns)</b>		6.32	6.04	-4.59	.00002	.001	
<b>ICU: drink</b>		0.60	0.26	-3.95	.0001	.005	
<b>Dominance (adjectives)</b>		3.91	5.40	3.66	.0004	.01	
<b>Uncertainties (Proportion of)</b>		0.02	0.04	3.52	.0007	.01	
<b>Frequency (verbs)</b>		6.13	6.65	3.43	.0009	.01	
<b>Nouns / Number of roots</b>		2.61	1.73	-3.43	.0009	.01	
<b>Semantic idea density (window length=10)</b>		0.64	0.61	-3.41	.0009	.01	
<b>Frequency (adjectives)</b>		4.13	5.34	3.03	.003	.03	
<b>Number of left children / number of roots</b>		0.06	0.04	-3.02	.003	.03	
<b>Brunet's W index</b>		6.30	6.55	2.91	.004	.04	
<b>Concreteness (verbs)</b>		3.20	2.99	-2.81	.006	.045	

Tableau 5. – Table 3: Group averages of statistically significant features.

Corrected P values with Benjamini-Hochberg correction and False discovery rate of .05.

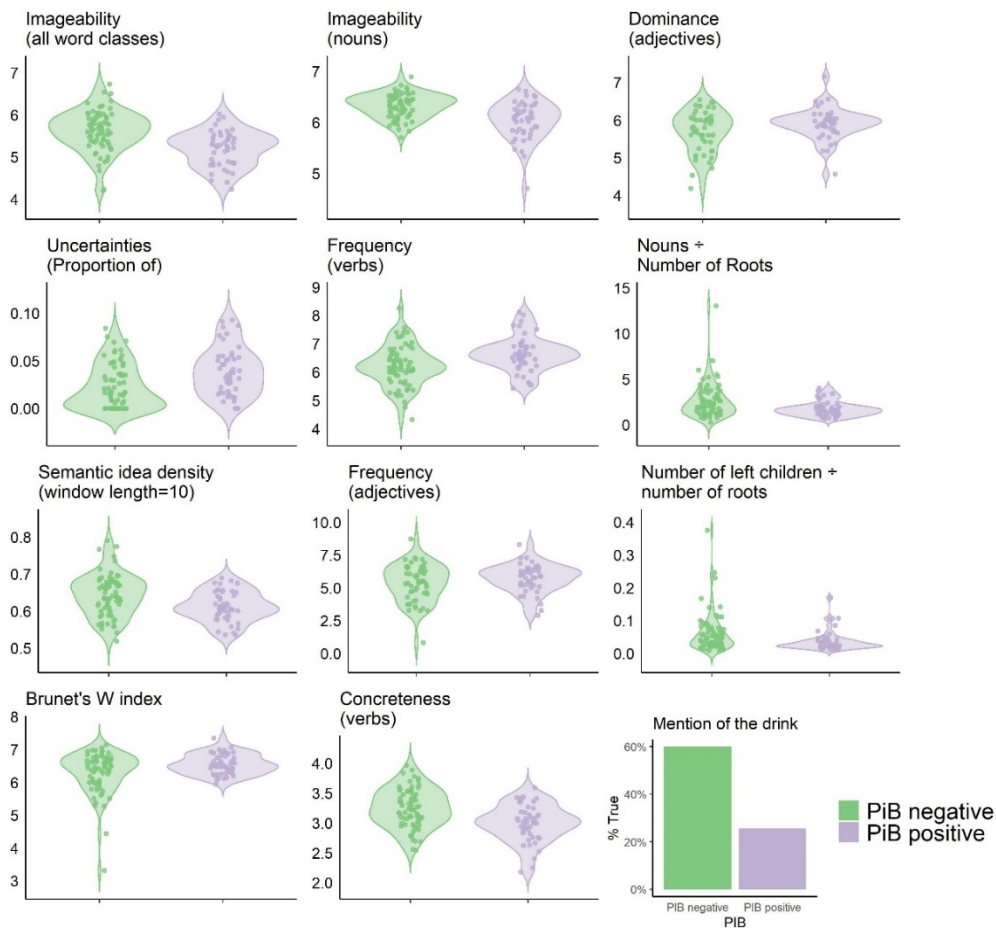


Figure 4. – Figure 1: Group distributions of statistically significant features.

All group differences are significant at  $p < .05$  after Benjamini-Hochberg correction for False discovery rate of .05. Most Y-axis labels were omitted to lighten the plot. Imageability, Dominance, Frequency, Concreteness are based on Likert scales from the psycholinguistic databases listed in section 2.5.1. Proportion of Uncertainties is the ratio of number of uncertainties in the transcript divided by the total number of words. Nouns / number of roots is the ratio of number of nouns divided by the number of sentence roots in the transcript. Number of left children / number of roots is the ratio of number of left children divided by number of sentence roots in the transcripts. Brunet's W index results from the equation in Table A1 and yields an inverted scale: a higher Brunet's W index corresponds to lower lexical diversity. Mention of the drink is plotted as the percentage of transcripts in which the drink was mentioned.

## ANNEX

Feature/feature family	Definition	Operationalisation	Number of features	Additional information
<b>Lexical features</b>				
Part of speech/grammatical class*: nouns, pronouns, verbs, adverbs, adjectives	Grammatical class, or part of speech (POS) of a given word	Number of instances of each grammatical class (nouns, pronouns, verbs, adverbs, adjectives) computed in the following 3 ways: in absolute numbers, relative to the total number of words and relative to total number of sentence roots. Computed with default implementation of spaCy POS tags.	5 variable x 3 computations = 15	Default values from spaCy are used, except for manually corrected anomalies reported in section 2.5.1
Deictic pronouns*	A term used in linguistic theory to subsume those features of language which refer directly to the personal, temporal or locational characteristics of the situation within which an utterance takes place, whose meaning is thus relative to that situation (Crystal, 2011)	The total number of instances of the following words: “This”, “That”, “Here” computed in the following 3 ways: in absolute numbers, relative to the total number of words and relative to total number of sentence roots.	1 variable x 3 computations = 3	March et al. (2006)
Type-Token Ratio (TTR)	Measure of lexical diversity relating total vocabulary to overall text length	Types/Tokens where Types is the number of different words used and Tokens is the text length expressed in number of words	1	A higher TTR higher indicates higher a lexical diversity
Honoré’s R	Measure of lexical diversity relating text length to total vocabulary and words spoken only once	$R = 100 * \log N / (1 - V1/V)$ where N is total text length in words, V is the total vocabulary and V1 is the number of words spoken only once	1	A higher Honoré statistic indicates higher lexical diversity
Brunet’s W index	Measure of lexical diversity relating text length to total vocabulary. Relatively independent of text length	$W = N * V^{-1.65}$ where N is total text length and V is total vocabulary used	1	A higher Brunet’s W indicates lower lexical diversity (inverted scale)
Familiarity†	The degree to which a word/part of speech is familiar	Subjective ratings of familiarity from the Glasgow norms (Scott et al., 2019) Average familiarity was computed for all words/nouns/verbs/adjectives in the transcript.	1 variable x 4 computations = 4	
Imageability†	The degree of effort involved in generating a mental image of the concept represented by the word	Subjective ratings of imageability from the Glasgow norms (Scott et al., 2019). Average imageability was computed for all words/nouns/verbs/adjectives in the transcript.	1 variable x 4 computations = 4	
Concreteness†	The degree to which the concept denoted by a word refers to a perceptible entity	Subjective ratings of concreteness from Brysbaert et al, 2014.	1 variable x 4 computations = 4	

Feature/feature family	Definition	Operationalisation	Number of features	Additional information
		Average concreteness was computed for all words/nouns/verbs/adjectives in the transcript.		
Age of Acquisition†	The subjective estimated age at which a word is learned. Age of acquisition (AoA) estimates are typically obtained by asking a group of participants at which age they learned various words	Subjective ratings of AoA from Kuperman et al. 2012. Average AoA was computed for all words/nouns/verbs/adjectives in the transcript.	1 variable x 4 computations = 4	
Frequency†	Rating on how often a word occurs in a language	Objective measure of word frequency from the SUBTLEX-US corpus (Brysbaert et New, 2009). Average frequency was computed for all words/nouns/verbs/adjectives in the transcript.	1 variable x 4 computations = 4	
Valence†	The degree of pleasantness of the emotions invoked by a word (happy, unhappy)	Subjective ratings of valence from Warriner et al. (2013). Average valence was computed for all words/nouns/verbs/adjectives in the transcript.	1 variable x 4 computations = 4	
Arousal†	The degree of intensity of emotion provoked by a word (excited, calm)	Subjective ratings of arousal from Warriner et al. (2013). Average arousal was computed for all words/nouns/verbs/adjectives in the transcript.	1 variable x 4 computations = 4	
Dominance†	The degree of control exerted by a word (weak/submissive, strong/dominant)	Subjective ratings of dominance from Warriner et al. (2013). Average dominance was computed for all words/nouns/verbs/adjectives in the transcript.	1 variable x 4 computations = 4	
<b>Syntactic features</b>				
Universal syntactic dependencies*	The directional dependency relation between a syntactic unit and the entities which constitute the relational structure of that unit (Crystal, 2011)	Number of instances of every syntactic dependency. Computed with spaCy dependencies (DEP). Computed in the following 3 ways: in absolute numbers, relative to the total number of words and relative to total number of sentence roots.	46 variables x 3 computations = 138	46 different dependencies occur in our corpus of transcripts <a href="https://universaldependencies.org/">https://universaldependencies.org/</a> <a href="https://spacy.io/usage/linguistic-features#dependency-parse">https://spacy.io/usage/linguistic-features#dependency-parse</a>
Left and right children*	The direct dependents of a word, which are connected to it by a single arc to the left or to the right in the dependency tree	Average number of left/right children of every word in the transcript. Computed with the spaCy default “n_left” and “n_right” commands. Computed in the following 3 ways: in absolute numbers, relative to the total number of words and relative to total number of sentence roots.	2 variables x 3 computations = 6	<a href="https://spacy.io/usage/linguistic-features#navigating">https://spacy.io/usage/linguistic-features#navigating</a>
Inflected verbs*	Verbs that are inflected in either tense, person, or number	Verbs in the transcript that do not correspond to their lemma as extracted by spaCy. Computed in the following 3 ways:	1 variable x 3 computations = 3	



Feature/feature family		Definition	Operationalisation	Number of features	Additional information
			in absolute numbers, relative to the total number of words and relative to total number of sentence roots.		
Subordinates*		Complex clauses involving subordination arise because a core or non-core dependent is realized as a clausal structure.	Total number of the four basic types of Universal Dependencies are concatenated from the spaCy syntactic dependencies: Clausal subjects (csubj). Clausal complements divided into those with obligatory control (xcomp) and those without (ccomp). Adverbial clause modifiers (advcl). Adnominal clause modifiers (acl) (with relative clauses as an important subtype in many languages). Computed in the following 3 ways: in absolute numbers, relative to the total number of words and relative to total number of sentence roots.	1 variable x 3 computations = 3	<a href="https://universaldependencies.org/u/overview/complex-syntax.html#subordination">https://universaldependencies.org/u/overview/complex-syntax.html#subordination</a>
Length of Sentence		Number of words in a sentence	Sentence boundaries are delimited by the default spaCy dependency parse	1	<a href="https://spacy.io/usage/linguistic-features#sbd">https://spacy.io/usage/linguistic-features#sbd</a>
Length and number of "noun chunks"*		Base phrases that have a noun as their head	Number of noun chunks and average noun chunk length by default spaCy implementation. Computed in the following 3 ways: in absolute numbers, relative to the total number of words and relative to total number of sentence roots.	2 variables x 3 computations = 6	<a href="https://spacy.io/usage/linguistic-features#noun-chunks">https://spacy.io/usage/linguistic-features#noun-chunks</a>
<b>Semantic features</b>					
36 separate information units (ICUs)		Separate subjects, places, objects, and actions that are represented in the picnic picture	Every separate ICU was computed as a Boolean feature (TRUE/FALSE) if it was mentioned (TRUE) or not (FALSE) by the patient.	36	List of ICUs from Jensen et al. (2006).
Total ICUs		Total number of ICUs that appear in the transcript	Sum of all TRUE ICUs	1	
<b>Pragmatic features</b>					
Uncertainties*		Words denoting uncertainty about the nature of some element of the picture	The number of instances of the following words: "think", "look", "like", "kind", "see", "maybe", "can", "something". Computed in the following 3 ways: in absolute numbers, relative to the total number of words and relative to total number of sentence roots.	1 variable x 3 computations = 3	Word list inspired by Garrard et al. (2014) and Rentoumi et al. (2014)
Filled pauses*		Hesitancy or pause in discourse marked by "uhm" or a variant ("hmm", "hum", etc.)	The number of instances of the following words: "uhm" All variants were normalized to "uhm".	1 variable x 3 computations = 3	

Feature/feature family	Definition	Operationalisation	Number of features	Additional information
		Computed in the following 3 ways: in absolute numbers, relative to the total number of words and relative to total number of sentence roots.		
Word-finding difficulties*	Words indicating a struggle to access a given word (tip-of-the-tongue phenomenon)	The number of instances of the following words: “know”, “remember”, “unable”. Computed in the following 3 ways: in absolute numbers, relative to the total number of words and relative to total number of sentence roots.	1 variable x 3 computations = 3	Word list inspired by Garrard et al. (2014) and Rentoumi et al. (2014)
Idea density (within sliding windows of length 3, 10, 20, 50)	The average density of (conceptually distinct) ideas conveyed within a sliding window.	Average cosine distance between all pairwise combinations of word embeddings within a sliding window. The word embeddings were extracted from the same spacy “en_core_web_lg” model that supported syntactic dependency and POS tagging. Within the sliding window, all pairwise cosine distances are averaged. Sliding windows of 3, 10, 20, and 50 words with a step size of half the window length were implemented, outputting four idea density scores per participant.	4	A conceptually similar procedure - with a different implementation - is presented in Ivensky (2019)
Efficiency	The ratio of transcript length to number of information units	Number of words/Total ICUs	1	

Tableau 6. – Table A1: list of all extracted features.

Asterisks (\*) mark features that were computed in absolute numbers, relative to the total number of words and relative to total number of sentence roots. Daggers (†) mark psycholinguistic features that were computed for all word categories, but also separately for nouns, for verbs, and for adjectives.

	PiB negative	PiB positive
<b>svPPA-l</b>	18 (90%)	2
<b>svPPA-r</b>	12 (92.3%)	1
<b>nfvPPA</b>	39 (97.5%)	1
<b>lvPPA</b>	1	43 (97.7%)

Tableau 7. – Table A2: distribution of clinical diagnoses by amyloid status

Feature	Selected within % of cross-validation folds
Imageability (nouns)	100,0
ICU: drink	100,0
Dominance (adjectives)	100,0
Frequency (verbs)	100,0
Semantic idea density (window length=10)	100,0
Frequency (adjectives)	100,0
Nouns / Number of roots	99,1
Uncertainties (Proportion of)	98,3
Imageability (all word classes)	97,4
Number of left children / number of roots	96,6
Brunet's W index	86,3
Concreteness (verbs)	47,0
Types* (absolute number of unique words)	18,8

Tableau 9. – Table A3: Features selected in at least 10% of folds during cross-validation

Note the almost perfect overlap with this table and the content of Table 3 in the manuscript. The features marked with an asterisk were excluded from group comparisons in Table 3 because of multicollinearity with other features but were nonetheless present in some folds during cross-validation.

## REFERENCES

- Brysbaert, M. et New, B. (2009, November 01). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English [journal article]. *Behavior Research Methods*, 41(4), 977-990. <https://doi.org/10.3758/brm.41.4.977>
- Crystal, D. (2011). *A dictionary of linguistics and phonetics* (vol. 30). John Wiley & Sons.
- Garrard, P., Rentoumi, V., Gesierich, B., Miller, B. et Gorno-Tempini, M. L. (2014). Machine learning approaches to diagnosis and laterality effects in semantic dementia discourse. *Cortex*, 55, 122-129.
- Ivensky, I. (2019). *Automatic detection of semantic dementia from scene description* [Université de Montréal].
- Jensen, A. M., Chenery, H. J. et Copland, D. A. (2006). A comparison of picture description abilities in individuals with vascular subcortical lesions and Huntington's disease. *Journal of Communication Disorders*, 39(1), 62-77.
- March, E. G., Wales, R. et Pattison, P. (2006). The uses of nouns and deixis in discourse production in Alzheimer's disease. *Journal of Neurolinguistics*, 19(4), 311-340.
- Rentoumi, V., Raoufian, L., Ahmed, S., de Jager, C. A. et Garrard, P. (2014). Features and machine learning classification of connected speech samples from patients with autopsy proven Alzheimer's disease with and without additional vascular pathology. *Journal of Alzheimer's Disease*, 42(s3), S3-S17.
- Scott, G. G., Keitel, A., Becirspahic, M., Yao, B. et Sereno, S. C. (2019). The Glasgow Norms: Ratings of 5,500 words on nine scales. *Behavior Research Methods*, 51(3), 1258-1270.
- Warriner, A. B., Kuperman, V. et Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, 45(4), 1191-1207.

## **Chapitre 3 – Discussion générale**

### **3.1 Sommaire et interprétation des résultats**

La présente thèse a été constituée de deux articles. Le premier visait à recenser systématiquement la littérature clinique existante au sujet des marqueurs discursifs de la MA typique en description d'image. Le second, basé sur le premier, consistait à repérer les meilleurs marqueurs discursifs d'une pathologie Alzheimer sous-jacente afin de tester le potentiel d'une analyse semi-automatisée à identifier la pathologie Alzheimer chez un groupe de patients atteints d'une APP. Dans la discussion qui suit, nous reviendrons sur les principaux résultats des deux articles et développerons leurs contributions scientifiques respectives. Nous présentons la contribution scientifique générale de la thèse, les limites des études, les pistes de recherche future, ainsi qu'une conclusion générale.

#### **3.1.1 Objectifs et résultats de l'Article 1**

L'objectif de l'article 1 était d'obtenir un portrait à la fois exhaustif, quantitatif et global des caractéristiques du discours suscité par description d'image qui permettent de différencier les patients MA de témoins en santé. Nous avons recensé une riche littérature et avons procédé à l'organisation des multiples variables testées dans six domaines linguistiques spécifiés et une catégorie mixte : production verbale, caractéristiques lexicales, structure syntaxique, fluidité, contenu sémantique, aspects pragmatiques et autres variables. La quantité importante d'études à recenser et leur lot de variables disparates (412 variables testées en 40 ans de recherche) nous a motivé à combiner des variables nominalement distinctes mais conceptuellement similaires, par exemple le ratio de pronoms et le ratio de noms/pronoms. La synthèse réalisée a permis de faire émerger un patron multidimensionnel de différences significatives par rapport aux personnes âgées en santé.

En bref, les résultats mettent en relief huit caractéristiques souvent étudiées : le débit de parole ralenti (*rate of speech*), la longueur réduite des énoncés (*length of utterance*), l'emploi accru de

pronoms (*use of pronouns*), la fréquence objective moyenne élevée (*frequency of words*), la plus grande répétition, la présence de difficultés d'accès lexical (*word-finding difficulty*), la réduction du contenu informatif du discours (*information units*) et une baisse de l'efficacité (*efficiency*). Ces caractéristiques ont été retenues pour discussion puisqu'elles étaient documentées dans au moins quatre études chacune et qu'elles présentaient une différence statistiquement significative dans une majorité d'études.

La mosaïque de changements langagiers résultante illustre bien le riche potentiel clinique d'un court échantillon de discours comme outil de dépistage. Nos résultats démontrent que plusieurs dimensions du discours sont affectées par la MA. De plus, lorsqu'on les considère individuellement, les variables significatives sont compatibles avec une ou plusieurs des perturbations cognitives observées chez les patients avec une MA sur le plan des fonctions exécutives (efficacité et densité idéatoire, répétitions), de la cognition sémantique (fréquence des mots employés, contenu informatif), de l'accès lexical (manque du mot et surutilisation de pronoms) et de la mémoire de travail (répétitions, simplification de la syntaxe). En somme, la revue systématique suggère qu'un court échantillon de discours peut donner une fenêtre sur le fonctionnement cognitif de la personne, ce qui sera confirmé entre temps par des études testant ce type d'hypothèse (Covatti Malcorra et al.; Kavé et Goral, 2018).

Les caractéristiques du discours les plus fréquemment indicatives d'une MA typique étaient majoritairement lexicosémantiques. Ceci est concordant avec deux déficits cognitifs bien documentés dans la MA, nommément les troubles d'accès lexical et les déficits sémantiques. Les troubles sémantiques peuvent être subtils mais précoces dans certains cas de maladie d'Alzheimer (Joubert et al., 2010; Joubert et al., 2017; Montembeault et al., 2017). Dans les études computationnelles, ce sont ces troubles sémantiques qui sont désignés comme les causes des



changements discursifs mesurés : pour une revue systématique, voir Martínez-Nicolás et al. (2021). Cependant, alors qu'ils sont typiquement objectivés dans des tâches de dénomination sous confrontation, les troubles d'accès lexical laissent également leur marque en description d'image (Kavé et Goral, 2016, 2018). Il semble donc que la contribution relative des troubles sémantiques par rapport aux troubles d'accès lexical dans la MA, qui n'est pas tranchée dans la littérature actuelle sur la MA (Joubert et al., 2017), ne l'est pas davantage dans l'étude des marqueurs discursifs.

Globalement, les changements discursifs qui caractérisent la MA sont pour la plupart assimilables à un des six des domaines langagiers identifiés. La majorité des variables recensées sont calculées manuellement par l'examineur. Ce constat est valable également dans d'autres tâches de discours, dans les études sur le TCL (Filiou et al., 2019), ainsi que dans le cas des APP (Boschi et al., 2017). La synthèse des preuves existantes effectuée dans l'article 1 suggère donc une perspective d'automatisation de l'analyse du discours.

### **3.1.2 Limites de l'article 1**

Nous ne pouvons exclure que l'ensemble des variables recensées et quantifiées en termes de tests statistiques significatifs est sujet aux biais de disponibilité et de publication. Effectivement, les variables les plus facilement extraites (biais de disponibilité) sont plus susceptibles de figurer dans les tests statistiques rapportés et donc d'être documentées dans la littérature. De plus, il n'est pas impossible que certains auteurs rapportent préférentiellement les tests statistiques effectués lorsque ces derniers sont significatifs (biais de publication), ce qui résulterait en une estimation trop optimiste du potentiel discriminant des variables recensées dans la revue systématique. À cet effet, une estimation des tailles d'effet moyennes aurait pu, sans corriger ces biais potentiels, permettre d'estimer la magnitude des différences de groupe rapportées et constituer une mesure plus

représentative de leur réel potentiel clinique. De plus, il est à noter que nous n'avons pas effectué d'analyse formelle de la qualité des études qui excèderait les recommandations PRISMA, par exemple les méthodes CONSORT (Bennett, 2005) ou TREND (Des Jarlais et al., 2004). Par définition, les études recueillies dans la revue systématique consistaient en un test d'hypothèse des différences statistiquement significatives entre un groupe clinique et un groupe de témoins en santé et ne constituaient donc pas des études de traitement/intervention. Ainsi, un petit nombre des études recensées auraient rempli ces standards de qualité puisque la majorité des devis expérimentaux n'étaient pas randomisés et étaient basés sur des critères d'inclusion/exclusion – dans les cas où ceux-ci étaient décrits - plus rudimentaires.

Par ailleurs, la méthode PRISMA employée pour la revue systématique ne prescrit pas comme nécessaire l'évaluation des articles à l'aveugle. Nous avons opté pour une stratégie de revue *indépendante* (mais non aveugle) d'inclusion/exclusion des études et de cueillette de données dans laquelle chacun des évaluateurs recensait dans un tableur les raisons pour lesquelles une étude était ou n'était pas sélectionnée. Les désaccords - dans les faits relativement rares - étaient résolus par consensus lors de communications régulières qui ont suivi la sélection initiale des articles et lors de l'ajout d'articles repêchés en bibliographie par un des évaluateurs. Nous n'avons toutefois pas quantifié le nombre et la nature de ces désaccords occasionnels.

La présence de conflits d'intérêt était décelée et l'étude exclue lorsque l'énoncé de ces conflits laissait planer un doute quant à un lien possible entre les bailleurs de fonds de l'étude et les données rapportées ou leur analyse. Ce type d'exclusion s'est révélé relativement simple à effectuer dans la mesure où la littérature scientifique publiée dans des revues réputées contre-indique habituellement ce type de conflit d'intérêt manifeste et exige des auteurs qu'ils en fournissent une description fidèle.

De plus, une étude de notre groupe (Filiou et al., 2019) a démontré que 38% des études recensées dans l'étude du discours en MA et TCL ne spécifiaient pas de critères diagnostiques et/ou de sévérité. Ces travaux ont montré un patron de variables légèrement différent selon la sévérité de la maladie. Plus spécifiquement, il semble que les mesures de production (par exemple, débit verbal), de fluidité (pauses et hésitations) et de sémantique (nombre d'éléments informatifs mentionnés) constituent les changements les plus précoces. Au stade léger-moderé, les perturbations syntaxiques (réduction de la longueur des énoncés) deviennent plus marquées. Dans l'article 1, 39% des études recensées avaient testé une forme ou une autre de l'effet de la sévérité de la MA sur les caractéristiques du discours. Une autre limite de l'article 1 réside donc en l'absence d'analyse complémentaire selon les degrés de sévérité, lacune qui paraît néanmoins partiellement comblée par notre contribution à des travaux se penchant spécifiquement sur la sévérité, à savoir Filiou et al. (2019).

Les résultats de la revue systématique de l'article 1 concernaient le *diagnostic clinique* de la MA typique, non nécessairement supporté par les biomarqueurs. À la seule exception de Ahmed, de Jager, et al. (2013), les études recensées ne validaient pas les changements mesurés par l'analyse du discours avec une confirmation (histopathologique ou par imagerie) de la présence de pathologie Alzheimer. Dans le cas des présentations atypiques de la MA et des APP (qui dépassaient le cadre de l'article 1), certaines études relient les caractéristiques du discours à l'imagerie anatomique (Ash et al., 2013; Wilson et al., 2010), mais aucune à notre connaissance n'établit de lien direct entre la présence/absence de pathologie Alzheimer confirmée et le profil de discours. On remarque donc une déconnexion entre la littérature sur le discours existante et les plus récents critères articulés autour des biomarqueurs (Dubois et al., 2014; McKhann et al., 2011; Morris et al., 2014).

### 3.1.3 Contributions de l'article 1

La portée des résultats de l'article 1 apparaît lorsqu'on les situe dans le contexte actuel de croissance exponentielle (de la Fuente Garcia et al., 2020a) de la recherche déployant l'IA dans l'étude du langage dans la MA (Martínez-Nicolás et al., 2021). Une nette majorité des études recensées dans l'étude 1 ne rapportent pas explicitement faire usage de techniques automatisées pour analyser le discours, ce qui reflète également l'état de la littérature sur le discours et le TCL (Filiou et al., 2019) et les APP (Boschi et al., 2017). Pourtant, des outils de TALN génériques récents et facilement accessibles ont montré leur potentiel d'intégration dans le paysage des (bio)marqueurs de neurodégénérescence supportés par l'IA, aux côtés de l'imagerie cérébrale, des échantillons sanguins et de la génomique (Myszczyńska et al., 2020). Cette intuition allait être confirmée dans les années suivant la publication de l'article 1. Une consultation de PubMed révèle que le volume d'articles contenant les mots-clés « natural language processing alzheimer » a triplé (passant de 32 à 86) dans les trois ans qui nous séparent actuellement de la publication de l'article 1.

Les outils de TALN disponibles en libre-accès ne sont pas conçus spécifiquement pour l'analyse du discours, et encore moins pour l'analyse du discours pathologique. Pour harnacher le potentiel de ces outils à l'étude d'échantillons cliniques, il faudra prévoir une supervision étroite pour s'assurer de la robustesse des résultats et au besoin ajouter les variables-clé. À cet effet, concernant la MA, on remarque que plusieurs variables discriminantes issues de l'article 1 ne sont pas extraites par les outils de TALN les plus courants, notamment les données psycholinguistiques et les unités d'information sémantique. Guidées par une expertise clinique résultant en une approche hybride, certaines variables importantes devront être greffées aux outils automatiques d'extraction de caractéristiques. L'explosion actuelle de la recherche dans le domaine marque donc un moment

charnière pour l'articulation des expertises (cliniques et computationnelles) dans le champ de recherche des marqueurs discursifs de la MA et d'autres maladies neurodégénératives.

D'ailleurs, certains commentaires critiques récents de l'analyse du discours dans les maladies neurodégénératives définissent trois axes qui font écho aux contributions de l'article 1 et à ses limites, qui sont appelées à être développées dans l'article 2 : a) l'importance d'une approche spécifique aux biomarqueurs pour les études computationnelles de classification (Holmlund et al., 2020) ; b) l'urgence de circonscrire des marqueurs fiables de la pathologie sous-jacente dans les cas de MA atypiques (Ruksenaite et al., 2021; Tippett, 2020) ; c) la primordialité de la connaissance du domaine (*domain knowledge*) dans les approches computationnelles (Balagopalan et al., 2021). En ce qui a trait au point a), des études récentes soulignaient au moment de débiter l'étude 2 que les patients avec une APP présentaient une pathologie Alzheimer sous-jacente dans une plus grande proportion que ce qui était admis auparavant (Bergeron et al., 2018; Villarejo-Galende et al., 2017). Le point b) est porté par un groupe de chercheurs en intervention spécialiste des APP qui appellent à un changement de paradigme dans la recherche portant sur les APP. Notamment, ils rapportent que des traitements modificateurs de la maladie ciblant notamment les tauopathies sont en cours et affirment que l'avenir de ces traitements réside dans une approche moléculaire au diagnostic des APP. Concernant le point c), l'article cité a démontré qu'une approche linguistique avec un nombre réduit de caractéristiques interprétables était au moins aussi performant qu'une approche entièrement computationnelle et reposant sur les toutes dernières avancées en TALN de données massives. De plus, des auteurs émergents ont également souligné l'importance de limiter (par la supervision d'un expert du domaine) le nombre de prédicteurs dans les modèles (Ostrand et al., 2021). Cette constatation a revêtu un caractère prémonitoire du développement de l'approche de sélection des variables, un trait distinctif de l'étude 2.

Dans l'orientation donnée à l'étude 2 telle que découlant des découvertes de l'étude 1, nous avons également identifié certains aspects pragmatiques et épistémologiques qui nous ont guidé vers l'abord d'un problème de classification différent (variante atypique plutôt que MA amnésique typique). Premièrement, il n'existait pas à l'époque de base de données équilibrées et de taille suffisante pour mener une expérience de classification novatrice avec une population MA vs témoins en santé. Des initiatives comme le ADDRESS challenge (Luz et al., 2020), qui proposent un jeu de données standardisé et équilibré, allaient voir le jour après nos efforts initiaux dans l'article 1. Également la disponibilité des données du Consortium Canadien en Neurodégénérescence liée au Vieillissement (CCNV) a été trop tardive pour nous permettre de les exploiter dans la présente thèse. Deuxièmement, l'identification d'une démence de type MA amnésique parmi un groupe de témoins en santé ne nous paraissait pas un problème qui cliniquement serait éclairci au premier abord par une tâche de description d'image, en particulier considérant que l'examen par imagerie amyloïde est restreint aux patients dont la présentation clinique laisse un doute quant à la pathologie sous-jacente. Troisièmement, l'étude 1 a montré un phénomène selon lequel plusieurs variables très similaires sont souvent testées, en particulier dans les études computationnelles. Il nous a donc semblé pertinent d'intégrer le dispositif de sélection de variables afin de préserver des variables discriminantes différentes les unes des autres et d'augmenter le pouvoir de prédiction de l'algorithme (Ostrand et al., 2021). Quatrièmement, l'étude 1 a mis en relief le potentiel de variables psycholinguistiques, dont le rôle prépondérant constitue une particularité de l'étude 2 dans le paysage actuel de l'analyse du discours dans les APP. En prenant en compte toutes ces considérations, une approche hybride computationnelle qui met de l'avant des connaissances cliniques raffinées issues de l'étude 1 s'est donc imposée.

En conclusion, l'article 1 a démontré qu'un diagnostic clinique de MA correspond à un profil de discours caractéristique que nous avons synthétisé, décrit et quantifié. Cet effort soulevait donc pour nous naturellement la question de l'existence d'un profil caractéristique - similaire ou distinct - dans une présentation langagière de la MA confirmée par imagerie amyloïde. Les objectifs et la méthodologie de l'article 2 ont donc été orientés par le potentiel d'une approche automatisée et supervisée de préciser la possibilité et la nature d'une expérience de classification par apprentissage automatique.

### **3.1.4 Objectifs et résultats de l'article 2**

Le dévoilement d'un résultat d'imagerie amyloïde positif a un impact mesurable sur le patient, ses proches et modifie souvent l'orientation des soins (Bensaïdane et al., 2016). L'article 2 visait donc à mobiliser les connaissances cumulées dans l'article 1, basées sur la comparaison de groupes de patients à des témoins en santé, vers un défi de classification plus représentatif d'une décision clinique où l'on cherche à établir l'étiologie de la symptomatologie du patient. Conséquemment, l'objectif de l'article 2 était double. Dans un premier temps, il s'agissait de tester l'utilité d'un court échantillon de description d'image pour prédire si un patient présentant des troubles de langage progressifs présente une imagerie amyloïde positive. Dans un second temps, il fallait cerner un patron de variables cliniquement observables qui supporterait la prise de décision quant au recours subséquent à des procédures coûteuses et invasives comme l'imagerie TEP ou la ponction lombaire. Nous avons adopté une approche hybride en extrayant automatiquement un très large éventail de caractéristiques de discours et procédé à une filtration, définie a priori, des caractéristiques les plus discriminantes.

Fidèlement aux résultats de l'article 1, le patron de caractéristiques discriminantes résultant était multidimensionnel, touchant les aspects lexicaux, syntaxiques, sémantiques et pragmatiques du

discours. En bref, les caractéristiques discriminantes identifiées étaient chez les patients PiB+ : une baisse de la valeur d'imagerie (*imageability*) et de la concrétude moyennes, une hausse de la fréquence objective moyenne et de la dominance moyenne des adjectifs, une plus faible diversité lexicale et une plus faible proportion de substantifs (*nouns proportion per sentence root*), une simplification de la syntaxe en l'absence d'agrammatisme (*left children per sentence root*), une tendance marquée à ne pas spécifier la boisson consommée dans l'image, ainsi que l'expression plus fréquente d'incertitude quant au contenu de l'image, ainsi qu'une plus faible densité sémantique moyenne. L'expérience de classification s'est avérée concluante, démontrant que l'analyse du discours peut renseigner la prise de décision clinique sur la pathologie sous-jacente d'un patient présentant des problèmes de langage prédominants.

Considérant l'hétérogénéité dans les présentations atypiques des maladies neurodégénératives comme l'APP et la MA, un dépistage précoce, peu cher et non invasif demeure une priorité importante afin d'orienter les patients vers les ressources spécialisées que leur état nécessite, mais également afin de leur épargner des procédures qui sont peu susceptibles de contribuer au diagnostic, au traitement et à la prise en charge de la maladie. Dans cette optique, l'article 2 a démontré qu'un court échantillon de discours peut identifier les patients avec une APP qui présentent une pathologie Alzheimer sous-jacente avec une précision, une sensibilité et une spécificité respectivement de 77%, 74% et 79%.

Un patron de variables observables cliniquement a émergé, touchant les dimensions lexicales, syntaxiques, sémantiques et pragmatiques du discours. Dans le contexte de l'adoption massive et soudaine de la télépratique pour l'examen neurologique et l'évaluation neuropsychologique, nos résultats sont en faveur de l'analyse du discours comme outil de dépistage pour les patients présentant des changements langagiers. Une des contributions majeures de l'article 2 est de reposer



sur un groupe clinique hétérogène bien caractérisé et de taille importante relativement à la littérature existante, ce qui a permis d'élaborer une réponse préliminaire à la question pressante de l'identification de la pathologie Alzheimer sous-jacente dans l'APP.

Concernant les retombées cliniques, une identification efficace des anomalies langagières pourrait permettre de référer les patients concernés à des interventions de rééducation. Une récente revue de littérature identifie trois types d'interventions documentées dans la littérature d'intervention en APP (Volkmer et al., 2020). Premièrement, des approches centrées sur le déficit spécifique à chaque variante de l'APP ont démontré leur efficacité, telles que les interventions ciblant l'accès lexical et la fluidité du discours. Deuxièmement, des approches dites compensatoires inspirées de l'aphasiologie vasculaire telles que l'utilisation de gestes, l'adoption de conduites facilitatrices par un proche aidant ou encore l'utilisation d'aides technologiques montrent également des résultats encourageants. Troisièmement, des résultats positifs ont émané d'études sur l'établissement de groupes de support, ce qui est concordant avec notre expérience clinique personnelle. De plus, comme compléments à la rééducation orthophonique, des thérapies de neurostimulation non invasives ont montré des résultats prometteurs et font l'objet d'essais cliniques (Ruksenaite et al., 2021). En particulier, la stimulation transcrânienne à courant continu aurait montré des bénéfices généralisables et durables en accès lexical (Cotelli et al., 2020) et la stimulation magnétique transcrânienne a fait l'objet d'un essai concluant en vlAPP concernant le fonctionnement cognitif global (Bereau et al., 2016).

Concernant l'intervention précoce (au stade de TCL), nous relevons le potentiel de la cueillette d'échantillons de discours comme un instrument clinique simple et informatif. En effet, contrairement à la TEP amyloïde, dont l'utilisation est restreinte à certains patients référés dans une poignée de centres spécialisés, la description d'image pourrait devenir un instrument de

dépistage de routine simple à administrer. Les cliniciens désireux de l'adopter pourront donc se référer à l'étude 1 et à l'étude 2 pour prendre connaissance de certaines variables d'intérêt. Il est à notre avis possible que les variables identifiées dans les deux études puissent se révéler des marqueurs fiables pour une population TCL, mais cette intuition doit faire l'objet d'études confirmatoires. Nous sommes néanmoins encouragés par des résultats computationnels récents qui exploitent des changements de discours qui seraient mesurables au stade de TCL (Fraser, Fors, et al., 2019; Fraser et al., 2018).

#### Efficacité comparée du discours par rapport au bilan cognitif/diagnostic clinique et aux biomarqueurs plasmatiques

Concernant l'efficacité du discours comme dispositif d'identification comparativement à d'autres biomarqueurs, nos résultats suggèrent un rendement avantageux du discours par rapport à deux autres méthodes d'identification de la pathologie sous-jacente. Un premier exemple basé sur l'évaluation cognitive approfondie démontre que la cooccurrence d'un diagnostic clinique de *vlAPP* et d'une dysfonction de la boucle phonologique prédit la pathologie Alzheimer post-mortem avec une sensibilité de 78% et une spécificité de 80% (Giannini et al., 2017). Un second exemple comparatif consiste en la prédiction par apprentissage automatique du résultat d'imagerie amyloïde via des biomarqueurs plasmatiques, qui atteignent une sensibilité de 78% et une spécificité de 77% (Ashton et al., 2019). Bien que les divergences méthodologiques et d'échantillonnage ne permettent pas une comparaison directe de nos résultats avec ces autres marqueurs, il demeure que les marqueurs discursifs tirés d'une simple description d'image offrent un rendement similaire à des procédures surspécialisées et coûteuses. Les marqueurs discursifs pourraient donc servir à identifier les patients APP les plus à risque d'un dépôt anormal de plaques amyloïdes dans les

situations cliniques où un examen neuropsychologique approfondi ou les biomarqueurs ne sont pas immédiatement envisageables.

#### Marqueurs discriminants de pathologie Alzheimer dans l'APP

L'article 2 a révélé un ensemble de marqueurs de discours lexicosémantiques et pragmatiques qui sont globalement compatibles avec deux caractéristiques neuropsychologiques associées à la vlAPP : des difficultés d'accès lexical et un trouble de la mémoire de travail verbale/boucle phonologique. Le substrat anatomique de ces troubles réside dans une atteinte de la matière blanche, principalement la portion antérieure du faisceau longitudinal inférieur, le faisceau unciné et le faisceau supérieur longitudinal ainsi que de la matière grise, notamment la portion postérieure du gyrus temporal supérieur et la région frontale inférieure (Montembeault et al., 2018).

#### Marqueurs non discriminants de pathologie Alzheimer dans l'APP

Les marqueurs significatifs de l'étude 2 ont été discutés dans le manuscrit de l'article publié. En revanche, les marqueurs discursifs extraits automatiquement n'ont pas tous contribué à la classification PiB+ vs PiB-, mais peuvent se révéler important dans le contexte d'un bilan cognitif approfondi. Nous discutons donc brièvement de quelques-uns de ces marqueurs non significatifs.

Certains marqueurs n'ont pas émergé probablement puisqu'ils sont plus typiques d'une atteinte sémantique (présents dans la vsAPP, typiquement PiB-) que de troubles d'accès lexical ou de mémoire de travail verbale (présents dans la vlAPP, typiquement PiB+). Lorsqu'on examine ces marqueurs à travers la lentille du diagnostic clinique (les variantes de l'APP), on remarque que les patients avec une vlAPP (typiquement PiB+) se situent, pour une variable donnée, à un niveau intermédiaire entre les patients avec une vsAPP et ceux avec une vfnAPP (deux groupes typiquement PiB-). Cette distribution explique probablement pourquoi les marqueurs suivants

n'ont pas émergé : usage de pronoms déictiques (ici, là, ceci, ça) et de commentaires liés à l'accès lexical (savoir, se rappeler), le nombre de clauses subordonnées, la longueur des clauses nominales et l'efficacité globale du discours. Effectivement, la littérature sur la vsAPP et des données recueillies au laboratoire suggèrent que ces phénomènes se distribuent sur un continuum d'intensité vnfAPP → vlAPP → vsAPP.

Un autre cas de figure est celui où des marqueurs sont pertinents pour la classification PiB+ vs PiB-, mais sont exclus de l'expérience de classification puisque très fortement corrélés à d'autres marqueurs déjà inclus dans le modèle (cf. Tableau 5 et 9). C'est le cas notamment des plus fortes valence et activation émotionnelles moyennes, qui sont aussi associées à une forte probabilité d'imagerie amyloïde-positif (PiB+). À noter, les trois aspects psycholinguistiques émotionnels (valence, activation et dominance) sont fortement intercorrélés en discours de description d'image. La dominance est le niveau de contrôle, mesuré subjectivement, qu'évoque un mot. Cette mesure est recueillie sur une échelle de Likert allant de « en contrôle » à « contrôlé » auprès d'une trentaine de volontaires (Warriner et al., 2013). Par exemple, dans nos échantillons, le mot *jolly* a une dominance élevée estimée à 7.4/10, alors que le mot *unable* a une dominance faible estimée à 3.08/10. La valence et l'activation, qui sont liées à l'utilisation de mots chargés émotionnellement, sont éclipsés par le marqueur de dominance tel que discutée plus haut et dans l'article 2. Cependant, l'absence du modèle de classification de la valence et de l'activation émotionnelle pour cette raison n'invalide pas l'interprétation anatomo-cognitive proposée pour expliquer leur pertinence dans la détection de la pathologie Alzheimer, notamment le meilleur traitement des concepts émotionnels au sein du cortex insulaire dans la vlAPP comparativement à la vnfAPP (Piguet et al., 2015).

Alors qu'on aurait pu s'attendre à ce que les patients avec une vlAPP forment des phrases raccourcies en raison de leurs problèmes de mémoire de travail auditivoverbale, la longueur

moyenne des phrases telle que mesurée par les librairies de TALN n'est pas discriminante. De façon étonnante, plusieurs patients avec un vnfAPP - pourtant réputés agrammatiques – parviennent à produire de « longues » phrases via la répétition de la conjonction « et », comme l'illustre l'exemple suivant : « ici un bateau *et* ça pêcheur *et* c'est un garçon et... ». Cette trouvaille en apparence étonnante révèle une limite importante des outils automatisés employés dans l'article 2 : les librairies de TALN dans leur configuration standard, dont l'algorithme de segmentation de phrases est tributaire d'un corpus d'entraînement de texte écrit (journaux, littérature, contenu web), peuvent être inadéquates pour segmenter des énoncés de discours de description d'image.

### **3.1.5 Limites de l'article 2**

Une première limite de l'article 2 est liée au fait que le protocole ne comprenait pas l'extraction de caractéristiques issues directement du signal sonore (par opposition aux transcriptions que nous avons employé), alors qu'on sait que certaines caractéristiques acoustiques permettent notamment de différencier la vnfPPA de la vlPPA (Haley et al., 2021) et les patients MA de témoins en santé (Slegers, Filiou, et al., 2018). Des travaux récents en ce sens ont démontré les avantages de l'intégration du signal acoustique dans la détection de l'APP et de la MA-TCL (Themistocleous, Eckerström, et al., 2020; Themistocleous, Ficek, et al., 2020).

Une deuxième limite concerne l'imagerie amyloïde elle-même. La TEP-PiB constitue une approximation de la pathophysiologie de la MA, celle-ci ne pouvant être établie de façon définitive qu'à l'autopsie. Bien que l'imagerie et les résultats de l'autopsie soient fortement corrélés, cette dernière révèle parfois que la pathologie Alzheimer détectée à l'imagerie amyloïde est secondaire à une autre pathologie liée à la DLFT (Santos-Santos et al., 2018).

Une troisième limite est liée à la non-représentativité sociodémographique de notre échantillon d'APP. Comme c'est souvent le cas dans la recherche sur le vieillissement cognitif normal et

pathologique, les participants de notre échantillon sont très éduqués (15-17 ans d'éducation en moyenne) et pas nécessairement représentatifs de toute l'étendue des profils socioculturels des personnes se présentant en clinique de mémoire.

### **3.1.6 Contributions de l'article 2**

En somme, le discours pourrait constituer un intermédiaire intéressant pour prédire la pathologie Alzheimer de façon rapide, économique et non invasive. L'étude a démontré qu'un patron de variables principalement lexicosémantiques et syntaxiques permet d'identifier avec une précision significative les patients avec une plus forte probabilité d'être positifs à l'imagerie amyloïde. Nous avons décrit un éventail de variables significatives que le clinicien peut repérer et qui sont compatibles avec les caractéristiques pathognomoniques du discours dans la vIAPP, c'est-à-dire de profondes difficultés d'accès lexical en l'absence d'agrammatisme. Du point de vue clinique, la classification syndromique (par variante) de l'APP demeure le meilleur prédicteur de la pathologie sous-jacente et permet notamment de distinguer d'autres sous-groupes de pathologie probable (notamment DLFT-tau et TDP-43) parmi les patients amyloïde-négatifs. Cependant, cette démarche diagnostique dépend de longues évaluations qui peuvent être difficiles à réaliser à distance et nécessitent une expertise clinique considérablement spécialisée qui n'est pas disponible dans la plupart des milieux de soins. Ainsi, une analyse rapide du discours telle que décrite dans l'article 2 pourrait permettre d'identifier les patients candidats à l'imagerie amyloïde en attendant qu'ils puissent subir une évaluation cognitive complète.

### **3.2 Contributions scientifiques générales des deux articles**

La portée de nos résultats s'étend donc à la nécessité, dans le contexte du développement d'approches automatisées facilitant le dépistage, non simplement de détecter la présence d'un processus pathologique indéterminé, mais plutôt d'isoler des marqueurs spécifiques à la

pathophysiologie sous-jacente (Holmlund et al., 2020). Pour ce faire, nous avons d'abord synthétisé les connaissances sur les marqueurs discursifs dans la description d'image qui discriminent entre des personnes avec un diagnostic clinique de MA typique et des témoins en santé. Puis, nous avons transféré ces connaissances vers un problème analogue, mais plus près de la réalité clinique : prédire la probabilité qu'un patient se présentant avec des symptômes langagiers présente un résultat de TEP amyloïde positif. Ce transfert de connaissances a mis en relief les caractéristiques du discours qui sont susceptibles de supporter la prise de décision clinique auprès des présentations langagières atypiques de MA : elles sont de nature lexicosémantique et syntaxique, mais également sémantique et pragmatique dans une certaine mesure. Le succès de l'expérience de classification basée sur ces caractéristiques du discours, ainsi que la description de celles-ci sous un angle clinique, ajoutent aux connaissances existantes.

L'article 1 et l'article 2 ont indiqué la prépondérance des marqueurs lexicosémantiques du discours, en particulier psycholinguistiques, pour détecter l'installation d'une pathologie Alzheimer. En effet, les changements psycholinguistiques se sont montrés systématiquement discriminants dans l'article 1 et parmi les caractéristiques les plus prédictives de la pathologie Alzheimer dans l'article 2. Une innovation de l'article 2 est donc d'inclure les plus récentes base de données et un large spectre de caractéristiques psycholinguistiques. En outre, notre laboratoire est bien positionné pour développer empiriquement des bases de données psycholinguistiques spécifiques à la population québécoise (Chedid, Brambati, et al., 2019; Chedid, Wilson, et al., 2019) et les mettre à disposition pour l'analyse du discours québécois francophone. Dans les expériences d'analyse automatisée du discours, ces variables psycholinguistiques spécialisées viennent compléter les variables issues d'une approche computationnelle qu'on qualifierait de conventionnelle.

À l'origine, l'adjonction des variables psycholinguistiques a été motivée par la littérature en neuropsychologie cognitive du vieillissement normal et pathologique, par exemple les études de Bird et al. (2000); Graves et al. (2007); Ralph et al. (1998). Notre démarche d'extraction et de sélection des variables est alignée avec les études comparatives récentes qui démontrent la supériorité d'une approche informée par l'expertise de domaine par rapport à une approche purement algorithmique (Balagopalan et al., 2021), en particulier lorsque les échantillons sont de petite taille (Soares et al., 2016). Afin toutefois de ne pas commettre de manipulation a posteriori des données, nous avons choisi un critère a priori de sélection des caractéristiques (*feature selection*). Ce mécanisme, couplé à l'extraction d'un vaste ensemble des caractéristiques guidée par l'expertise de domaine, a permis de mener une expérience de classification interprétable et reproductible. De plus, nos méthodes sont à notre avis susceptibles d'être généralisables à d'autres jeux de données pour identifier les variables discriminantes dans d'autres populations cliniques.

Au-delà de ces résultats spécifiques, notre contribution démontre l'intérêt de travailler à l'interface de l'IA et des neurosciences du vieillissement. Une approche hybride a permis de déployer des outils de TALN disponibles en libre-accès en combinaison avec des variables issues des sciences cognitives et neurosciences cliniques. L'apprentissage machine permet une automatisation à deux niveaux dans l'article 2 : à l'étape d'extraction des caractéristiques et à celle de la classification des échantillons. Dans la présente thèse, toutefois, l'utilisation de l'IA demeure un moyen technique servant à raffiner les connaissances en neurosciences cliniques.

L'approche a permis de traiter avec succès d'une question jusqu'ici jamais abordée dans la littérature existante, à savoir l'identification, strictement à partir du discours, des patients avec une APP étant le plus susceptibles de présenter une imagerie amyloïde positive. En ce sens, les contributions des articles consistent à supporter la prise de décision médicale au sujet de procédures



coûteuses et invasives. Notre dispositif pourrait être utilisé en télépratique ou dans un contexte où les ressources cliniques sont limitées.

### **Avenues futures**

Sur le plan technique et conformément aux pratiques recommandées pour l'évaluation des technologies biométriques émergentes, il serait utile d'établir un jeu de données commun et de réserver un échantillon de test séparé de l'échantillon d'entraînement (Goldsack et al., 2020) dans l'éventualité où on voudrait reproduire, étendre ou reproduire l'article 2. De plus, il serait pertinent de procéder à l'adaptation de domaine (Daume III et Marcu, 2006; Han et Eisenstein, 2019) de nos modèles d'IA, opération consistant, suite à un pré-entraînement sur un corpus général, à reparamétrer l'algorithme sur un ensemble de données plus représentatives de son cas d'application (pour un exemple lié à l'étude du discours dans les maladies neurodégénératives, voir Vásquez-Correa et al. (2021)). Concrètement, par exemple, on pourrait recourir à des vecteurs entraînés sur un corpus de description d'image s'il était possible d'en colliger un de taille suffisante. De façon plus réaliste, on pourrait utiliser en pré-entraînement un corpus de sous-titres de films (Van Paridon et Thompson, 2021) qui sont susceptibles, avec leur contenu dialogique, de mieux refléter la structure linguistique du discours de description d'image. Cette adaptation pourrait augmenter la performance des outils de TALN actuels qui sont habituellement entraînés sur des corpus de texte écrit (Wikipédia, presse quotidienne, littérature).

Sur le plan neurocognitif, la part relative des troubles d'accès lexical relativement aux troubles sémantiques dans la « signature » discursive de patients avec une MA ou une APP dépasse le cadre de la présente thèse. Des travaux connexes devraient être poursuivis en appariant les bases neuronales (atrophies des lobes temporaux antérieurs ou de la jonction temporopariétale) de ces déficits aux caractéristiques du discours qui y sont associées afin de bien dégager les marqueurs

discursifs les plus spécifiques de l'un et l'autre de ces déficits neuropsychologiques. En lien avec les mesures neuropsychologiques traditionnelles, Boller et al. (1991) ont démontré la supériorité de la dénomination d'images pour prédire un déclin cognitif plus rapide dans la MA, alors que Mickes et al. (2007) ont établi que les scores en dénomination sont plus bas chez les patients qui convertissent vers une MA, deux ans avant le diagnostic. Une étude comparative des différentes épreuves pour prédire la conversion TCL-MA a été réalisée par Belleville et al. (2017). Dans une approche analogue et en incluant des mesures de discours, il serait intéressant de comparer l'efficacité de différents outils psychométriques pour prédire l'imagerie amyloïde ou un autre biomarqueur.

En ce qui concerne la pertinence des outils computationnels pour la singularité socioculturelle québécoise, les échantillons discursifs devraient en outre être culturellement adaptés à la population de référence (Holmlund et al., 2021). Dans le contexte québécois, il importe de documenter les facteurs comme le bilinguisme ou le plurilinguisme, puisque ces facteurs ont montré des résultats protecteurs chez les immigrants et les francophones montréalais, retardant chez les plurilingues de quelques années le diagnostic de la MA (Chertkow et al., 2010). En ce sens, la rareté des données normatives adaptées, qui est déploré depuis quelques années dans les études computationnelles (Masrani et al., 2017; Noorian et al., 2017), pourrait être palliée par des initiatives comme le Consortium Canadien en Neurodégénérescence associée au Vieillissement (CCNV). En effet, contrairement aux radiologistes par exemple, les cliniciens en neuropsychiatrie doivent intégrer des données nuancées, notamment psychosociales, afin de parvenir à un diagnostic, ce que la machine est peu susceptible d'arriver à faire sans supervision. On note néanmoins que certains tentent d'intégrer le dossier informatisé de l'utilisateur à la prédiction de l'algorithme concernant la

présence et la sévérité d'une atteinte cognitive (Vyas et al., 2021), ce qui n'est pas sans soulever de complexes questions épistémiques, éthiques et légales.

## **Conclusion**

Dans le contexte de l'expansion rapide des approches automatisées à l'analyse du discours et dans le but d'éclairer la sélection des variables pertinentes dans ce type d'études, nous avons brossé un portrait quantitatif exhaustif des variables qui différencient les patients avec une MA de témoins en santé. Soutenus par cette démarche et l'adjonction de variables issues de la neuropsychologie cognitive des APP, nous avons montré qu'il est possible de prédire dans une certaine mesure quels patients avec des troubles du langage présentent un résultat d'imagerie amyloïde positif. Nous avons ensuite documenté et commenté dans une perspective clinique les variables les plus prédictives du résultat d'imagerie amyloïde.



## Références bibliographiques

- Ahmed, S., de Jager, C. A., Haigh, A.-M. et Garrard, P. (2013). Semantic processing in connected speech at a uniformly early stage of autopsy-confirmed Alzheimer's disease. *Neuropsychology*, 27(1), 79.
- Ahmed, S., Haigh, A.-M. F., de Jager, C. A. et Garrard, P. (2013). Connected speech as a marker of disease progression in autopsy-proven Alzheimer's disease. *Brain*, 136(12), 3727-3737.
- Amieva, H., Jacqmin-Gadda, H., Orgogozo, J.-M., Le Carret, N., Helmer, C., Letenneur, L., Barberger-Gateau, P., Fabrigoule, C. et Dartigues, J.-F. (2005). The 9 year cognitive decline before dementia of the Alzheimer type: a prospective population-based study. *Brain*, 128(5), 1093-1101.
- Ash, S., Evans, E., O'shea, J., Powers, J., Boller, A., Weinberg, D., Haley, J., McMillan, C., Irwin, D. J. et Rascovsky, K. (2013). Differentiating primary progressive aphasia in a brief sample of connected speech. *Neurology*, 81(4), 329-336.
- Ashton, N. J., Nevado-Holgado, A. J., Barber, I. S., Lynham, S., Gupta, V., Chatterjee, P., Goozee, K., Hone, E., Pedrini, S. et Blennow, K. (2019). A plasma protein classifier for predicting amyloid burden for preclinical Alzheimer's disease. *Science advances*, 5(2), eaau7220.
- Auclair-Ouellet, N. (2015). Inflectional morphology in primary progressive aphasia and Alzheimer's disease: A systematic review. *Journal of Neurolinguistics*, 34, 41-64.
- Bahar-Fuchs, A., Villemagne, V., Ong, K., Chetelat, G., Lamb, F., Reiningner, C. B., Woodward, M. et Rowe, C. C. (2013). Prediction of amyloid- $\beta$  pathology in amnesic mild cognitive impairment with neuropsychological tests. *Journal of Alzheimer's Disease*, 33(2), 451-462.
- Balagopalan, A., Eyre, B., Robin, J., Rudzicz, F. et Novikova, J. (2021). Comparing Pre-trained and Feature-Based Models for Prediction of Alzheimer's Disease Based on Speech. *Frontiers in Aging Neuroscience*, 13, 189.
- Balagopalan, A., Eyre, B., Rudzicz, F. et Novikova, J. J. a. p. a. (2020). To BERT or Not To BERT: Comparing Speech and Language-based Approaches for Alzheimer's Disease Detection.
- Balthazar, M. L. F., Cendes, F. et Damasceno, B. P. (2008). Semantic error patterns on the Boston Naming Test in normal aging, amnesic mild cognitive impairment, and mild Alzheimer's disease: Is there semantic disruption? *Neuropsychology*, 22(6), 703-709.  
<https://doi.org/10.1037/a0012919>

- Battista, P., Salvatore, C., Berlingeri, M., Cerasa, A. et Castiglioni, I. (2020). Artificial Intelligence And Neuropsychological Measures: The Case Of Alzheimer's Disease. *Neuroscience and Biobehavioral Reviews*.
- Bayles, K. A. et Tomoeda, C. K. (1991). Caregiver Report of Prevalence and Appearance Order of Linguistic Symptoms in Alzheimer's Patients<sup>1</sup>. *The Gerontologist*, 31(2), 210-216. <https://doi.org/10.1093/geront/31.2.210> %J The Gerontologist
- Belleville, S., Fouquet, C., Hudon, C., Zomahoun, H. T. V. et Croteau, J. (2017). Neuropsychological measures that predict progression from mild cognitive impairment to Alzheimer's type dementia in older adults: a systematic review and meta-analysis. *Neuropsychology Review*, 27(4), 328-353.
- Beltrami, D., Gagliardi, G., Rossini Favretti, R., Ghidoni, E., Tamburini, F. et Calzà, L. (2018). Speech analysis by natural language processing techniques: A possible tool for very early detection of cognitive decline? *Frontiers in Aging Neuroscience*, 10, 369.
- Bennett, J. A. J. N. R. (2005). The consolidated standards of reporting trials (CONSORT): Guidelines for reporting randomized trials. 54(2), 128-132.
- Bensaïdane, M. R., Beaugregard, J.-M., Poulin, S., Buteau, F.-A., Guimond, J., Bergeron, D., Verret, L., Fortin, M.-P., Houde, M. et Bouchard, R. W. (2016). Clinical utility of amyloid PET imaging in the differential diagnosis of atypical dementias and its impact on caregivers. *Journal of Alzheimer's Disease*, 52(4), 1251-1262.
- Bereau, M., Magnin, E., Nicolier, M., Berthet, L., Dariel, E., Ferreira, S., Sylvestre, G., Monnin, J., Chopard, G. et Bouladour, H. J. E. n. (2016). Left prefrontal repetitive transcranial magnetic stimulation in a logopenic variant of primary progressive aphasia: a case report. 76(1-2), 12-18.
- Bergeron, D., Gorno-Tempini, M. L., Rabinovici, G. D., Santos-Santos, M. A., Seeley, W., Miller, B. L., Pijnenburg, Y., Keulen, M. A., Groot, C. et van Berckel, B. N. (2018). Prevalence of amyloid- $\beta$  pathology in distinct variants of primary progressive aphasia. *Annals of Neurology*, 84(5), 729-740.
- Bird, H., Ralph, M. A. L., Patterson, K. et Hodges, J. R. (2000). The rise and fall of frequency and imageability: Noun and verb production in semantic dementia. *Brain and Language*, 73(1), 17-49.

- Boller, F., Becker, J. T., Holland, A. L., Forbes, M. M., Hood, P. C. et McGonigle-Gibson, K. L. J. C. (1991). Predictors of decline in Alzheimer's disease. *27*(1), 9-17.
- Boschi, V., Catricalà, E., Consonni, M., Chesi, C., Moro, A. et Cappa, S. F. (2017). Connected Speech in Neurodegenerative Language Disorders: A Review. *Frontiers in Psychology*, *8*.
- Braak, H. et Braak, E. J. A. n. (1991). Neuropathological staging of Alzheimer-related changes. *82*(4), 239-259.
- Brysbaert, M. et New, B. (2009, November 01). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English [journal article]. *Behavior Research Methods*, *41*(4), 977-990. <https://doi.org/10.3758/brm.41.4.977>
- Brysbaert, M., Warriner, A. B. et Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904-911.
- Bucks, R. S., Singh, S., Cuerden, J. M. et Wilcock, G. K. (2000). Analysis of spontaneous, conversational speech in dementia of Alzheimer type: Evaluation of an objective technique for analysing lexical performance. *Aphasiology*, *14*(1), 71-91.
- Canada, S. (2021). <https://www.statcan.gc.ca/fr/science-donnees/ressources/terminologie>
- Canada, S. A. d. (2016). *Prévalence et coûts financiers des maladies cognitives au Canada*.
- Chedid, G., Brambati, S. M., Bedetti, C., Rey, A. E., Wilson, M. A. et Vallet, G. T. J. B. r. m. (2019). Visual and auditory perceptual strength norms for 3,596 French nouns and their relationship with other psycholinguistic variables. *51*(5), 2094-2105.
- Chedid, G., Wilson, M. A., Bedetti, C., Rey, A. E., Vallet, G. T. et Brambati, S. M. J. B. r. m. (2019). Norms of conceptual familiarity for 3,596 French nouns and their contribution in lexical decision. *51*(5), 2238-2247.
- Cherney, L. R. et Canter, G. J. (1993). Informational content in the discourse of patients with probable Alzheimer's disease and patients with right brain damage. *Clinical aphasiology*, *21*, 123-134.
- Chertkow, H., Whitehead, V., Phillips, N., Wolfson, C., Atherton, J., Bergman, H. J. A. D. et Disorders, A. (2010). Multilingualism (but not always bilingualism) delays the onset of Alzheimer disease: evidence from a bilingual community. *24*(2), 118-125.

- Cho, S., Nevler, N., Ash, S., Shellikeri, S., Irwin, D. J., Massimo, L., Rascovsky, K., Olm, C., Grossman, M. et Liberman, M. J. C. (2021). Automated analysis of lexical features in frontotemporal degeneration.
- Clarke, N., Foltz, P. et Garrard, P. (2020, 2020/05/19/). How to do things with (thousands of) words: Computational approaches to discourse analysis in Alzheimer's disease. *Cortex*. <https://doi.org/https://doi.org/10.1016/j.cortex.2020.05.001>
- Cotelli, M., Manenti, R., Ferrari, C., Gobbi, E., Macis, A., Cappa, S. F. J. N. et Reviews, B. (2020). Effectiveness of language training and non-invasive brain stimulation on oral and written naming performance in Primary Progressive Aphasia: A meta-analysis and systematic review. *108*, 498-525.
- Covatti Malcorra, B. L., Mota, N. B., Weissheimer, J., Schilling, L. P., Wilson, M. A. et Hübner, L. C. J. J. o. A. s. D. Low Speech Connectedness in Alzheimer's Disease is Associated with Poorer Semantic Memory Performance. (Preprint), 1-8.
- Croisile, B., Ska, B., Brabant, M.-J., Duchene, A., Lepage, Y., Aimard, G. et Trillet, M. (1996). Comparative study of oral and written picture description in patients with Alzheimer's disease. *Brain and Language*, *53*(1), 1-19.
- Crystal, D. (2011). *A dictionary of linguistics and phonetics* (vol. 30). John Wiley & Sons.
- Cummings, J. L., Benson, D. F., Hill, M. A. et Read, S. (1985). Aphasia in dementia of the Alzheimer type. *Neurology*, *35*(3), 394-394.
- Daume III, H. et Marcu, D. (2006). Domain adaptation for statistical classifiers. *Journal of artificial Intelligence research*, *26*, 101-126.
- de la Fuente Garcia, S., Ritchie, C. et Luz, S. (2020a). Artificial Intelligence, Speech, and Language Processing Approaches to Monitoring Alzheimer's Disease: A Systematic Review. *Journal of Alzheimer's Disease*, (Preprint), 1-27.
- de la Fuente Garcia, S., Ritchie, C. et Luz, S. J. J. o. A. s. D. (2020b). Artificial Intelligence, Speech, and Language Processing Approaches to Monitoring Alzheimer's Disease: A Systematic Review. (Preprint), 1-27.
- Des Jarlais, D. C., Lyles, C., Crepaz, N. et health, T. G. J. A. j. o. p. (2004). Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *94*(3), 361-366.



- Dubois, B., Feldman, H. H., Jacova, C., Cummings, J. L., DeKosky, S. T., Barberger-Gateau, P., Delacourte, A., Frisoni, G., Fox, N. C. et Galasko, D. (2010). Revising the definition of Alzheimer's disease: a new lexicon. *The Lancet Neurology*, 9(11), 1118-1127.
- Dubois, B., Feldman, H. H., Jacova, C., Hampel, H., Molinuevo, J. L., Blennow, K., DeKosky, S. T., Gauthier, S., Selkoe, D. et Bateman, R. J. T. L. N. (2014). Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. *13*(6), 614-629.
- Dubois, B. et Michon, A. (2015). *Démences*. Doin-John Libbey Eurotext.
- Eckerström, C., Olsson, E., Bjerke, M., Malmgren, H., Edman, Å., Wallin, A. et Nordlund, A. J. J. o. A. s. D. (2013). A combination of neuropsychological, neuroimaging, and cerebrospinal fluid markers predicts conversion from mild cognitive impairment to dementia. *36*(3), 421-431.
- Ehrlich, J. S., Obler, L. K. et Clark, L. (1997). Ideational and semantic contributions to narrative production in adults with dementia of the Alzheimer's type. *Journal of Communication Disorders*, 30(2), 79-99.
- Filiou, R.-P., Bier, N., Slegers, A., Houzé, B., Belchior, P. et Brambati, S. M. (2019). Connected speech assessment in the early detection of Alzheimer's disease and mild cognitive impairment: a scoping review. *Aphasiology*, 1-33. <https://doi.org/10.1080/02687038.2019.1608502>
- Fink, H. A., Hemmy, L. S., Linskens, E. J., Silverman, P. C., MacDonald, R., McCarten, J. R., Talley, K. M., Desai, P. J., Forte, M. L. et Miller, M. A. (2020). Diagnosis and Treatment of Clinical Alzheimer's-Type Dementia: A Systematic Review.
- Fraser, K. C., Fors, K. L., Kokkinakis, D. J. C. S. et Language. (2019). Multilingual word embeddings for the assessment of narrative speech in mild cognitive impairment. *53*, 121-139.
- Fraser, K. C., Linz, N., Li, B., Fors, K. L., Rudzicz, F., König, A., Alexandersson, J., Robert, P. et Kokkinakis, D. (2019). Multilingual prediction of Alzheimer's disease through domain adaptation and concept-based language modelling. Dans. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers).
- Fraser, K. C., Lundholm Fors, K., Eckerström, M., Öhman, F. et Kokkinakis, D. J. F. i. a. n. (2019). Predicting MCI status from multimodal language data using cascaded classifiers. *11*, 205.

- Fraser, K. C., Lundholm Fors, K., Eckerström, M., Themistocleous, C. et Kokkinakis, D. (2018). Improving the sensitivity and specificity of MCI screening with linguistic information. Dans. LREC workshop: RaPID-2. Miyazaki, Japan.
- Fraser, K. C., Meltzer, J. A., Graham, N. L., Leonard, C., Hirst, G., Black, S. E. et Rochon, E. (2014, 6//). Automated classification of primary progressive aphasia subtypes from narrative speech transcripts. *Cortex*, 55, 43-60. <https://doi.org/http://dx.doi.org/10.1016/j.cortex.2012.12.006>
- Fraser, K. C., Meltzer, J. A. et Rudzicz, F. (2016). Linguistic features identify Alzheimer's disease in narrative speech. *Journal of Alzheimer's Disease*, 49(2), 407-422.
- Gandy, S., Knopman, D. S. et Sano, M. J. M. N. (2021). Talking points for physicians, patients and caregivers considering Aduhelm® infusion and the accelerated pathway for its approval by the FDA. *16*(1), 1-4.
- Garrard, P., Rentoumi, V., Gesierich, B., Miller, B. et Gorno-Tempini, M. L. (2014). Machine learning approaches to diagnosis and laterality effects in semantic dementia discourse. *Cortex*, 55, 122-129.
- Giannini, L. A., Irwin, D. J., McMillan, C. T., Ash, S., Rascovsky, K., Wolk, D. A., Van Deerlin, V. M., Lee, E. B., Trojanowski, J. Q. et Grossman, M. (2017). Clinical marker for Alzheimer disease pathology in logopenic primary progressive aphasia. *Neurology*, 88(24), 2276-2284.
- Giles, E., Patterson, K. et Hodges, J. R. (1996). Performance on the Boston Cookie Theft picture description task in patients with early dementia of the Alzheimer's type: missing information. *Aphasiology*, 10(4), 395-408.
- Goldsack, J. C., Coravos, A., Bakker, J. P., Bent, B., Dowling, A. V., Fitzer-Attas, C., Godfrey, A., Godino, J. G., Gujar, N. et Izmailova, E. J. n. d. M. (2020). Verification, analytical validation, and clinical validation (V3): the foundation of determining fit-for-purpose for Biometric Monitoring Technologies (BioMeTs). 3(1), 1-15.
- Goodglass, H., Kaplan, E. et Barresi, B. (2000). Boston Diagnostic Aphasia Examination-(BDAE-3) San Antonio. TX: Psychological Corporation.
- Gorno-Tempini, M., Hillis, A., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. e., emsp14, al, Ogar, J., Rohrer, J., Black, S. et Boeve, B. F. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, 76(11), 1006-1014.

- Gorno-Tempini, M. L., Brambati, S. M., Ginex, V., Ogar, J., Dronkers, N. F., Marcone, A., Perani, D., Garibotto, V., Cappa, S. et Miller, B. (2008). The logopenic/phonological variant of primary progressive aphasia. *Neurology*, *71*(16), 1227-1234.
- Gorno-Tempini, M. L., Dronkers, N. F., Rankin, K. P., Ogar, J. M., Phengrasamy, L., Rosen, H. J., Johnson, J. K., Weiner, M. W. et Miller, B. L. (2004). Cognition and anatomy in three variants of primary progressive aphasia. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, *55*(3), 335-346.
- Graves, W. W., Grabowski, T. J., Mehta, S. et Gordon, J. K. (2007). A neural signature of phonological access: distinguishing the effects of word frequency from familiarity and length in overt picture naming. *Journal of Cognitive Neuroscience*, *19*(4), 617-631.
- Groves-Wright, K., Neils-Strunjas, J., Burnett, R. et O'Neill, M. J. (2004). A comparison of verbal and written language in Alzheimer's disease. *Journal of Communication Disorders*, *37*(2), 109-130.
- Haley, K. L., Jacks, A., Jarrett, J., Ray, T., Cunningham, K. T., Gorno-Tempini, M. L., Henry, M. L. J. J. o. S., Language, et Research, H. (2021). Speech Metrics and Samples That Differentiate Between Nonfluent/Agrammatic and Logopenic Variants of Primary Progressive Aphasia. 1-22.
- Han, X. et Eisenstein, J. J. a. p. a. (2019). Unsupervised domain adaptation of contextualized embeddings for sequence labeling.
- Hastie, T., Tibshirani, R. et Friedman, J. J. C. o. (2009). The elements of statistical learnin. 33.
- Hernández-Domínguez, L., Ratté, S., Sierra-Martínez, G. et Roche-Bergua, A. (2018). Computer-based evaluation of Alzheimer's disease and mild cognitive impairment patients during a picture description task. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, *10*, 260-268.
- Hier, D. B., Hagenlocker, K. et Shindler, A. G. (1985). Language disintegration in dementia: Effects of etiology and severity. *Brain and Language*, *25*(1), 117-133.
- Hodges, J. R., Salmon, D. P. et Butters, N. (1991). The nature of the naming deficit in Alzheimer's and Huntington's disease. *Brain*, *114*(4), 1547-1558.
- Holmlund, T. B., Diaz-Asper, C. et Elvevåg, B. (2020). The reality of doing things with (thousands of) words in applied research and clinical settings: A commentary on Clarke et al.(2019). *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*.

Holmlund, T. B., Diaz-Asper, C. et Elvevåg, B. (2021, 2021/03/01/). The reality of doing things with (thousands of) words in applied research and clinical settings: A commentary on Clarke et al. (2020). *Cortex*, 136, 150-156.  
<https://doi.org/https://doi.org/10.1016/j.cortex.2020.08.024>

[Record #90480 is using a reference type undefined in this output style.]

Illes, J. (1989). Neurolinguistic features of spontaneous language production dissociate three forms of neurodegenerative disease: Alzheimer's, Huntington's, and Parkinson's. *Brain and Language*, 37(4), 628-642.

Ivensky, I. (2019). *Automatic detection of semantic dementia from scene description* [Université de Montréal].

Jarrold, W. L., Peintner, B., Yeh, E., Krasnow, R., Javitz, H. S. et Swan, G. E. (2010). Language analytics for assessing brain health: Cognitive impairment, depression and pre-symptomatic Alzheimer's disease. Dans. International Conference on Brain Informatics.

Jensen, A. M., Chenery, H. J. et Copland, D. A. (2006). A comparison of picture description abilities in individuals with vascular subcortical lesions and Huntington's disease. *Journal of Communication Disorders*, 39(1), 62-77.

Johnson, K. A., Minoshima, S., Bohnen, N. I., Donohoe, K. J., Foster, N. L., Herscovitch, P., Karlawish, J. H., Rowe, C. C., Carrillo, M. C. et Hartley, D. M. (2013). Appropriate use criteria for amyloid PET: a report of the Amyloid Imaging Task Force, the Society of Nuclear Medicine and Molecular Imaging, and the Alzheimer's Association. *Journal of Nuclear Medicine*, 54(3), 476-490.

Joubert, S., Brambati, S. M., Ansado, J., Barbeau, E. J., Felician, O., Didic, M., Lacombe, J., Goldstein, R., Chayer, C. et Kergoat, M.-J. (2010, 3//). The cognitive and neural expression of semantic memory impairment in mild cognitive impairment and early Alzheimer's disease. *Neuropsychologia*, 48(4), 978-988.  
<https://doi.org/http://dx.doi.org/10.1016/j.neuropsychologia.2009.11.019>

Joubert, S., Vallet, G. T., Montembeault, M., Boukadi, M., Wilson, M. A., Rouleau, I. et Brambati, S. M. (2017). Comprehension of concrete and abstract words in semantic variant primary progressive aphasia and Alzheimer's disease: A behavioral and neuroimaging study. *Brain and Language*, 170, 93-102.

- Kamath, V., Sutherland, E. R. et Chaney, G.-A. (2020). A Meta-Analysis of Neuropsychological Functioning in the Logopenic Variant of Primary Progressive Aphasia: Comparison with the Semantic and Non-Fluent Variants. *Journal of the International Neuropsychological Society*, 26(3), 322-330.
- Karlawish, J. et Grill, J. D. J. N. R. N. (2021). The approval of Aduhelm risks eroding public trust in Alzheimer research and the FDA. 1-2.
- Kavé, G. et Dassa, A. (2018). Severity of Alzheimer's disease and language features in picture descriptions. *Aphasiology*, 32(1), 27-40.
- Kavé, G. et Goral, M. (2016). Word retrieval in picture descriptions produced by individuals with Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology*, 1-9.
- Kavé, G. et Goral, M. (2018). Word retrieval in connected speech in Alzheimer's disease: a review with meta-analyses. *Aphasiology*, 32(1), 4-26.
- Kertesz, A. (1982). *Western aphasia battery test manual*. Psychological Corp.
- Kuperman, V., Stadthagen-Gonzalez, H. et Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, 44(4), 978-990.
- Laforce, R., Rosa-Neto, P., Soucy, J.-P., Rabinovici, G. D., Dubois, B. et Gauthier, S. (2016). Canadian consensus guidelines on use of amyloid imaging in Canada: update and future directions from the specialized task force on amyloid imaging in Canada. *Canadian Journal of Neurological Sciences*, 43(4), 503-512.
- Land Jr, W. H. et Schaffer, J. D. (2019). *The Art and Science of Machine Intelligence: With An Innovative Application for Alzheimer's Detection from Speech*. Springer.
- Laske, C., Sohrabi, H. R., Frost, S. M., López-de-Ipiña, K., Garrard, P., Buscema, M., Dauwels, J., Soekadar, S. R., Mueller, S. et Linnemann, C. (2015). Innovative diagnostic tools for early detection of Alzheimer's disease. *Alzheimer's & Dementia*, 11(5), 561-578.
- Liu, C.-C., Kanekiyo, T., Xu, H. et Bu, G. J. N. R. N. (2013). Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. 9(2), 106-118.
- Luz, S., Haider, F., de la Fuente, S., Fromm, D. et MacWhinney, B. (2020). Alzheimer's Dementia Recognition through Spontaneous Speech: The ADReSS Challenge. *arXiv preprint arXiv:2004.06833*.
- March, E. G., Wales, R. et Pattison, P. (2006). The uses of nouns and deixis in discourse production in Alzheimer's disease. *Journal of Neurolinguistics*, 19(4), 311-340.

- Martínez-Nicolás, I., Llorente, T. E., Martínez-Sánchez, F. et Meilán, J. J. G. J. F. i. P. (2021). Ten Years of Research on Automatic Voice and Speech Analysis of People With Alzheimer's Disease and Mild Cognitive Impairment: A Systematic Review Article. *12*, 645.
- Masrani, V., Murray, G., Field, T. S. et Carenini, G. (2017). Domain adaptation for detecting mild cognitive impairment. Dans. Canadian Conference on Artificial Intelligence.
- McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Kawas, C. H., Klunk, W. E., Koroshetz, W. J., Manly, J. J. et Mayeux, R. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3), 263-269.
- Mesulam, M. (1982). Slowly progressive aphasia without generalized dementia. *Annals of Neurology*, 11(6), 592-598.
- Mesulam, M. (1987). Primary progressive aphasia—differentiation from Alzheimer's disease. *Annals of Neurology*, 22(4), 533-534.
- Mesulam, M. (2001). Primary progressive aphasia. *Annals of Neurology*, 49(4), 425-432.
- Mickes, L., Wixted, J. T., Fennema-Notestine, C., Galasko, D., Bondi, M. W., Thal, L. J. et Salmon, D. P. J. N. (2007). Progressive impairment on neuropsychological tasks in a longitudinal study of preclinical Alzheimer's disease. *21*(6), 696.
- Mitrushina, M., Boone, K. B., Razani, J. et DELia, L. F. (2005). *Handbook of normative data for neuropsychological assessment*. Oxford University Press.
- Mittal, A., Sahoo, S., Datar, A., Kadiwala, J., Shalu, H. et Mathew, J. J. a. p. a. (2020). Multi-Modal Detection of Alzheimer's Disease from Speech and Text.
- Möller, H.-J. et Graeber, M. (1998). The case described by Alois Alzheimer in 1911. *European Archives of Psychiatry and Clinical Neuroscience*, 248(3), 111-122.
- Montembeault, M., Brambati, S., Joubert, S., Boukadi, M., Chapleau, M., Laforce, R. J., Wilson, M., Macoir, J. et Rouleau, I. (2017). Naming unique entities in the semantic variant of primary progressive aphasia and Alzheimer's disease: Towards a better understanding of the semantic impairment. *Neuropsychologia*, 95, 11-20.
- Montembeault, M., Brambati, S. M., Gorno-Tempini, M. L. et Migliaccio, R. (2018). Clinical, anatomical, and pathological features in the three variants of primary progressive aphasia: a review. *Frontiers in Neurology*, 9, 692.

- Morris, J. C., Blennow, K., Frölich, L., Nordberg, A., Soininen, H., Waldemar, G., Wahlund, L. O. et Dubois, B. J. J. o. i. m. (2014). Harmonized diagnostic criteria for Alzheimer's disease: recommendations. *275*(3), 204-213.
- Mueller, K. D., Hermann, B., Mecollari, J. et Turkstra, L. S. (2018). Connected speech and language in mild cognitive impairment and Alzheimer's disease: A review of picture description tasks. *Journal of Clinical and Experimental Neuropsychology*, 1-23.
- Murdoch, B. E., Chenery, H. J., Wilks, V., Boyle, R. S. J. B. et language. (1987). Language disorders in dementia of the Alzheimer type. *31*(1), 122-137.
- Murray, J., Schneider, J., Banerjee, S. et Mann, A. (1999). EURO CARE: a cross-national study of co-resident spouse carers for people with Alzheimer's disease: II—a qualitative analysis of the experience of caregiving. *International Journal of Geriatric Psychiatry*, *14*(8), 662-667.
- Myszczyńska, M. A., Ojiamies, P. N., Lacoste, A. M. B., Neil, D., Saffari, A., Mead, R., Hautbergue, G. M., Holbrook, J. D. et Ferraiuolo, L. (2020, 2020/07/15). Applications of machine learning to diagnosis and treatment of neurodegenerative diseases. *Nature Reviews Neurology*. <https://doi.org/10.1038/s41582-020-0377-8>
- Nicholas, M., Obler, L. K., Albert, M. L. et Helm-Estabrooks, N. (1985). Empty speech in Alzheimer's disease and fluent aphasia. *Journal of Speech, Language, and Hearing Research*, *28*(3), 405-410.
- Nicholas, M., Obler, L. K., Au, R. et Albert, M. L. (1996, 8//). On the Nature of Naming Errors in Aging and Dementia: A Study of Semantic Relatedness. *Brain and Language*, *54*(2), 184-195. <https://doi.org/http://dx.doi.org/10.1006/brln.1996.0070>
- Noorian, Z., Pou-Prom, C. et Rudzicz, F. (2017). On the importance of normative data in speech-based assessment. *arXiv preprint arXiv:1712.00069*.
- Orange, J. B. et Colton-Hudson, A. (1998). Enhancing communication in dementia of the Alzheimer's type. *Topics in Geriatric Rehabilitation*, *14*(2), 56-75.
- Ostrand, R., Gunstad, J. J. J. o. G. P. et Neurology. (2021). Using automatic assessment of speech production to predict current and future cognitive function in older adults. *34*(5), 357-369.
- Padhee, S., Illendula, A., Sadler, M., Shalin, V. L., Banerjee, T., Thirunarayan, K. et Romine, W. L. (2020). Predicting Early Indicators of Cognitive Decline from Verbal Utterances. Dans. 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM).

- Peintner, B., Jarrold, W., Vergyri, D., Richey, C., Tempini, M. L. G. et Ogar, J. (2008). Learning diagnostic models using speech and language measures. Dans. Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE.
- Petersen, R. C., Roberts, R. O., Knopman, D. S., Boeve, B. F., Geda, Y. E., Ivnik, R. J., Smith, G. E. et Jack, C. R. J. A. o. n. (2009). Mild cognitive impairment: ten years later. *66*(12), 1447-1455.
- Piaceri, I., Nacmias, B. et Sorbi, S. J. F. B. (2013). Genetics of familial and sporadic Alzheimer's disease. *5*(1), 167-177.
- Piguet, O., Leyton, C. E., Gleeson, L. D., Hoon, C. et Hodges, J. R. J. J. o. A. s. D. (2015). Memory and emotion processing performance contributes to the diagnosis of non-semantic primary progressive aphasia syndromes. *44*(2), 541-547.
- Rabinovici, G., Furst, A., O'neil, J., Racine, C., Mormino, E., Baker, S., Chetty, S., Patel, P., Pagliaro, T. et Klunk, W. J. N. (2007). 11C-PIB PET imaging in Alzheimer disease and frontotemporal lobar degeneration. *68*(15), 1205-1212.
- Rabinovici, G., Rosen, H., Alkalay, A., Kornak, J., Furst, A., Agarwal, N., Mormino, E., O'Neil, J., Janabi, M. et Karydas, A. (2011). Amyloid vs FDG-PET in the differential diagnosis of AD and FTLD. *Neurology*, *77*(23), 2034-2042.
- Rabinovici, G. D., Gatzonis, C., Apgar, C., Chaudhary, K., Gareen, I., Hanna, L., Hendrix, J., Hillner, B. E., Olson, C. et Lesman-Segev, O. H. (2019). Association of amyloid positron emission tomography with subsequent change in clinical management among medicare beneficiaries with mild cognitive impairment or dementia. *JAMA*, *321*(13), 1286-1294.
- Rabinovici, G. D., Jagust, W. J., Furst, A. J., Ogar, J. M., Racine, C. A., Mormino, E. C., O'Neil, J. P., Lal, R. A., Dronkers, N. F. et Miller, B. L. (2008). A $\beta$  amyloid and glucose metabolism in three variants of primary progressive aphasia. *Annals of neurology*, *64*(4), 388-401.
- Ralph, M. A. L., Graham, K. S., Ellis, A. W. et Hodges, J. R. (1998). Naming in semantic dementia—what matters? *Neuropsychologia*, *36*(8), 775-784.
- Rentoumi, V., Raoufian, L., Ahmed, S., de Jager, C. A. et Garrard, P. (2014). Features and machine learning classification of connected speech samples from patients with autopsy proven



- Alzheimer's disease with and without additional vascular pathology. *Journal of Alzheimer's Disease*, 42(s3), S3-S17.
- Roark, B., Mitchell, M., Hosom, J.-P., Hollingshead, K. et Kaye, J. (2011). Spoken language derived measures for detecting mild cognitive impairment. *IEEE transactions on audio, speech, and language processing*, 19(7), 2081-2090.
- Rodrigue, K., Kennedy, K., Devous, M., Rieck, J., Hebrank, A., Diaz-Arrastia, R., Mathews, D. et Park, D. J. N. (2012).  $\beta$ -Amyloid burden in healthy aging: regional distribution and cognitive consequences. *78(6)*, 387-395.
- Rogalski, E., Sridhar, J., Rader, B., Martersteck, A., Chen, K., Cobia, D., Thompson, C. K., Weintraub, S., Bigio, E. H. et Mesulam, M.-M. (2016). Aphasic variant of Alzheimer disease: clinical, anatomic, and genetic features. *Neurology*, 87(13), 1337-1343.
- Ruksenaite, J., Volkmer, A., Jiang, J., Johnson, J. C., Marshall, C. R., Warren, J. D., Hardy, C. J. J. C. N. et Reports, N. (2021). Primary Progressive Aphasia: Toward a Pathophysiological Synthesis. *21(3)*, 1-12.
- Sanborn, V., Ostrand, R., Ciesla, J. et Gunstad, J. J. A. N. A. (2020). Automated assessment of speech production and prediction of MCI in older adults. 1-8.
- Santos-Santos, M. A., Rabinovici, G. D., Iaccarino, L., Ayakta, N., Tammewar, G., Lobach, I., Henry, M. L., Hubbard, I., Mandelli, M. L. et Spinelli, E. (2018). Rates of amyloid imaging positivity in patients with primary progressive aphasia. *JAMA neurology*, 75(3), 342-352.
- Scott, G. G., Keitel, A., Becirspahic, M., Yao, B. et Sereno, S. C. (2019). The Glasgow Norms: Ratings of 5,500 words on nine scales. *Behavior Research Methods*, 51(3), 1258-1270.
- Shimada, M., Meguro, K., Yamazaki, H., Horikawa, A., Hayasaka, C., Yamaguchi, S., Yamaguchi, K., Katsuyama, N., Nakano, M. et Yamadori, A. (1998). Impaired verbal description ability assessed by the picture description task in Alzheimer's disease. *Archives of Gerontology and Geriatrics*, 27(1), 57-65.
- Ska, B. et Duong, A. (2005). Communication, discours et démence. *Psychologie & Neuropsychiatrie du Vieillissement*, 3(2), 125-133.
- Slegers, A., Cole, J., Joubert, S., Escudier, F., Seni, A., Bédirian, V., Charbonneau, S., Denault, C., Imbeault, H., Scherzer, P., Gagnon, J. et Rouleau, I. (2018). Normes québécoises pour une version abrégée de l'Échelle de dénomination de Boston à 30 items. *App. 2*, 95-108.

- Slegers, A., Filiou, R., Montembeault, M. et Brambati, S. (2018). Connected Speech Features from Picture Description in Alzheimer's Disease: A Systematic Review. *Journal of Alzheimer's disease: JAD*, 65(2), 519.
- Smith, S. R., Chenery, H. J. et Murdoch, B. E. (1989). Semantic abilities in dementia of the Alzheimer type. II. Grammatical semantics. *Brain and Language*, 36(4), 533-542.
- Smith, S. R., Murdoch, B. E. et Chenery, H. J. J. A. J. o. H. C. D. (1987). Language disorders associated with dementia of the Alzheimer type: a review. *15*(1), 49-70.
- Soares, I., Dias, J., Rocha, H., do Carmo Lopes, M. et Ferreira, B. (2016). Feature selection in small databases: a medical-case study. Dans. XIV Mediterranean Conference on Medical and Biological Engineering and Computing 2016.
- Taler, V. et Phillips, N. A. (2008, 2008/06/04). Language performance in Alzheimer's disease and mild cognitive impairment: A comparative review. *Journal of Clinical and Experimental Neuropsychology*, 30(5), 501-556. <https://doi.org/10.1080/13803390701550128>
- Tan, R. H., Kril, J. J., Yang, Y., Tom, N., Hodges, J. R., Villemagne, V. L., Rowe, C. C., Leyton, C. E., Kwok, J. B. et Ittner, L. M. (2017). Assessment of amyloid  $\beta$  in pathologically confirmed frontotemporal dementia syndromes. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 9, 10-20.
- Themistocleous, C., Eckerström, M. et Kokkinakis, D. J. P. o. (2020). Voice quality and speech fluency distinguish individuals with Mild Cognitive Impairment from Healthy Controls. *15*(7), e0236009.
- Themistocleous, C., Ficek, B., Webster, K., den Ouden, D.-B., Hillis, A. E. et Tsapkini, K. J. b. (2020). Automatic subtyping of individuals with Primary Progressive Aphasia.
- Themistocleous, C., Webster, K., Afthinos, A. et Tsapkini, K. J. A. J. o. S.-L. P. (2020). Part of speech production in patients with Primary Progressive Aphasia: An analysis based on natural language processing. 1-15.
- Tippett, D. C. (2020). Classification of primary progressive aphasia: challenges and complexities. *F1000Research*, 9.
- Van Paridon, J. et Thompson, B. J. B. r. m. (2021). subs2vec: Word embeddings from subtitles in 55 languages. *53*(2), 629-655.

- Vásquez-Correa, J. C., Rios-Urrego, C. D., Arias-Vergara, T., Schuster, M., Rusz, J., Nöth, E. et Orozco-Arroyave, J. R. J. P. R. L. (2021). Transfer learning helps to improve the accuracy to classify patients with different speech disorders in different languages.
- Vieira, R. T., Caixeta, L., Machado, S., Silva, A. C., Nardi, A. E., Arias-Carrión, O., Carta, M. G. J. C. p., CP, e. i. m. h. et EMH. (2013). Epidemiology of early-onset dementia: a review of the literature. *9*, 88.
- Villarejo-Galende, A., Llamas-Velasco, S., Gómez-Grande, A., Puertas-Martin, V., Contador, I., Sarandeses, P., González-Sánchez, M., Trincado, R., Pilkington, P. et Ruiz-Solis, S. (2017). Amyloid pet in primary progressive aphasia: case series and systematic review of the literature. *Journal of Neurology*, *264*(1), 121-130.
- Villemagne, V. L., Doré, V., Burnham, S. C., Masters, C. L. et Rowe, C. C. (2018). Imaging tau and amyloid- $\beta$  proteinopathies in Alzheimer disease and other conditions. *Nature Reviews Neurology*, *14*(4), 225-236.
- Villeneuve, S., Rabinovici, G. D., Cohn-Sheehy, B. I., Madison, C., Ayakta, N., Ghosh, P. M., La Joie, R., Arthur-Bentil, S. K., Vogel, J. W. et Marks, S. M. (2015). Existing Pittsburgh Compound-B positron emission tomography thresholds are too high: statistical and pathological evaluation. *Brain*, *138*(7), 2020-2033.
- Volkmer, A., Rogalski, E., Henry, M., Taylor-Rubin, C., Ruggero, L., Khayum, R., Kindell, J., Gorno-Tempini, M. L., Warren, J. D. et Rohrer, J. D. J. P. n. (2020). Speech and language therapy approaches to managing primary progressive aphasia. *20*(2), 154-161.
- Vyas, A., Aisopos, F., Vidal, M.-E., Garrard, P. et Paliouras, G. (2021). Identifying the Presence and Severity of Dementia by Applying Interpretable Machine Learning Techniques on Structured Clinical Records.
- Warriner, A. B., Kuperman, V. et Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191-1207.
- Wilson, S. M., Henry, M. L., Besbris, M., Ogar, J. M., Dronkers, N. F., Jarrold, W., Miller, B. L. et Gorno-Tempini, M. L. (2010). Connected speech production in three variants of primary progressive aphasia. *Brain*, *133*(7), 2069-2088.
- Witte, M. M., Foster, N. L., Fleisher, A. S., Williams, M. M., Quaid, K., Wasserman, M., Hunt, G., Roberts, J. S., Rabinovici, G. D. et Levenson, J. L. (2015). Clinical use of amyloid-

- positron emission tomography neuroimaging: practical and bioethical considerations. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 1(3), 358-367.
- Wolk, D. A., Klunk, W. E. J. C. n. et reports, n. (2009). Update on amyloid imaging: from healthy aging to Alzheimer's disease. 9(5), 345-352.
- Zhou, L., Fraser, K. C. et Rudzicz, F. (2016). Speech recognition in Alzheimer's disease and in its assessment. *Interspeech 2016*, 1948-1952.
- Zhu, Y. et Liang, X. J. a. p. a. (2020). Exploiting Fully Convolutional Network and Visualization Techniques on Spontaneous Speech for Dementia Detection.
- Zimmerer, V. C., Hardy, C. J., Eastman, J., Dutta, S., Varnet, L., Bond, R. L., Russell, L., Rohrer, J. D., Warren, J. D. et Varley, R. A. J. C. (2020). Automated profiling of spontaneous speech in primary progressive aphasia and behavioral-variant frontotemporal dementia: an approach based on usage-frequency. 133, 103-119.