

Contribution de l'analyse de corpus d'échanges conversationnels à la compréhension des troubles de la communication verbale dans la maladie d'Alzheimer

Contribution with the analysis of conversation corpus for understanding of conversation of verbal communication disorders in the Alzheimer's disease

Thi Mai Tran, Maïté Boye, Sandrine Mejias et Natalia Grabar



Édition électronique

URL : <http://journals.openedition.org/corpus/4203>

DOI : 10.4000/corpus.4203

ISSN : 1765-3126

Éditeur

Bases ; corpus et langage - UMR 6039

Référence électronique

Thi Mai Tran, Maïté Boye, Sandrine Mejias et Natalia Grabar, « Contribution de l'analyse de corpus d'échanges conversationnels à la compréhension des troubles de la communication verbale dans la maladie d'Alzheimer », *Corpus* [En ligne], 19 | 2019, mis en ligne le 01 avril 2019, consulté le 08 septembre 2020. URL : <http://journals.openedition.org/corpus/4203> ; DOI : <https://doi.org/10.4000/corpus.4203>

Ce document a été généré automatiquement le 8 septembre 2020.

© Tous droits réservés

Contribution de l'analyse de corpus d'échanges conversationnels à la compréhension des troubles de la communication verbale dans la maladie d'Alzheimer

Contribution with the analysis of conversation corpus for understanding of conversation of verbal communication disorders in the Alzheimer's disease

Thi Mai Tran, Maïté Boye, Sandrine Mejias et Natalia Grabar

Introduction

La maladie d'Alzheimer, pathologie cérébrale corticale liée au vieillissement, touche en France plus de 850 000 personnes. Elle est la première cause de troubles cognitifs majeurs (DSM-V, APA, 2013) de la personne âgée. Cette maladie neurodégénérative se caractérise par l'apparition insidieuse et progressive de troubles cognitifs, émotionnels et comportementaux conduisant à une perte de l'autonomie et un état de dépendance caractérisant la phase d'état de la maladie. Les troubles du langage font partie, avec les troubles de la mémoire et les troubles des fonctions exécutives, des premiers signes cliniques cognitifs de la maladie (Cummings *et al.* 1985, Derouesné 2006). Ils sont à l'origine de troubles de la communication perturbant progressivement les interactions sociales (Ska & Duong 2005 ; Khosravi 2011).

L'identification d'atteintes de la communication verbale contribue au diagnostic pluridisciplinaire de la maladie (Dubois *et al.* 2014 ; Sagot *et al.* 2015). Comme les autres déficits cognitifs, les troubles du langage sont le plus souvent évalués en situation de tests, au moyen d'épreuves ciblées (ex : dénomination, répétition, lecture à voix haute, discours narratif). Si ces épreuves permettent de mettre en évidence des déficits linguistiques, elles n'apportent pas d'informations concernant leurs répercussions

fonctionnelles sur la communication du patient dans ses interactions quotidiennes (Herbet *et al.* 2008 ; Lee 2011). Or, ces informations sont importantes pour évaluer le niveau d'autonomie du malade, ses besoins ainsi que les aides qui peuvent lui être apportées. Dans cette perspective, nous nous sommes intéressés à la situation de conversation, certes plus complexe à analyser en raison de la multitude de facteurs en jeu, mais plus représentative des échanges verbaux de la vie de tous les jours.

Après avoir rappelé les principales caractéristiques des troubles du langage et de la communication dans la maladie d'Alzheimer (désormais MA), nous présentons le protocole mis en place pour recueillir les productions langagières spontanées de dix personnes âgées (avec et sans MA). La sémiologie des troubles du langage dans ce contexte neurodégénératif se prêtant à une analyse automatisée, nous présenterons les outils utilisés ainsi que les résultats obtenus auprès de nos deux groupes de sujets. Nous discuterons enfin les pistes cliniques qu'apporte l'analyse de corpus conversationnels dans l'évaluation et la prise en charge des troubles neurodégénératifs acquis du langage chez la personne âgée.

1. Langage, communication et maladie d'Alzheimer

Si le langage est considéré comme une des fonctions cognitives résistant le mieux au vieillissement normal, il est décrit comme systématiquement atteint à la phase d'état de la MA (Cardebat *et al.* 1991). La maladie, décrite par Alzheimer en 1907, se caractérise par une atrophie cérébrale progressive en lien avec deux processus neurodégénératifs simultanés : la formation de plaques séniles et l'accumulation localisée de dégénérescences neurofibrillaires. Dans la phase préclinique, les lésions se constituent progressivement (initialement dans la région de l'hippocampe et des noyaux amygdaliens) sans donner de signes cliniques. Les débuts de la maladie se manifestent par l'apparition, pour les formes typiques, de troubles mnésiques, de modifications émotionnelles et d'une désorientation spatiale et temporelle. Avec la progression des lésions vers les aires cérébrales associatives postérieures, le tableau clinique s'enrichit graduellement d'autres troubles cognitifs pouvant toucher le langage, les fonctions exécutives, les gnosies et/ou les praxies, l'ensemble de ces déficits ayant des retentissements sur la vie sociale et relationnelle du sujet et conduisant à une réduction progressive de l'autonomie.

La spécificité des troubles du langage dans la MA est que les déficits linguistiques ne sont pas isolés et qu'ils s'intègrent dans une atteinte cognitive globale (mesurée notamment par le MMS (Flostein *et al.* 1975). Cela signifie que les troubles de la communication dans la MA ont une origine plurifactorielle et peuvent par exemple s'expliquer à la fois par des déficits linguistiques (ex. manques du mot, paraphasies), mnésiques (ex. difficulté à se rappeler ce qui a été dit), exécutifs (ex. difficulté à initier et organiser son discours) ou gnosiques (ex. problème de reconnaissance de l'interlocuteur). Il est noté également une hétérogénéité des manifestations cliniques initiales et des différences interindividuelles dans l'évolution de la maladie (rythme, sémiologie, sévérité), les troubles du langage pouvant être présents très précocement chez certains patients, survenir plus tardivement chez d'autres ou encore ne pas toujours être proportionnels aux autres troubles cognitifs (Eustache *et al.* 1993 ; Cardebat *et al.* 1995 ; Joannette *et al.* 2006). Toutefois, la dégradation progressive du langage suit une séquence relativement ordonnée que nous rappelons brièvement.

Le plus souvent discrets au début (recherches lexicales et pauses plus fréquentes que dans le discours ordinaire [Gayraud *et al.* 2011]) et souvent compensés (utilisation de périphrases et de termes généraux), les déficits linguistiques s'intensifient rendant progressivement les échanges problématiques (réduction de la fluence, appauvrissement du lexique et de la syntaxe, production de paraphrasies essentiellement sémantiques, présence de répétitions, de persévérations, d'incohérences et de digressions dans le discours). Ces troubles expressifs s'associent dès la phase modérée à des troubles de la compréhension d'abord des phrases complexes puis des mots mais aussi des intentions communicatives de l'interlocuteur pour aboutir, à la phase sévère, à une atteinte globale de tous les composants du langage et conduisant à une réduction majeure voire une absence de communication verbale aux stades ultimes. Au fil du temps, les énoncés référentiels, porteurs d'informations, vont diminuer au profit des énoncés modalisateurs (Nespoulous 1980), mentionnant parfois de façon explicite les difficultés linguistiques (Duong *et al.* 2003), et de routines langagières et formules toutes faites (Wray 2013), permettant de maintenir l'interaction de plus en plus appauvrie sémantiquement. Avec l'évolution de la maladie, la qualité de l'échange dépendra de l'investissement de l'interlocuteur et des aides communicationnelles qu'il pourra apporter au malade. C'est pourquoi il semble nécessaire de prendre en compte des situations de communication naturelle et plus particulièrement conversationnelles pour décrire le langage et la communication des personnes avec MA.

Le tableau 1 ci-dessous reprend les principales caractéristiques de la communication verbale au cours de l'évolution de la maladie et indique les possibilités conversationnelles à ses différents stades.

Tableau 1. Évolution de la communication verbale aux différents stades de la MA

Stades	Compétences linguistiques atteintes	Compétences linguistiques préservées	Échange conversationnel	Troubles cognitifs associés
Léger 26-24 < MMS < 20	Atteinte lexico-sémantique en production Atteinte discursive et pragmatique	Fluence Dimension phonético-phonologique Dimension morphosyntaxique	Possible avec petits troubles d'évocation lexicale et quelques paraphrasies Langage référentiel et modalisateur	Troubles de la mémoire Troubles des fonctions exécutives (initiative, inhibition, organisation, planification, programmation) Troubles de l'attention Troubles gestuels (apraxie) Troubles de la reconnaissance visuelle (agnosie visuelle, prosopagnosie) Troubles émotionnels et comportementaux Désorientation spatio-temporelle
	Handicap communicationnel discret			

Modéré 10 < MMS < 20	Diminution de la fluence Aggravation de l'atteinte lexicosémantique en production s'étendant à la compréhension Aggravation de l'atteinte discursive et pragmatique	Les aspects phonétophonologiques et morphosyntaxiques se dégradent très progressivement mais restent mieux préservés que les aspects lexicosémantiques	Échange de plus en plus difficile avec le soutien de l'interlocuteur (manque d'informativité, perte du fil de la conversation, problèmes d'adaptation au contexte et/ou à l'interlocuteur) Diminution du langage référentiel au profit du langage modalisateur
	Handicap communicationnel léger à modéré		
Sévère MMS < 10	Atteinte de tous les niveaux linguistiques (en production et en compréhension) Réduction de l'expression pouvant aller jusqu'au mutisme, atteinte majeure de l'informativité et de l'intelligibilité	Langage modalisateur et automatismes langagiers	Échange très difficile (supporté par l'interlocuteur et s'appuyant sur la communication non verbale) voire impossible
	Handicap communicationnel sévère		

2. Objectifs et hypothèses

Notre objectif est d'étudier la communication verbale des patients MA en situation de conversation. Notre étude vise à compléter la réflexion diagnostique et thérapeutique réalisée à partir des bilans de langage et des situations de tests en déterminant ce que peut apporter l'analyse de corpus conversationnel à l'identification de critères spécifiques en situation naturelle de communication.

Nous faisons l'hypothèse que (1) les troubles lexicosémantiques mis en évidence dans les épreuves cliniques se retrouvent dans les corpus de langage conversationnel et que (2) la situation de langage conversationnel est plus adaptée pour comprendre les répercussions fonctionnelles des troubles du langage sur la communication de la personne MA.

3. Sujets, matériel et méthode

Dix femmes âgées, de 83 à 102 ans (âge moyen = 90 ans \pm 6.42), de même niveau d'éducation (NSC = 1), vivant en EHPAD (Établissement d'Hébergement pour Personnes Agées Dépendantes) ont participé à cette étude. Toutes étaient décrites comme communicantes par le personnel de l'établissement et avaient une audition et une vue normales ou corrigées. Chacune avait bénéficié d'un bilan neuropsychologique dans la structure datant de moins de six mois.

Les cinq participantes témoins répondaient aux critères du GREFEX ([Roussel et Godefroy 2008] absence d'antécédents neurologiques ou psychiatriques pouvant être à l'origine de dysfonctionnements cognitifs) et avaient un MMS dans la norme (MMS moyen = 26/30 \pm 1) (Kalafat *et al.*, 2003). Les cinq autres participantes présentaient une maladie d'Alzheimer à un stade léger à modérément sévère (MMS moyen = 15/30 \pm 3.5) avec une compréhension simple satisfaisante (évaluée par la réalisation des ordres issus du Protocole Montréal-Toulouse d'examen de l'aphasie [Nespoulous *et al.* 1992]).

Les deux groupes de participantes sont appariés en âge ($t(8) = -.82$; $p = .435$) et présentent des scores au MMS différents ($t(8) = 6.77$; $p < .001$), les participantes présentant une MA ayant des résultats significativement inférieurs.

Le protocole, réalisé sur leur lieu de vie, dans un local calme, comprenait trois situations standardisées avec le même interlocuteur, familier des participantes :

- Une situation d'évaluation de la production lexicale avec une épreuve courte de dénomination d'images notée sur 36, issue de l'Exa-Dé (Bachy-Langedock 1988).
- Une situation de discours narratif oral (DNO) avec l'épreuve issue du Protocole Montréal-Toulouse d'examen de l'aphasie (Nespoulous *et al.* 1992). Les sujets étaient invités à raconter une scène imagée représentant le vol d'une banque. Cinq aspects étaient évalués sur une échelle de 1 à 5 : fluence, fluidité, production du langage (présence d'erreurs lexicales ou de transformations), élaboration du récit et informativité pour un score global de 25 points.
- Une situation d'échange conversationnel (tâche du protocole COMAZ¹ : échange à partir d'une thématique choisie par les participantes et avec des relances ciblées de la part de l'interlocuteur) dont l'analyse automatique, portant sur le contenu de l'échange et l'interaction verbale, sera détaillée ensuite.

Grâce aux deux premières situations, il a été permis de confirmer la présence de troubles lexico-sémantiques et discursifs chez les sujets MA alors que ces troubles étaient absents chez les sujets témoins (cf. Tableau 2 ci-dessous).

Tableau 2. Caractéristiques de la population et résultats aux épreuves de langage

Sujets témoins	Âge (ans)	MMS /30	Exa-Dé /36	DNO /25	Sujets MA	Âge (ans)	MMS /30	Exa-Dé /36	DNO /25
N1	83	26	36	25	MA1	89	20	16	11
N2	83	25	36	25	MA2	90	16	18	19
N3	84	25	34	24	MA3	90	12	29	15
N4	92	27	36	23	MA4	93	15	24	16
N5	102	27	36	25	MA5	99	11	15	20

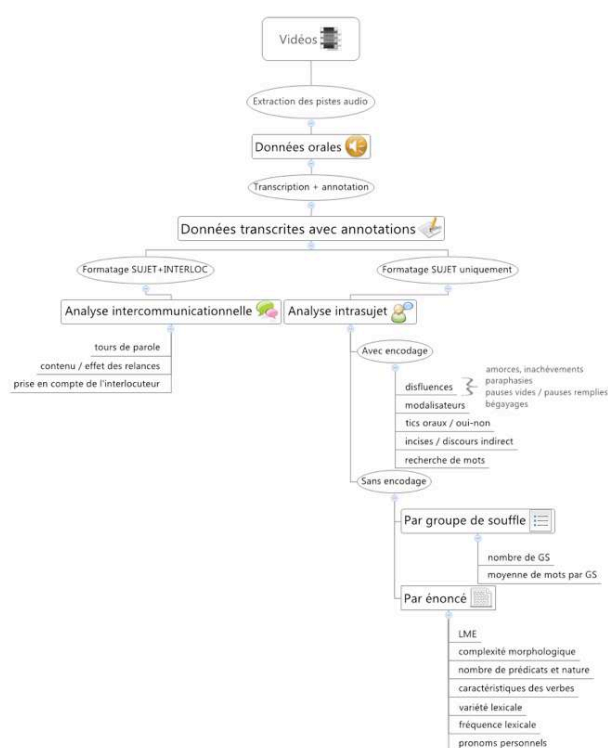
Ainsi, les profils dénominatif et narratif des sujets MA correspondaient aux descriptions habituelles des troubles décrits dans les premiers stades de la maladie (Croisile *et al.* 1996 ; Berrewaerts *et al.* 2003 ; Tran *et al.* 2012). En dénomination (Exa-Dé), les témoins avaient un score moyen de $35,5/36 \pm .89$ tandis que les performances moyennes des sujets MA correspondaient à un score moyen de $20,4/36 \pm 5.94$, significativement plus faible ($t(8) = 5.66$; $p \leq .001$). Chez les sujets MA, les erreurs étaient essentiellement sémantiques (ex. « seau » pour *arrosoir*) et visuelles (ex. « miroir » pour *loupe*) avec la production de dénominations génériques (ex. « oiseau » pour *perroquet*), de circonlocutions (ex. « pour les poissons » pour *aquarium*) et de modalisations (ex. « ah, je le connais mais j'arrive pas à le dire »). Dans le discours narratif oral (DNO), le score moyen des témoins ($24,4/25 \pm 0,85$) était significativement supérieur à celui des participantes MA ($16,2/25 \pm 3,56$; $t(8) = 4.99$; $p = .001$). Une participante MA a eu du mal à interpréter la scène imagée, deux autres ont rencontré des difficultés dans l'organisation de leur discours et deux ont produit un discours pauvre et descriptif comparativement aux productions des sujets témoins.

Ces résultats confirment également les caractéristiques du langage des sujets MA dans la première et la seconde partie d'évolution de la maladie (stade débutant à modéré) avec une expression construite syntaxiquement, fluente et intelligible (car correctement articulée et ne comprenant pas de transformations phonétiques ou phonologiques) pour laquelle un traitement automatique peut être envisagé.

Pour la situation d'échange conversationnel, deux photographies en noir et blanc (l'une évoquant un départ en train au début des congés payés et l'autre une balade à vélo d'un groupe d'amis à la même époque) étaient proposées aux participantes. La photographie préférée par le sujet servait d'amorce pour débiter une conversation autour de la thématique choisie, les sujets abordés au cours de l'échange pouvant ensuite évoluer librement. L'objectif était d'obtenir une production langagière spontanée et en quantité suffisante pour pouvoir procéder ensuite à une analyse automatique des corpus recueillis. Un peu plus tard, la discussion était redirigée vers la photographie non choisie afin d'amener le sujet à s'intéresser au point de vue de son interlocuteur et à l'inciter à produire un discours plus argumentatif et l'amener à donner son avis sur les idées exprimées par son interlocuteur. La conversation, filmée à l'aide de deux caméras sur pied, durait de 20 à 40 minutes. L'analyse a porté sur les dix dernières minutes de chaque étape (photo choisie par le sujet et photo choisie par l'interlocuteur) comportant plus de productions et d'informations. Ceci correspond au total à 200 minutes d'échanges.

Concernant la transcription et le traitement des corpus conversationnels, les données source (données orales extraites des vidéos) ont été transcrites afin d'obtenir un support textuel. La transcription a été réalisée grâce au logiciel *Transcriber* (Barras *et al.* 1998). Il permet d'annoter facilement le texte pour ajouter des informations complémentaires à la transcription brute correspondant notamment aux marques de l'oral, c'est-à-dire aux phénomènes de dysfluences spécifiques de la langue parlée (Blanche-Benveniste & Bilger 1999) comme les hésitations, pauses, répétitions, amorces ou inachèvements. Les traitements automatiques ont été effectués à partir des fichiers transcrits et formatés pour correspondre aux critères d'entrée des différents modules. La figure 1 reprend ces différentes étapes.

Figure 1. Étapes du traitement des corpus conversationnels



Le temps de parole des locuteurs, le nombre de tours de parole, la quantité de mots produits, le débit de parole ont été calculés manuellement. La lemmatisation, les analyses morphosyntaxiques et morphologiques ont été effectuées par les logiciels de Traitement Automatique des Langues : TreeTagger (Schmid 1994), FLEMM (Namer 2000) et DériF (Namer 2009), respectivement. L'exploitation de leurs résultats a permis de mesurer la diversité lexicale de chaque locuteur.

L'annexe 1 illustre les transcriptions réalisées par la présentation de deux extraits de conversations (l'une avec la participante témoin N5 et l'autre avec la participante MA5). Le tableau 3 ci-dessous récapitule les différentes variables retenues dans l'analyse.

Tableau 3. Données extraites des corpus conversationnels

Fichier source	Données calculées	Données observées
Fichier *.stm	Temps de parole du sujet : - seul - en simultané avec l'interlocuteur	- Chevauchements des temps de parole du sujet et de l'interlocuteur
Fichier *.typ formaté avec SUJET+INTERLOC et encodage	- Nombre de tours de parole - Nombre de pauses vides	- Contenu et effet des relances - Prise en compte de l'interlocuteur - Initiative
Fichier *.typ formaté SUJET seul sans encodage par groupe de souffle	- Nombre de groupes de souffle - Moyenne de mots par groupe de souffle	- Pause inter-tours

Fichier *.typ formaté SUJET seul sans encodage un énoncé par ligne	<ul style="list-style-type: none"> - LME (Longueur Moyenne d'Énoncé) - Diversité lexicale (nombre de lemmes/nombre total de mots) - Énoncés OUI/NON 	<ul style="list-style-type: none"> - Informativité - Richesse des propos
Fichier *.typ formaté SUJET seul avec encodage	<ul style="list-style-type: none"> - Nombre de dysfluences selon leur nature - Nombre d'incises - Nombre de passages en discours indirect 	<ul style="list-style-type: none"> - Analyse des dysfluences - Analyse des modalisateurs

Nous avons ainsi centré nos analyses sur les deux aspects suivants :

1. La dynamique de l'interaction verbale (tours de parole, temps de parole, chevauchements et débit).
2. Le contenu de l'échange à travers l'étude des données lexicales (nombre de mots produits, diversité lexicale), des données syntaxiques (LME, incises, discours rapporté, énoncés *oui-non*, pronoms personnels) et des dysfluences (amorces, pauses, répétitions, autocorrections, inachèvements).

4. Résultats

Le tableau 4 ci-dessous reprend les principales données analysées dans les corpus conversationnels de nos deux groupes de sujets ainsi que le score Z des différentes variables étudiées.

Tableau 4. Résultats des différentes variables en fonction du groupe de sujets

VARIABLES ANALYSÉES	TÉMOINS (N= 5)	MA (N = 5)
	Moyenne (ET) [min-max]	Moyenne (ET) [min-max]
1. Dynamique de l'interaction verbale		
a) Temps		
Temps de parole seul (sec.) Score Z	791.13 (177.81) [598.37-1038.74] -.46 (.109) [-.72-1.98]	640.00 (119.55) [464.04 - 788.01] -.46 (.73) [-1.54 - .44]
Temps de parole total (sec.) Score Z	181.80 (79.29) [42.00-230.00] -.56 (.94) [-.60-1.78]	673.74 (120,79) [476.18-791.22] -.56 (.76) [-1.81-.18]
b) Quantité		
Chevauchements (nombre) Score Z	1.01 (.75) [.09-2.12] .33 (1.17) [-1.12-2.06]	0.59 (.50) [.05 -1.26] -.33 (.78) [-1.18-.72]
Débit de parole (nombre de mots par minute) Score Z	175.29 (29.72) [132.43-200.99] .35 (1.14) [-1.30-1.33]	157.33(21.02) [132.17-184.23] -.35 (.81) [-1.31-.69]

Tours de parole (nombre de tours) Score Z	181.80 (79.29) [42-230] -.09 (1.09) [-2.01-.57]	194.80 (74.01) [73-255] .09 (1.02) [-1.59-.92]
2. Contenu de l'échange		
(a) Données lexicales		
Mots produits (nombre) Score Z	2404.40 (284.47) [2114-2804] .66 (.60) [.04-1.49]	1778.40 (431.27) [1144-2318] -.66 (.90) [-1.98-.47]
Diversité lexicale (nombre de substantifs/ verbes et adjectifs sur total des mots) Score Z	302.40 (39.42) [243-354] .81 (.48) [-.43-1.93]	168.60 (51.53) [98-221] -.81 (.62) [-1.66--.18]
(b) Données syntaxiques		
Longueur moyenne d'énoncés (LME) Score Z	7.26 (1.75) [5.18-9.79] .63(.90) [-.43-1.93]	4.79 (1.31) [2.94-6.12] -.63(.67) [-1.58-.05]
Incises Score Z	4.8 (3.70) [1-11] .56 (1.15) [-.62-2.48]	1.20 (1.30) [0-3] -.56 (.40) [-.93-.00]
Discours rapporté direct Score Z	14.20 (7.05) [6-25] .77 (.86) [-.23-2.09]	1.60 (1.52) [0-4] -.77 (.18) [-.96--.48]
Énoncés oui/non Score Z	52.4 (32.80) [6-89] -.37 (.66) [-1.30-.37]	88.8 (60.39) [5-170] -.37 (1.22) [-1.32-2]
Pronoms personnels (1 ^{er} personne du singulier) Score Z	67.20 5 (30.78) [22-100] -.10 (.89) [-1.41-.85]	74.20 (41.34) [45-146] .10 (1.20) [-.74-2.18]
(c) Dysfluences		
Répétitions, autocorrections, amorces, hésitations RAAH Score Z	63.4 (22.07) [42-95] .67 (.87) [-.18-1.91]	29.60 (15.76) [7-46] -.67 (.62) [-1.56--.02]
Pauses vides Score Z	31 (14.71) [17-48] -.44 (.38) [-.79-.00]	65 (49.91) [14-142] .44 (1.28) [-.87-2.41]
Pauses pleines (≥ 1 sec.) Score Z	28.00 (11.00) [19-45] .64 (.86) [-.06-1.97]	11.60 (8.82) [1-23] -.64 (.69) [-1.47-.25]
Inachèvement de mots Score Z	8.4 (6.88) [0-19] .49 (1.20) [-.98-2.34]	2.8 (2.59) [0-7] -.49 (.45) [-.98-.24]
Inachèvement de phrases Score Z	19.4 (5.13) [13-26] .03 (.51) [-.61-.69]	18.8 (14.18) [8-43] -0.3 (1.41) [-1.10-2.38]

L'ensemble des données a été analysé avec des modèles linéaires généralisés (GLM) et des comparaisons de moyennes pour échantillons indépendants à l'aide du logiciel SPSS version 25 (IBM Corp. Released 2017) afin de répondre à nos hypothèses et tester statistiquement quels éléments du corpus (mesurés grâce aux différentes variables composant « la dynamique de l'interaction verbale » et « le contenu de l'échange ») différaient selon les groupes de participants (témoins et MA).

Il est bien entendu important de garder à l'esprit que la taille de notre échantillon (5 sujets dans chaque groupe) est restreinte. Il convient donc de considérer les résultats et leur généralisation avec prudence.

Afin d'ajuster les séries de valeurs extraites des corpus conversationnels pour rendre les variables comparables, nous avons choisi de normaliser les données (en scores Z). Les scores aux différentes variables et leurs scores Z figurent dans le tableau 4 (moyennes, écarts-types, résultats minimum et maximum).

Dans un premier temps, nous nous sommes donc intéressés à la **dynamique de l'interaction verbale** selon (a) **des critères de temps**, mesurée au travers des scores Z des 2 variables « temps de parole seul et temps de parole total » pour nos 2 groupes de participants (témoins et MA). Les résultats ont montré qu'il n'y a pas de différence entre le temps de parole seul et le temps de parole total ($p = 1$), ni de différence entre les groupes ($p = .104$).

En ce qui concerne la **dynamique de l'interaction verbale** mesurée selon (b) **des critères de quantité**, les résultats ne montrent pas de différences entre les scores Z des 3 variables « chevauchements, tours de parole et débit » ($p = 1$) ni de différence entre nos 2 groupes de participants (témoins et MA ; $p = .523$).

Dans un deuxième temps, nous nous sommes intéressés au **contenu de l'échange** à travers l'étude :

(a) des données lexicales des participants témoins et MA.

Les données lexicales ont été mesurées au travers des scores Z des 2 variables (« mots produits, diversité lexicale ») pour nos 2 groupes de participants (témoins et MA). L'analyse met en évidence un effet de groupe ($F(1,8) = 14.08$; $p = .006$; $\eta^2 = .638$). On relève ainsi une différence de quantité et de diversité des mots produits chez les participants témoins ($M = .733 \pm .276$) par rapport aux participants MA ($M = -.733 \pm .276$). Les participants MA produisant globalement moins de données lexicales.

(b) des données syntaxiques des participants témoins et MA.

Dans ce cas précis, les données syntaxiques ont été mesurées par le biais des scores Z des 5 variables « LME, incises, discours rapporté direct, énoncés oui/non et pronoms personnels » pour nos 2 groupes de participants (témoins et MA). Cette analyse révèle un effet de groupe ($F(1,8) = 8.670$; $p = .019$; $\eta^2 = .520$). Les participants MA produisent globalement moins de données syntaxiques par rapport aux participants témoins ($M = -.298 \pm .143$; $M = .298 \pm .143$; pour les participants MA et témoins, respectivement). Une interaction est également présente entre données syntaxiques et groupe ($F(4,32) = 3.03$; $p = .032$; $\eta^2 = .275$). En effet, on observe une différence significative entre les deux groupes en faveur des participants du groupes témoins pour les variables « LME » ($t(8) = -2.53$; $p = .035$), « discours rapporté direct » ($t(8) = -3.907$; $p = .004$) et une différence marginale pour la variable « incises » ($t(8) = -2.051$; $p = .074$). On n'observe pas de différence significative entre les 2 groupes de participants pour les variables « énoncés oui/non » et « pronom personnel » ($ps \geq .27$).

(c) des dysfluences des participants témoins et MA.

Les dysfluences ont été mesurées par le biais des scores Z des 5 variables « RAAH, Pauses vides, Pauses pleines, Inachèvements de mots, Inachèvements de phrases » pour nos 2 groupes de participants (témoins et MA). Ici, l'analyse met en évidence une interaction entre dysfluences et groupe ($F(4,32) = 2.89$; $p = .038$; $\eta^2 = .266$) et un effet marginal du groupe ($F(1,8) = 4.252$; $p = .07$; $\eta^2 = .347$). La comparaison des

performances des 2 groupes pour chaque type de dysfluences met en évidence une différence significative pour la variable « RAAH » ($t(8) = -2,79$; $p = .024$) et « pauses pleines » ($t(8) = -2,60$; $p = .032$) en faveur des participants témoins. Aucune autre différence significative n'est relevée ($ps \geq .13$).

En résumé, l'analyse des 20 minutes de conversation par sujet révèle que les participants MA produisent moins de mots que les participants témoins, que leur débit de parole est plus lent et que le contenu de leurs messages est moins riche lexicalement et syntaxiquement. Si la dynamique de l'échange verbal reste globalement préservée, les participants MA sont moins engagés dans l'échange et s'expriment de manière plus succincte que leur interlocuteur avec moins de développements (incises, discours rapporté direct, précisions dans les propos).

5. Discussion

Les résultats quantitatifs et qualitatifs de nos analyses des corpus conversationnels montrent que la communication verbale des participants MA est quantitativement et qualitativement différente de celle des participants témoins (voir illustration en annexes) : les sujets MA parlent moins, produisent des énoncés plus courts et moins élaborés. Ils ont tendance à employer des mots fréquents et produisent moins de mots différents. Ils prennent moins d'initiatives dans l'échange, restent davantage centrés sur leur point de vue et prennent peu de distance par rapport à ce qu'ils disent. Ceci peut s'expliquer par les déficits linguistiques décrits dans la littérature (troubles lexicosémantiques, troubles discursifs et pragmatiques). Cela peut également s'expliquer par les autres troubles cognitifs associés à la MA : les déficits de la mémoire peuvent rendre compte d'une simplification de la syntaxe, de l'utilisation réduite d'incises et de discours rapporté nécessitant tous deux d'insérer une proposition dans le discours tout en sachant reprendre ensuite le fil de la conversation. Les déficits exécutifs, le manque d'élaboration des propos et leur relative pauvreté d'un point de vue sémantique contribuent également à ces résultats.

Une autre explication intéressante de ces résultats pourrait être la façon dont les sujets MA s'adaptent à leurs troubles et mobilisent leurs compétences préservées pour maintenir l'échange. L'hypothèse d'une stratégie communicative émerge de l'analyse des dysfluences. Compte-tenu de la présence de déficits linguistiques, mis en évidence dans les épreuves de dénomination et de discours narratif oral, une augmentation des dysfluences (amorces, pauses, autocorrections) aurait pu être attendue chez les participantes MA comme c'est le cas dans les tests linguistiques où les cliniciens recherchent des manifestations déficitaires (délais dans les réponses, erreurs, recherches lexicales, circonlocutions et modalisations). Il semble cependant que là où les sujets témoins, investis dans l'échange, précisent leur pensée et enrichissent leur discours, en cherchant des formulations plus précises (à l'origine de pauses pleines, d'amorces, d'hésitations), les sujets MA s'efforcent de maintenir l'interaction en prenant moins de risques (évitant de produire trop d'erreurs ou de réparations trop complexes à gérer) et en s'appuyant davantage sur les propositions du partenaire de communication.

À nos yeux, seule l'analyse de corpus conversationnels permet d'identifier ces comportements verbaux. Or ceux-ci peuvent tout à la fois, de manière complémentaire aux évaluations cliniques, participer au diagnostic et au projet thérapeutique. En effet,

une réduction des initiatives verbales, un appauvrissement du contenu des propos, une communication verbale recourant davantage à des routines conversationnelles et reposant en grande partie sur l'interlocuteur constitueraient des indicateurs cliniques des troubles du langage. De plus, ils informeraient de façon plus directe sur le retentissement des déficits linguistiques et cognitifs sur la communication et la participation sociale des patients. Ces comportements verbaux se traduiraient par une communication verbale réduite, appauvrie, moins dynamique et interactive sur le plan du contenu avec un recours plus important aux automatismes langagiers et une position plus en retrait du locuteur. D'un point de vue thérapeutique, la compréhension de ces comportements aiderait l'interlocuteur à fournir les aides adaptées au patient afin de favoriser et de soutenir son expression verbale (ex. Small et Perry 2012).

Conclusion

La prise en compte de critères conversationnels, dans le cadre d'un travail méticuleux sur corpus, ouvre de nouvelles perspectives pour l'étude de la communication verbale des MA. En effet, dans la littérature, les troubles du langage des sujets MA sont généralement recherchés à partir de manifestations de surface déficitaires comme les paraphasies, les pauses, les autocorrections, les répétitions et les énoncés inachevés et ceci généralement par l'intermédiaire d'épreuves de langage éloignées des situations de communication naturelle. Notre étude montre que l'absence de ces critères en situation de conversation n'implique pas l'absence de troubles mais indique que ceux-ci peuvent se manifester autrement dans la conversation, au travers d'un comportement verbal adaptatif plus centré sur l'échange et le maintien de l'interaction plutôt que sur la transmission d'informations précises. Ce travail fournit aux cliniciens d'autres pistes d'exploration en matière de diagnostic et de thérapie et confirme la complémentarité des deux approches clinique et écologique dans le suivi et la prise en charge des personnes souffrant de MA.

BIBLIOGRAPHIE

- American Psychiatric Association (APA) (2013). *Diagnostic and Statistical Manual of mental disorders*, 5^e édition. Washington DC : American Psychiatric Association.
- Bachy-Langedock N. (1988). *Batterie d'examen des troubles de la dénomination*. Bruxelles : Editest.
- Barra C., Geoffrois E., Wu Z. & Liberman M. (1998). « Transcriber : a free tool for segmenting, labeling and transcribing speech », *Conference on Language Resources and Evaluation (LREC)*, p. 1373-1376.
- Berrewaerts J., Hupet M. & Feyereisen P. (2003). « Langage et démence : examen des capacités pragmatiques dans la maladie d'Alzheimer », *Revue de Neuropsychologie* 13(2) : 165-207.
- Blanche-Benveniste C. & Bilger M. (1999). « Français parlé-oral spontané. Quelques réflexions », *Revue Française de Linguistique Appliquée*, Dossier "L'oral spontané" 4(2) : 21-30.

- Bombois S., Dubois B. & Pasquier F. (2015). « Approche clinique de la maladie d'Alzheimer », in Dubois B. et Mandon-Michon A. (dir.) *Démences*. Rueil-Malmaison : Doin, 148-156.
- Cardebat D., Aithamon B. & Puel M. (1995). « Les troubles du langage dans les démences de type Alzheimer », in Eustache F. & Agniel A. *Neuropsychologie clinique des démences : évaluations et prises en charge*. Marseille : Solal éditeur, 213-223.
- Chapman R & Miller J. (1984). *SALT : Systematic analysis of language transcripts*. Madison, WI : University of Wisconsin.
- 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 and language* 53 : 1-19.
- Cummings J.L., Benson D.F., Hill M.A. & Read S. (1985). « Aphasia in dementia of Alzheimer type », *Neurology* 35 : 394-397.
- Derouesné C. (2006). « Maladie d'Alzheimer : données épidémiologiques, neuropathologiques et cliniques », in Belin C., Ergis A.-M. & Moreaud O. (éd.) *Actualités sur les démences : aspects cliniques et neuropsychologiques*. Marseille : Solal, 25-34.
- Dubois B., Feldman H.H., Jacova C., Hampel H., Molinuevo J.L., Blennow K., DeKosky S.T., Gauthier S., Selkoe D., Bateman R., Cappa S., Crutch S., Engelborghs S., Frisoni G.B., Fox N.C., Galasko D., Habert M.O., Jicha G.A., Nordberg A., Pasquier F., Rabinovici G., Robert P., Rowe C., Salloway S., Sarazin M., Epelbaum S., de Souza L.C., Vellas B., Visser P.J., Schneider L., Stern Y., Scheltens P. & Cummings J.L. (2014). « Advancing research diagnostic criteria for Alzheimer's disease : the IWG-2 criteria », *The Lancet Neurology* 13 : 614-629.
- Duong A., Tardif A. & Ska B. (2003). « Discourse about discourse : What is it and how does it progress in Alzheimer's disease ? », *Brain and cognition* 53 : 177-180.
- Eustache F., Agniel A., Dary M., Viallard G., Puel M., Démonet J.-F., Rascol A. & Lechevallier B. (1993). « Sériation chronologique des symptômes comportementaux et instrumentaux dans les démences de type Alzheimer », *Revue de Neuropsychologie* 3 : 37-61.
- Folstein M.F., Folstein S.E. & McHugh P.R. (1975). « "Mini-mental State" : A practical method for grading the cognitive state of patients for the clinician », *Journal of Psychiatric Research* 12 : 189-198.
- Herber R., Hickin J., Howaerd D., Osborne F. & Best W. (2008). « Do picture naming tests provide a valid assessment of everyday functional lexical retrieval ? », *Aphasiology* 22 : 184-203.
- Joanette Y., Kahlaoui K., Champagne-Lavau M. & Ska B. (2006). « Troubles du langage et de la communication dans la maladie d'Alzheimer : description clinique et prise en charge », in Belin C., Ergis A.-M., Moreaud O. (éd.) *Actualités sur les démences : aspects cliniques et neuropsychologiques*. Marseille : Solal, 223-241.
- 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 », *Clinical Linguistics & Phonetics* 25(3) : 198-209.
- IBM Corp. Released 2017. IBM SPSS Statistics for Mac OS, Version 25.0. Armonk, NY : IBM Corp.
- Kalafat M., Hugonot-Diener L., Poitrenaud J. (2003). « Standardisation et étalonnage français du "Mini Mental State" (MMS) version GRECO », *Revue de Neuropsychologie* 13(2) : 209-36.
- Khosravi M. (2011). *La communication lors de la maladie d'Alzheimer et des troubles apparentés* (3^e éd.). Rueil-Malmaison : Doin.

Lee H. (2012). *Langage et maladie d'Alzheimer : analyse multidimensionnelle d'un discours pathologique*. Thèse de Doctorat en Sciences du Langage. Université de Montpellier III.

Namer F. (2000). « FLEMM : Un analyseur flexionnel du Français à base de règles », *TAL* 41(2) : 523-548.

Namer F. (2009). « Morphologie, Lexique et TAL : l'analyseur DériF », *TIC et Sciences cognitives*. London : Hermes Sciences Publishing.

Nespoulous J.-L. (1980). « De deux comportements verbaux de base : référentiel et modalisateur. De leur dissociation dans le discours aphasique », *Cahiers de psychologie* 23 : 195-210.

Nespoulous J.-L., Lecours A.R., Lafond D., Lemay A., Puel M., Joannette Y., Cot F. & Rascol A. (1992). *Protocole Montréal-Toulouse d'examen linguistique de l'aphasie*. Isbergues : Ortho-Editions.

Roussel M. & Godefroy O. (2008). « La batterie GREFEX : données normatives », in Godefroy O. et GREFEX, *Fonctions exécutives et pathologies neurologiques*. Marseille : Solal, 231-252.

Schmid E. (1995). « Probabilistic Part-of-Speech Tagging Using Decision Trees », *Proceedings of International Conference on New Methods in Language Processing*.

Sagot C., Tran T.M. & Pariente J. (2012). « Développement d'une batterie francophone pour l'évaluation des troubles du langage dans les maladies neurodégénératives : quels enseignements de 10 ans de recherche sur les aphasies primaires progressives ? », *Revue Française de Linguistique Appliquée* XVII : 117-133.

Small J. & Perry J. A. (2012). « Training family care partners to communicate effectively with persons with Alzheimer's disease : The TRACED program », *Canadian Journal of Speech Language Pathology and Audiology* 36(4) : 332-350.

Ska B. & Duong A. (2005). « Communication, discours et démence », *Psychologie et NeuroPsychiatrie du Vieillessement* 3(2) : 125-133.

Tran T.M., Dasse P., Letellier L., Lubjinkowic C., Thery J., Mackowiak M.-A. (2012). « Les troubles du langage inauguraux et démence : étude des troubles lexicaux auprès de 28 patients au stade débutant de la maladie d'Alzheimer », in Neveu F., Muni Toke V., Blumenthal P., Klinger T., Ligas P., Prévost S., Testond-Bonnard (éd.), *3^e Congrès Mondial de Linguistique Française- CMLF 2012*. Paris, Institut de Linguistique française, 1659-1676.

Wray A. (2013). « Formulaic language », *Language Teaching* 46 : 316-334

ANNEXES

Extraits de corpus

a) Conventions de transcription

<<SUJET>> : le sujet participant parle

<<INTERLOC>> : l'interlocuteur parle

[] : indique les informations sur les spécificités de la prononciation ou les bruits, dont on peut perdre la trace lors de la transcription.

Explication des cas qui apparaissent dans les extraits proposés :

[p] : une pause égale ou supérieure à 1 seconde

[b] : présence d'un bruit lors de la locution

[rire] : présence du rire de la part du locuteur

[pron=pi] : prononciation inintelligible faible

[pron=euh] : dysfluence avec « euh »

enfin+[lex=oral] : production lexicale spécifique à l'oral qui porte sur le mot qui précède (ici, « enfin »)

a+[pron=disflu:autocorrection] : présence d'une autocorrection du mot qui précède (ici, « a »)

pas+[pron=disflu:répétition] : présence d'une répétition du mot qui précède (ici, « pas »)

b) Extrait du corpus du sujet N5

<<SUJET>>

Alors donc [pron=euh].

[pron=pi]

Mais enfin+[lex=oral] bon+[lex=oral] [pron=euh] comme je vous dis

on n'a pas eu de voiture pendant longtemps.

[p]

Il y a+[lex=oral] quand on est venu à Ronchin

ben ma fille a+[pron=disflu:autocorrection] avait eu son bac et puis elle était
secrétaire de direction d'un laboratoire à Haubourdin.

Alors elle devait prendre le train tous les matins

[p]

On a décidé

[b]

de+[pron=disflu:autocorrection]

d'acheter une voiture.

Alors c'est elle qui a pris les leçons et [pron=pi] mon mari ne conduisait pas.

Et alors pour s'habituer

avec un monsieur de Ronchin qui allait dans le même établissement qu'elle
ils faisaient la route en voiture.

Il conduisait très bien.

Et alors donc après c'est mon mari qui a conduit.

Mais

avec lui j'avais peur parce qu'il conduisait vite.

<<INTERLOC>>

Ah.

<<SUJET>>

Il conduisait vite.

Et puis je trouvais qu'il ne conduisait pas+[pron=disflu:répétition] pas très bien.

Et alors une fois

on était parti en vacances

et alors partout où on passait

ça sentait le hareng fumé.

Vous savez quand on fait du hareng. [pron=pi]

Ah ça sent le hareng.

Ca sent le hareng fumé.

Alors il avait oublié de desserrer son frein à main.

<<INTERLOC>>

Ah mon Dieu.

[p]

<<SUJET>>

Alors

<<INTERLOC>>

Ah oui ça sentait bien le+[pron=disflu:autocorrection] ça sentait le crâmé.

[rire]

<<SUJET>>

Oh lalalala+[lex=oral]

<<INTERLOC>>

[rire]

c) Extrait du corpus du sujet MA5

Et vous, vous partiez en vacances J ?

<<SUJET>>

Oh il y a longtemps que je n'ai plus été en vacances.

<<INTERLOC>>

Non.

<<SUJET>>

Oui.

[p]

<<INTERLOC>>

Vous partiez où ?

<<SUJET>>

On partait oh jamais loin hein. Je travaillais tous les+[pron=disflu:répétition] tous les jours moi hein

avec mes journaux.

Tous les jours je travaillais alors.

On n'allait pas loin vous savez.

<<INTERLOC>>

Vous faisiez quoi comme travail ?

Avec les journaux vous faisiez quoi ?

<<SUJET>>

Ben+[lex=oral] je les distribuais.

<<INTERLOC>>

Ah.

Tous les jours tous les jours.

<<SUJET>>

Tous les jours.

[p]

Fallait être là hein par tous les temps [pron=euh] de l'année hein.

Et on se faisait attraper quand on était en retard encore.

<<INTERLOC>>

Ah bah ça.

<<SUJET>>

[rire]

C'est vrai hein.

[p]

<<INTERLOC>>

Ca n'a pas changé ça.

<<SUJET>>

Ah non.

Ben non les gens ils sont à leur maison.

Il pleut ben il pleut.

Mais celui qui est dehors aussi.

<<INTERLOC>>

C'est sûr.

<<SUJET>>

Qui est obligé pour son travail.

<<INTERLOC>>

Oui.

[p]

<<SUJET>>

Oh ben ça va, il ne faut pas se plaindre.

<<INTERLOC>>

Non.

<<SUJET>>

Voilà.

<<INTERLOC>>

Vous avez fait des kilomètres et des kilomètres.

<<SUJET>>

Hein ?

<<INTERLOC>>

Vous avez fait des kilomètres et des kilomètres.

<<SUJET>>

Ah oui, c'est vrai, c'est vrai, c'est vrai.

[rire]

<<INTERLOC>>

C'est pour ça que vous avez de belles jambes.

<<SUJET>>

Ah oui oui oui.

NOTES

1. Communication verbale et non verbale dans la Maladie d'Alzheimer : approche intégrative et fonctionnelle. Projet de recherche (COMAZ) sous la responsabilité de Loris Schiaratura (Université de Lille) et financé par CPER-MESHS : <http://www.meshs.fr/page.php?r=14&id=1387&lang=fr>.

RÉSUMÉS

Les troubles du langage font partie des troubles cognitifs présents dans la maladie d'Alzheimer. Ils sont le plus souvent étudiés dans des tâches cliniques ciblées, éloignées des situations de communication naturelle. Leur but est de mettre en évidence les déficits linguistiques comme par exemple les troubles lexico-sémantiques présents dès le début de l'évolution et spécifiques de la maladie. Nous avons choisi de compléter les épreuves classiques d'évaluation du langage par l'étude d'une situation d'échange conversationnel en essayant de dégager de nouveaux critères cliniques fonctionnels. Notre étude porte sur l'analyse automatique de corpus conversationnels recueillis auprès de 10 personnes âgées (moyenne d'âge = 90 ans) dont 5 présentant une maladie d'Alzheimer à un stade débutant à modéré (MMS moyen = 15/30). Nos analyses montrent des différences quantitatives et qualitatives entre les deux populations étudiées et proposent de s'intéresser davantage à l'interaction verbale pour mieux décrire le retentissement des troubles du langage sur la communication verbale des sujets avec maladie d'Alzheimer.

Language pathologies are part of cognitive disorders specific to the Alzheimer disease. Most often, they are studied through precise clinical tasks, which are very different from situations of natural communication. Their purpose is to find evidence of linguistic deficits such as lexical and semantic troubles occurring at the early stages and specific of the disorder. We propose to enrich the classical tests used for the evaluation of language capacity with the study of conversational exchanges thanks to which we expect to find new clinical functional criteria. Our study addresses an automatic analysis of a conversational corpus collected from 10 elderly participants (mean age = 90 years) among whom 5 participants have the Alzheimer disease at early to moderate stages (mean MMS = 15/30). Our analysis shows the existence of quantitative and qualitative differences between the two studied populations and proposes to take advantage of verbal interactions in order to better describe the impact of language troubles on verbal communication of people with the Alzheimer disease.

INDEX

Mots-clés : Communication verbale, Conversation, Maladie d'Alzheimer, TAL

Keywords : Verbal communication, Conversation, Alzheimer disease, NLP

AUTEURS

THI MAI TRAN

Département d'orthophonie - Faculté de Médecine, Université de Lille, Lille France
CNRS UMR 8163 STL - Savoirs Textes Langage, Université de Lille, Lille, France

MAÏTÉ BOYE

Département d'orthophonie - Faculté de Médecine, Université de Lille, Lille France

SANDRINE MEJIAS

Département d'orthophonie - Faculté de Médecine, Université de Lille, Lille France
CNRS, CHU Lille, UMR 9193 - SCALab - Sciences Cognitives et Sciences Affectives, Université de Lille, Lille, France

NATALIA GRABAR

CNRS UMR 8163 STL - Savoirs Textes Langage, Université de Lille, Lille, France