

Vieillesse normale et maladie d'Alzheimer : analyse comparative de la narration semi-dirigée au niveau lexical

Hyeran Lee

Praxiling UMR5267 - CNRS & Université de Montpellier
hye-ran.lee@univ-monpt3.fr

Résumé L'objectif de cette étude est de comparer et d'analyser les discours oraux produits par 19 patients atteints de la maladie d'Alzheimer et 19 sujets contrôles appariés en âge, sexe, et niveau socioculturel. Différentes mesures lexicales – diversité du vocabulaire, catégories grammaticales, segments répétés, fréquence lexicale, disfluences – ont été réalisées pour caractériser le déficit linguistique observé dans la maladie d'Alzheimer. Les résultats montrent qu'il existe des convergences et des divergences entre les deux populations de sujets.

Abstract The aim of the present study is to compare and analyze the oral discourse produced by 19 patients diagnosed with Alzheimer's disease and 19 age-, sex-, socio-cultural level-matched healthy controls. Different lexical measures – vocabulary diversity, grammatical categories, repeated segments, lexical frequency, disfluencies – were used to characterize the linguistic decline in Alzheimer's disease. Results show that there are similarities and differences in language ability between these two groups of subjects.

Mots-clés : Maladie d'Alzheimer, Vieillesse cognitive, Analyse du discours oral, Analyse lexicale

Keywords: Alzheimer's disease, Cognitive Aging, Speech analysis, Lexical analysis

1 Introduction

La maladie d'Alzheimer (MA) est une maladie neurodégénérative entraînant la perte progressive et irréversible des fonctions cognitives (amnésie, aphasie, apraxie, agnosie). Elle est la forme de démence la plus fréquente chez les personnes âgées de plus de 65 ans (60% des cas de démences séniles). Aujourd'hui, 860.000 Français souffrent de la maladie d'Alzheimer et, avec le vieillissement de la population, on estime que ce chiffre sera multiplié par deux à horizon 2040 (Ferri et *al.*, 2005). C'est pourquoi la recherche visant à combattre cette pathologie constitue un véritable enjeu de santé publique.

Le déficit linguistique est décrit dès la première observation clinique de la maladie d'Alzheimer (Alzheimer, 1907). Aujourd'hui, la dégradation de la capacité linguistique est reconnue comme une des manifestations les plus précoces et les plus prédominantes de la pathologie (Taler et *al.*, 2008). Par ailleurs, de nombreuses études ont constaté une diminution des performances langagières lors de la phase pré-symptomatique de la maladie d'Alzheimer (Snowdon et *al.*, 1996, Powell et *al.*, 2006, Amieva et *al.*, 2008).

Cependant, malgré l'importance des troubles du langage chez les patients atteints de la maladie d'Alzheimer, la plupart des échelles d'évaluation utilisées pour poser le diagnostic probable de la maladie d'Alzheimer comportent peu des critères linguistiques (Ferria et *al.*, 2009). Les tests linguistiques standardisés mesurant la fluence verbale et/ ou les performances en dénomination d'images sont certes sensibles au déclin linguistique associé à la maladie d'Alzheimer (Small et *al.*, 2008), toutefois, il convient de souligner que ce genre de test fournit des informations limitées quant aux capacités de communication réelle des patients. L'analyse du discours permettrait de caractériser les phénomènes langagiers naturels pouvant servir d'évaluation quantitative et qualitative de l'altération linguistique propre à la maladie d'Alzheimer. De même elle peut contribuer de manière efficace au développement d'une prise en charge non médicamenteuse des patients en permettant, par exemple, d'améliorer les outils de remédiation orthophonique.

La neuropsychologie part du postulat que les troubles neuropsychologiques sont révélateurs de l'organisation sous-jacente de l'architecture de la cognition (Belin et *al.*, 2006). Le discours étant le produit de l'activité cognitive, il constitue un objet d'étude privilégié pour l'exploration des mécanismes cognitifs. Le langage est probablement la fonction cognitive la plus résistante au vieillissement (Mathey et *al.*, 2008), ainsi une étude comparative portant sur une analyse quantitative et qualitative des productions de patients Alzheimer et de sujets âgés sains est un champ d'investigation pertinent et prometteur pour dresser un tableau précis de la détérioration *vs.* la préservation des capacités linguistiques dans les démences de type Alzheimer.

Afin de mettre en perspective l'apport de l'analyse du discours à la recherche sur la maladie d'Alzheimer, nous nous intéressons ici tout particulièrement aux aspects lexicaux. La compétence lexicale est considérée comme un facteur déterminant de la performance verbale (Smith, 1941) ; or la dimension lexico-sémantique du langage est le domaine le plus vulnérable dans le cadre de la maladie d'Alzheimer (Kemper et *al.*, 2009) alors qu'elle est la plus résistante dans le vieillissement cognitif normal (Van der Linden et *al.*, 1994). La caractérisation des aspects lexicaux dans le discours de patients Alzheimer et dans celui des sujets âgés sains doit ainsi nous permettre de relever des variables pertinentes pour la description du déficit linguistique dans le vieillissement normal *vs.* pathologique.

2 Méthode

2.1 Sujets

Les données ont été recueillies auprès d'une population contrôle constituée de personnes âgées saines et de patients atteints de la maladie d'Alzheimer à différents stades de démence (léger et modéré).

Le groupe de contrôle est constitué de 19 sujets âgés ne présentant aucune pathologie cognitive connue – score au MMSE (Folstein et *al.*, 1975) égal à 30/30. Les patients (n= 19) ont été recrutés en collaboration avec le Centre Hospitalier de Coutances, le Centre gériatrio-psychiatrique de Montpellier, et les maisons de retraite de Montpellier et sa région accueillant des patients atteints de la maladie d'Alzheimer.

Tous les sujets sont de langue maternelle française. Les deux groupes de sujets sont appariés en âge, sexe et niveau socioculturel – grille de Poitrenaud (Kalafat et *al.*, 2003) –.

Cette étude est conforme au cadre juridique du CNRS et a été menée auprès de sujets volontaires. Un formulaire de consentement rappelant les objectifs de l'étude et les conditions expérimentales de l'enquête a fait l'objet de la signature du sujet, de leurs référents familiaux et/ou de leurs tuteurs légaux.

Le profil des sujets est récapitulé dans le tableau 1.

	Patients n=19			Contrôles n=19			p
	(F=11, M=8)			(F=11, M=8)			
	Moyen	Ecart-type	Etendue	Moyen	Ecart-type	Etendue	
Age	75.58	5.17	67-84	75.79	5.58	66-86	n.s
NSC	2.53	0.96	1-4	2.84	1.01	1-4	n.s
MMS	22.32	2.85	16-26	30.00	0.00	30-30	<0.0001

Tableau 1 : Profil des sujets

2.2 Procédure

Les données orales proviennent d'une série d'entretiens individuels semi-dirigés. Deux enquêteurs entraînés se sont entretenus avec les sujets de contrôle (SC) entre juillet et octobre 2009. Trois autres enquêteurs ont procédé à des entretiens auprès des patients atteints de la maladie d'Alzheimer (MA) entre décembre 2009 et février 2010.

Dans un premier temps, toutes les informations nécessaires à l'appariement du groupe contrôle et du groupe de patients ont été renseignées (i.e. niveau d'éducation, activité professionnelle, antécédents médicaux personnels et familiaux (dépression, alcoolisme...), traitements antérieurs et actuels, histoire de la maladie (date et condition du diagnostic, prises en charge éventuelle...), changement de comportement du sujet, retentissement des troubles sur les activités quotidiennes, etc.

Dans un second temps une batterie de pré-tests (MMSE, Gnosies visuelles du PEGV, DO 80, Empans, F-nart) ont été administrés. Enfin, nous avons demandé aux participants à l'étude d'évoquer oralement un souvenir personnel.

Tous les entretiens ont été enregistrés numériquement (44KHz). Ces discours oraux ont été transcrits manuellement avec le logiciel Transcriber avec une transcription orthographique standard de type GARS (Blanche-Benveniste, 1998).

Notre corpus comporte 13, 808 occurrences. Pour éliminer l'effet de la longueur du corpus sur nos résultats, nous avons équilibré la taille du corpus entre le groupe de patients (6901 occurrences) et de contrôle (6907 occurrences) et de chaque sujet (moyenne 363 occurrences).

2.3 Mesures linguistiques

Nous avons sélectionné certaines mesures lexicales pour examiner des patterns lexicaux de nos deux populations.

- **Diversité du vocabulaire**

La mesure de la richesse lexicale est utilisée dans divers domaines de recherche linguistique comme le développement langagier, l'apprentissage d'une langue étrangère, l'altération de la capacité linguistique dans la pathologie. Pour étudier si la maladie d'Alzheimer a un effet sur la productivité langagière, nous avons calculé l'indice de diversité du vocabulaire en utilisant le programme VOCD (McKee et al., 2000) implémenté dans le logiciel CLAN (MacWhinney, 1995). Ce programme permet de contourner l'inconvénient de la mesure classique de Type/Token Ratio, très sensible à la longueur de corpus, en calculant statistiquement la diversité de types utilisés par rapport aux occurrences produites sur l'échantillonnage aléatoire au cours d'une interaction.

- **Catégorie grammaticale**

La classe grammaticale dont relève les mots produits a été étudiée afin de déterminer si certaines catégories sont favorisées ou déficitaires selon la présence ou l'absence de la maladie. Pour cela, nous avons effectué le calcul de la proportion de la classe grammaticale (nom, pronom, verbe, adjectif, adverbe, déterminant, conjonction, préposition, interjection) de chaque texte à l'aide de logiciel *Cordial analyseur* version 14.0 (Synapse développement).

- **Segments répétés**

Les segments répétés, suites de formes identiques dont la fréquence est supérieure à 2 dans le texte, ont également été calculé avec le logiciel *Lexico3* version 3.45 (SYLED-CLA2T) dans le but d'étudier l'organisation et la structuration des mots.

- **Fréquence lexicale**

Nous avons comparé la fréquence lexicale pour chacun des deux groupes de sujets. Pour ce faire, nous avons employé la base de donnée *Lexique 3* (New et al., 2001), fréquence de mot à l'oral calculée, à l'aide de 9,474 sous-titres de films, représentant 50 millions de mots. Nous avons d'abord attribué le rang aux mots de haute fréquence à basse fréquence de cette base. Puis, nous avons lemmatisé notre corpus afin d'éviter les ambiguïtés (e.g. « je marche » : verbe ; « la marche » : nom). Enfin, nous avons comparé le taux d'utilisation de 800 lemmes oraux les plus fréquents de chaque texte.

- **Disfluen**

Nous avons considéré la répétition immédiate de mots, le mot fragmenté, la paraphrasie phonémique comme des indicateurs de disfluen. Pour savoir si la difficulté de réalisation des mots au cours de la production orale surgit à travers ces indicateurs, nous avons mesuré le pourcentage de ces disfluences sur le nombre total d'occurrences.

Nous avons utilisé le test non-paramétrique de Mann-Whitney U test pour l'analyse statistique.

3 Résultats

- **Diversité du vocabulaire**

L'indice de diversité du vocabulaire mesuré par VOCD montre que le score moyen du groupe de MA est plus faible par rapport à celui du groupe de SC, 81,63 et 86,04 respectivement. Cependant, cette différence n'est pas significative ($p=0,55$).

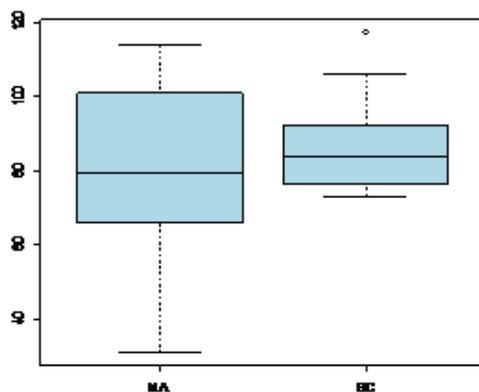


Figure 1 : Boîte à moustaches du score VOCD en fonction des groupes

- **Catégorie grammaticale**

Nous avons relevé par l'intermédiaire du logiciel *Cordial Analyseur* les différents types de catégories grammaticales utilisées par le panel. Seules les catégories pronom, adjectif, et préposition ont une différence significative entre les deux groupes. En effet, nous avons constaté une sur-utilisation de pronoms ($p < 0,05$) et une sous-utilisation d'adjectifs ($p < 0,001$) et de prépositions ($p < 0,05$) chez les patients Alzheimer. L'analyse plus fine de ces différents types grammaticaux a révélé que ces différences sont marquées par l'emploi plus faible de l'adjectif épithète ($p < 0,001$) et l'utilisation plus élevée du pronom personnel de 1^{ère} et 2^{ème} personne ($p < 0,05$) par les patients.

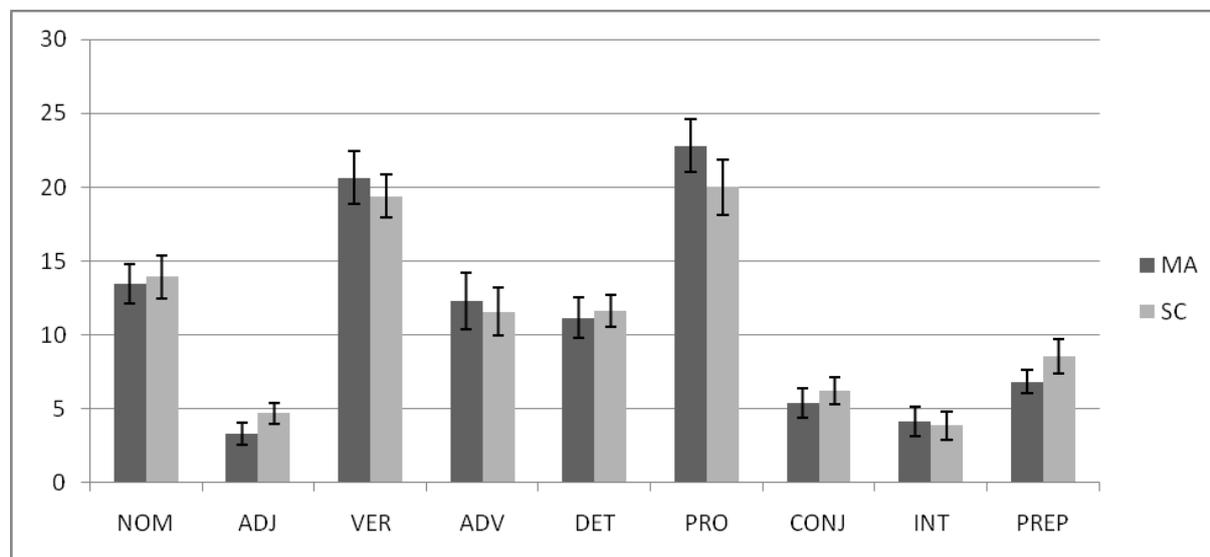


Figure 2 : Histogramme des catégories grammaticales en fonction des groupes

- **Segments répétés**

La figure 3 montre l'histogramme de pourcentage moyen de segments répétés sur le nombre total d'occurrence de chaque groupe. Même si des segments répétés sont plus présents dans le discours des patients (13,04 %) par rapport aux sujets sains (10,28 %), cette différence n'est pas significative statistiquement ($p = 0,21$). Cette tendance est la même pour les segments répétés de différentes longueurs (Longueur 2, 3, 4, 5, 6, 7).

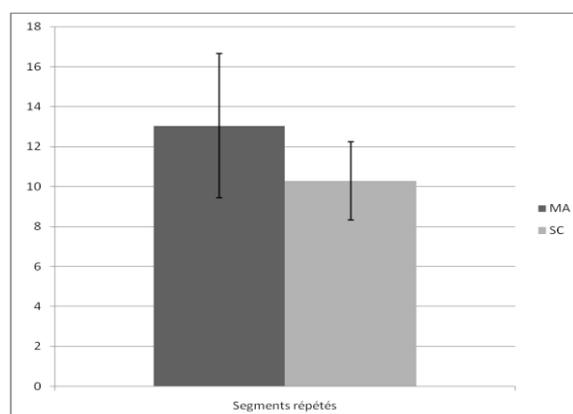


Figure 3 : Pourcentage de segments répétés

- **Fréquence lexicale**

La mesure de fréquence lexicale montre que les patients atteints de la maladie d'Alzheimer ont tendance à utiliser davantage de mots fréquents que les sujets de contrôle. Nous avons constaté que les sujets Alzheimer utilisent davantage les 200 lemmes les plus fréquents que les personnes âgées saines. Au delà des 500 lemmes les plus fréquents, les sujets MA recourent moins aux mots à faible fréquence que les sujets âgés sains pour construire leur discours ($p < 0,05$).

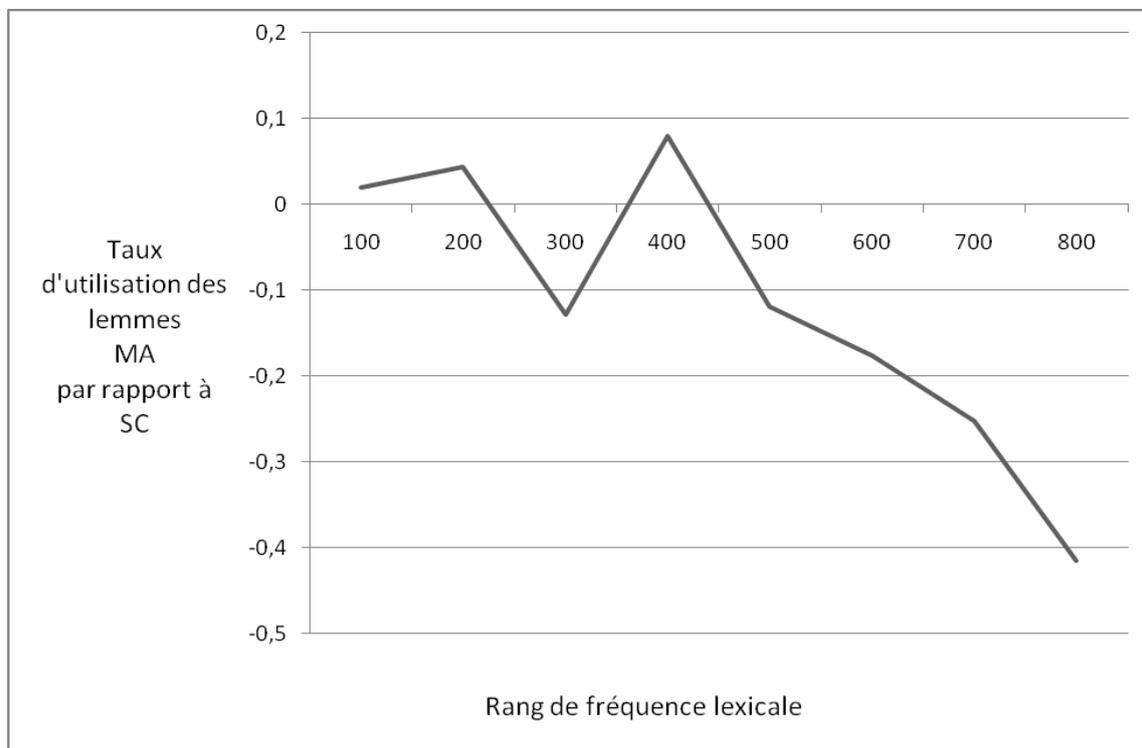


Figure 4 : Différence de taux d'utilisation des lemmes de MA par rapport à SC

- **Disfluence**

Le pourcentage moyen de répétition, de mot fragmenté, et de paraphrasie phonémique est plus faible chez les sujets de contrôle en comparaison à celui des patients : moyenne de 3,0 % pour les sujets sains et de 3,6 % pour les patients en répétition ; 0,49 % pour les sujets sains et 0,59 % pour les contrôles en mots fragmentés ; et 0,01 % pour le groupe de contrôle et 0,08 % pour le groupe de patients en paraphrasie phonémique. Cependant, leur faible fréquence dans le corpus étudié ne permet pas d'assurer que le phénomène est statistiquement significatif.

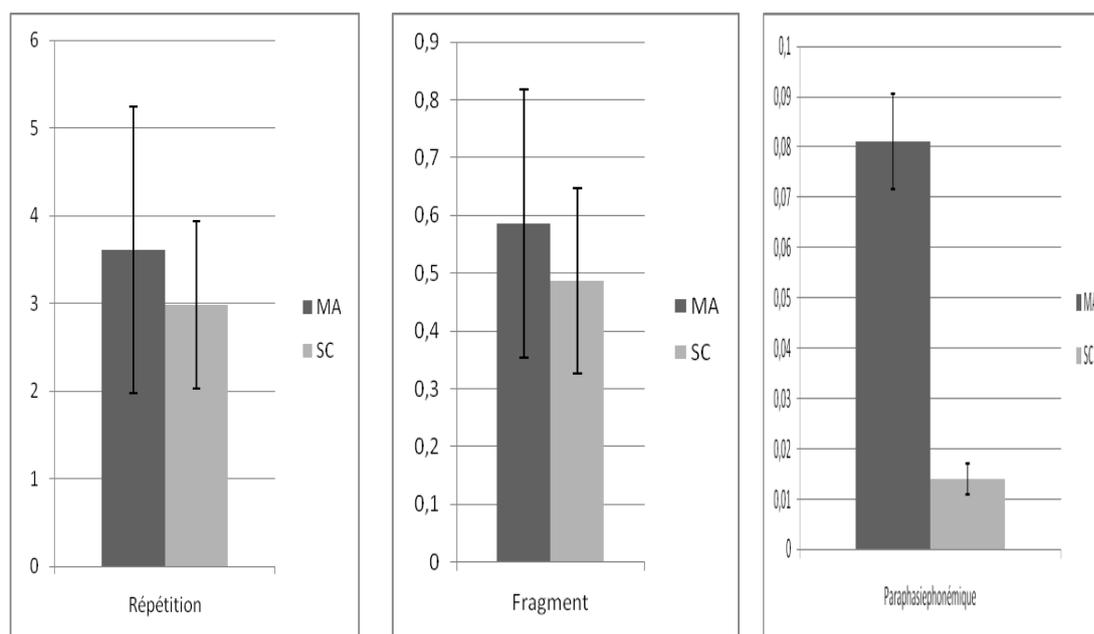


Figure 5 : Pourcentage de dysfluence dans le discours en fonction des groupes

4 Discussion

Nos résultats montrent quelques convergences et divergences de la capacité linguistique entre les deux groupes de sujets.

Contrairement aux études précédentes (Guetos et al., 2003, Groves-Wright et al., 2004), où l'altération de la capacité lexicale dans la maladie d'Alzheimer est systématiquement signalée, nos résultats montrent une légère différence de cette capacité entre le groupe de patients et le groupe de contrôles. Même si le score de VOCD est plus bas chez les patients Alzheimer, ils possèdent une diversité de vocabulaire comparable à celle des sujets sains. L'utilisation de segments répétés peut révéler une stratégie d'économie cognitive par l'emploi de structures préconstruites. Une tendance à l'utilisation de segments répétés est plus marquée chez les patients Alzheimer mais cette tendance n'est pas validée statistiquement. Enfin, la répétition immédiate, le mot fragmenté, et la paraphrasie phonémique considérés comme des disfluences sont plus présents dans le discours des patients. En revanche, nous avons observé la dynamique de réparation (e.g. en-*, enfant ; argre-*, arrière-grand-mère) ou la plasticité de substitution (e.g. ma mère qui tombait mal-* mourante) chez les patients comme chez les sujets sains.

Tout de même, l'effet de la fréquence lexicale est plus prononcé dans le discours produit par les patients Alzheimer. En effet, les patients recourent de moins en moins aux mots de basse fréquence par rapport aux sujets de contrôle et cette différence est marquée significativement. De nombreuses études ont montré que le lexique de haute fréquence est plus facile d'accès et plus rapidement récupérable par rapport au lexique de basse fréquence. L'utilisation préférentielle de lexique de fréquence élevée par les patients peut être interprétée comme une manifestation de la maladie d'une part et peut expliquer la disfluence contrôlée d'autre part. En effet, selon Garrad et al. (2005), la réduction de la ressource cognitive dans la maladie d'Alzheimer peut être reflétée par l'utilisation des mots fréquents. Aussi, l'étude d'Astell et al. (1996) a démontré qu'il y a moins d'occurrences de mot sur le bout de la langue sur les mots de haute fréquence chez les patients atteints de la maladie d'Alzheimer.

Vieillesse normale et maladie d'Alzheimer : analyse de la narration semi-dirigée au niveau lexical

Nous avons également observé que certaines catégories grammaticales sont sous-représentées dans le discours des patients, à savoir la préposition et l'adjectif ; alors que le pronom, est sur-employé par rapport aux discours produits par les sujets sains. En particulier, l'adjectif épithète est l'objet d'une sous-utilisation dans le discours des patients. Ainsi, le discours des patients Alzheimer n'en demeure pas moins correct mais il s'avère moins riche que celui des sujets sains d'un point de vue adjectival. Comme dans l'étude d'Almor et al. (1999), une sur-représentation des pronoms dans le discours des patients a été observée dans notre étude. Nous avons remarqué que cette augmentation est due aux pronoms personnels de la 1^{ère} et de la 2^{ème} personne. Il ne s'agit donc pas de pronoms anaphoriques mais de pronoms déictiques. Cela semble révéler chez les patients la volonté de maintenir le lien communicationnel. Cependant, une étude approfondie de ce point permettrait de mieux comprendre la nature de cette sur-utilisation des pronoms personnels.

5 Conclusion

La présente étude a examiné la narration semi-dirigée produite par les patients atteints de la maladie d'Alzheimer et des sujets de contrôle dans le but de déterminer les caractéristiques lexicales propres à la maladie d'Alzheimer dans un contexte discursif plus libre que les tests linguistiques standards.

Certes, l'analyse lexicale n'est pas suffisante pour décrire finement le profil linguistique des patients atteints de la maladie d'Alzheimer. Cependant, cette étude montre qu'une étude comparative entre le vieillissement normal et pathologique du langage permettrait une meilleure connaissance des troubles langagiers et des capacités linguistiques préservées dans les pathologies de type Alzheimer.

La présence d'un changement linguistique dans le discours des patients souffrant de la maladie d'Alzheimer peut être considérée comme un marqueur pertinent du déclin cognitif et pourrait contribuer à discriminer le vieillissement cognitif normal du vieillissement pathologique. Une description fine de la performance linguistique dans le vieillissement normal et pathologique permettrait de dresser une carte des détériorations et des préservations de la capacité langagière. Cela pourrait améliorer la gamme de diagnostic précoce de la maladie d'Alzheimer mais également offrir des pistes pour une prise en charge non médicamenteuse.

Références

ALMOR A., KEMPLER D., MACDONALD M., ANDERSEN E., TYLER L. (1999). Why do Alzheimer patients have difficulty with pronouns? Working memory, semantics, and reference in comprehension and production in Alzheimer's disease. *Brain and language* 67, 202-227.

ALZHEIMER A. (1907). Über eine eigenartige Erkrankung der Hirnrinde. *Allgemeine Zeitschrift für Psychiatrie und Psychisch-Gerichtlich Medizin* 64, 146-148.

AMIEVA H., LE GOFF M., MILLET X., ORGOGOZO J., BARBERGER-GATEAU P., JACQMIN-GADDA H., DARTIGUES J. (2008). Prodromal Alzheimer's disease : successive emergence of the clinical symptoms. *Annals of neurology* 64, 492-498.

ASTELL A., HARLEY T. (1996). Tip-of-the-Tongue stages and lexical access in dementia. *Brain and language* 54(2), 196-215.

BELIN C., ERGIS, A., MOREAUD O. (2006). *Actualités sur les démences : aspects cliniques et neuropsychologiques*. Marseille : Solal.

BLANCHE-BENVENISTE C. (1998). *Approches de la langue parlée en français*. Paris : Ophrys.

CUETOS F., MARTINEZ T., MARTINEZ C., IZURA C., ELLIS A. (2003). Lexical processing in Spanish patients with probable Alzheimer's disease. *Cognitive brain research* 17(3), 549-561.

FERRI C., PRINCE M., BRAYNE C., BRODATY H., FRATIGLIONI L., GRANGULI M., et al. (2005). Global prevalence of dementia : a Delphi consensus study. *Lancet* 366, 2112-2117.

FERRIA S., IHL R., ROBERT P. WINBLAD B., GATZ G., TENNIGKEIT F., GAUTHIER S. (2009). Severe Impairment Battery Language scale: A language-assessment tool for Alzheimer's disease patients. *Alzheimer's & Dementia* 5, 375-379.

FOLSTEIN M., FOLSTEIN S., MAC HUGH P. (1975). Mini Mental State : a practical method for grading the cognitive state of patients for the clinician. *J Psych Res* 12, 189-198.

GARRARD P., MALONEY L., HODGES J., PATTERSON K. (2005). The effects of very early Alzheimer's disease on the characteristics of writing by a renowned author. *Brain* 128(2), 250-260.

GROVES-WRIGHT K., NEILS-STRUNJAS J., BURNETT R., O'NEILL M. (2004). A comparison of verbal and written language in Alzheimer's disease. *Journal of communication disorders* 37, 109-130.

KALAFAT M., HUGONOT-DIENER L., POITRENAUD J. (2003). Etalonnage français du MMS version GRECO. *Revue neuropsychologie* 13, 209-236.

KEMPER S., ALTMANN L. (2009). Dementia and language. In *Encyclopedia of neuroscience*. Oxford : Academic Press. 409-414.

MACWHINNEY B. (1995). *The CHILDES project* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

MATHEY S., POSTAL V. (2008). Le langage. In *Neuropsychologie du vieillissement normal et pathologique*. Issy-les Moulineaux : Elsevier Masson. 79-102.

MCKEE G., MALVERN D., RICHARDS B. (2000). Measuring vocabulary diversity using dedicated software. *Literary and linguistic computing* 15 (3), 323-337.

NEW B., PALLIER C., FERRAND L., MATOS R. (2001). Une base de données lexicales du français contemporain sur Internet : LEXIQUE. *L'année psychologie* 101, 447-462. <http://www.lexique.org>

POWELL M., SMITH G., KNOPMAN D., PARISI J., BOEVE B., et al. (2006). Cognitive measures predict pathologic Alzheimer disease. *Arch. Neurol.* 63, 865-868.

SMALL J., SANDHU N. (2008). Episodic and semantic memory influences on picture naming in Alzheimer's disease. *Brain and language* 104 (1), 1-9.

Vieillesse normale et maladie d'Alzheimer : analyse de la narration semi-dirigée au niveau lexical

SMITH M. (1941). Measurement of the size of general English vocabulary through the elementary grades and high school. *Genetic psychological monographs* 24, 311-345.

SNOWDON D., KEMPER S., MORTIMER J., GREINER L., WEKSTEIN D., MARKESBERY W. (1996). Linguistic ability in early life and cognitive function and Alzheimer's disease in late life: findings from the Nun Study, *JAMA* 275, 528-532.

TALER V., PHILLIPS N. (2008). Language performance in Alzheimer's disease and mild cognitive impairment: A comparative review. *Journal of Clinical and Experimental Neuropsychology* 30, 501-556.

VAN DER LINDEN M., HUPET M. (1994). *Le vieillissement cognitif*. Paris: PUF.